A Signal Perturbation Free Whitening-Rotation-Based Semiblind Approach for MIMO Channel Estimation

Feng Wan, Wei-Ping Zhu, *Senior Member, IEEE*, and M. N. S. Swamy, *Fellow, IEEE*

***Abstract*—It was shown in our previous work that, in the noise-free case, the whitening-rotation (WR)-based MIMO channel estimation algorithm is subject to a signal perturbation error, justifying that the WR-based method is efficient only in the low signal-to-noise ratio (SNR) case. In this paper, a very efficient signal-perturbation-free WR-based approach is proposed for semiblind channel estimation of MIMO systems. A novel transmit scheme is developed based on the eigenvalue decomposition of the correlation matrix of the transmitted signal. The new scheme is to send a small volume of data bearing the information of the correlation matrix to the receiver for the cancellation of the signal perturbation error so as to improve the performance of the WR-based method in the case of high SNRs. Then, a perturbation analysis of the proposed WR-based semiblind method with the new transmit scheme is conducted, leading to a closed-form expression for the mean square error (MSE) of the channel estimate. Computer simulations show that the proposed approach significantly outperforms the original WR-based method as well as some other channel estimation methods for all SNR levels.**

***Index Terms*—Channel estimation, MIMO system, perturbation analysis, semiblind, subspace.**

I. Introduction

The MIMO (multiple-input-multiple-output) technique has been considered as one of the key technologies for the development of the next-generation wireless communication systems. With multiple transmit and multiple receive antennas, MIMO systems can provide either a diversity gain to combat signal fading or a capacity gain, called spatial multiplexing, to make an efficient use of the channel resources [1], [2]. It means that, with MIMO techniques, higher data rate and better performance can be achieved without increasing the total transmission power or the bandwidth. On the other hand, the performance of MIMO systems depends largely upon the availability of the knowledge of the channel. Thus, an accurate estimation of the wireless channel is of crucial importance to MIMO systems [3].

Based on known pilots, the MIMO channel can be estimated by employing different kinds of training-based algorithms such as the least square (LS), the maximum likelihood (ML), the maximum a posteriori (MAP) and the minimum mean square error (MMSE) algorithms [3]. In contrast to training-based methods, blind channel estimation algorithms like those proposed in [4]–[8], can achieve a better spectral efficiency by use of the second-order statistics, correlative coding or other properties of user data. By combining the training-based and blind algorithms, semiblind channel estimation techniques can potentially enhance the quality of MIMO channel estimation [9]–[14]. With a small number of training symbols, problems such as ambiguities and mis-convergence of the blind methods can be solved by semiblind techniques. On the other hand, the use of the available information data yields an improved accuracy of the channel estimation.

More recently, a whitening-rotation (WR)-based semiblind algorithm has been proposed for frequency-flat MIMO channel estimation [10], [15]–[18]. The idea of this algorithm was first briefly presented in [10] as a simplified version of a semiblind algorithm for the estimation of frequency-selective MIMO channels. It was then fully disclosed in [15]–[17], in which a training-based constrained ML method was developed for the estimation of an ambiguity matrix for general MIMO systems. This method was later further extended for MIMO systems with maximum ratio transmission in [18]. The WR-based semiblind algorithm consists of two steps: 1) estimation of a whitening matrix utilizing information data; and 2) estimation of a unitary rotation matrix using pilots. The Cramer-Rao bound (CRB) of this semiblind technique shows that it can achieve a better channel estimation performance than the conventional LS method, when the number of receive antennas is greater than or equal to the number of transmit antennas. However, the perturbation analysis of the WR-based channel estimation method conducted in our previous work [19] shows that, in the noise-free case, the blind part of the WR-based method is subject to a signal perturbation error, justifying that the WR-based method is efficient only in the case of low signal-to-noise ratios (SNRs).

A nulling-based semiblind MIMO channel estimation approach, which can achieve a better performance in moderate to high SNR cases, was developed in [10], [11], and [20]. Instead of estimating the whitening matrix, this method uses the information data to obtain a blind constraint for the channel matrix, which is then combined with a training-based LS cost function so as to produce a semiblind solution for the MIMO channel response. This method can be considered as a modified LS solution involving a weighted blind constraint. It has been shown in [20] that the semiblind method provides a better channel estimation performance over the pure training-based LS method. In our previous work [19], the superiority of this

Manuscript received July 25, 2008; accepted February 07, 2009. First published March 31, 2009; current version published July 15, 2009. The associate editor coordinating the review of this paper and approving it for publication was Dr. Athanasios P. Liavas. This work was supported partly by the Natural Sciences and Engineering Research Council (NSERC) of Canada, and partly by the Regroupement Stratégique en Microsystèmes du Québec (ReSMIQ).

The authors are with the Center for Signal Processing and Communications, Department of Electrical and Computer Engineering, Concordia University, Montreal, Quebec, Canada H3G 1M8 (e-mail: f_wan@ece.concordia.ca; weiping@ece.concordia.ca; swamy@ece.concordia.ca).

Digital Object Identifier 10.1109/TSP.2009.2020021

1053-587X/$25.00 © 2009 IEEE

approach has been theoretically proved, and an appealing scheme of determining the weighting factor employed to trade off the least square and the blind criteria has been proposed.

In this paper, we propose a novel signal-perturbation-free transmit scheme to improve the performance of the WR-based method in the moderate to high SNR cases. By utilizing the eigenvalue decomposition (EVD) of the transmit signal perturbation matrix, a very efficient transmit scheme is designed for the elimination of the signal perturbation error in the receiver, leading to a signal-perturbation-free WR-based semiblind approach. It is shown that the new approach provides a much better performance than the original WR-based method as well as the nulling-based method does for all SNR cases.

In the second part of this paper, we apply the perturbation theory [9], [21]–[29] to the analysis of the new WR-based semiblind method with the proposed transmit scheme. The first-order perturbation theory is employed to analyze the first step of the WR-based method, leading to an expression of the perturbation term of the whitening matrix. For the second step of the WR-based method, a perturbation analysis of the singular value decomposition (SVD) of a square matrix involving only the signal subspace needs to be conducted. However, the existing perturbation theory like the one proposed in [9], [21]–[29] deals with the case where the noise subspace also exists. In this paper, a novel perturbation analysis theory is developed for the SVD of the square matrix without the noise subspace, leading to an expression of the perturbation term of the rotation matrix. By utilizing the perturbation terms of the whitening matrix and the rotation matrix, a closed-form expression for the MSE of the new WR-based method with the proposed transmit scheme is then derived.

The rest of the paper is organized as follows. Section II gives a brief review of the original WR-based semiblind channel estimation method and justifies why the WR-based method is efficient only in the case of low SNRs. Section III presents a novel signal-perturbation-free transmit scheme for the WR-based method that can completely eliminate the signal perturbation error at the receiver in the noise-free case. Section IV provides a perturbation analysis of the new WR-based approach with the proposed transmit scheme and the derivation of a closed-form expression for the MSE of the proposed method. Section V comprises a number of experiments validating the significant advantages provided by the new WR-based approach over the original WR-based method as well as the nulling-based method. Finally, Section VI highlights some of the distinct features of the proposed approach.

Throughout the paper, we adopt the following notations:

$\dagger$	Pseudoinverse;
$\otimes$	Kronecker product;
T	transpose;
H	complex conjugate transpose;
$*$	linear convolution;
$\circledast$	circular convolution;
$\|\| \|\|_F$	Frobenius norm;
$\circ$	Hadamard product;
$\delta(\,)$	Delta function;
$\mathrm{diag}()$	a stacking of the diagonal elements of the involved matrix into a vector;
$\mathrm{vec}()$	a stacking of the columns of the involved matrix into a vector, which has the following properties: $\mathrm{vec}[\mathbf{ABC}] = (\mathbf{C}^T \otimes \mathbf{A})\mathrm{vec}(\mathbf{B}))$, and $\mathrm{vec}[\mathbf{AB}] = (\mathbf{I}^T \otimes \mathbf{A})\mathrm{vec}(\mathbf{B}) = (\mathbf{B}^T \otimes \mathbf{I})\mathrm{vec}(\mathbf{A})$.

II. The Whitening-Rotation-Based Semiblind MIMO Channel Estimation

Consider a spatial-multiplexed MIMO system with N_T transmit and $N_R (\geq N_T)$ receive antennas. Suppose that the frequency-flat fading MIMO channel is characterized by an $N_R \times N_T$ matrix $\mathbf{H}$, whose (i_R, i_T)th element h_{i_T, i_R} represents the channel response from the i_Tth transmit antenna to the i_Rth receive antenna. Given the transmitted signal vector $\mathbf{x}(n) \triangleq [x_1(n), \cdots, x_{N_T}(n)]^T$ whose elements are independent identically distributed (i.i.d.) Gaussian random variables with zero mean and unit variance $\sigma_x^2 = 1$, the received signal vector $\mathbf{y}(n) \triangleq [y_1(n), \cdots, y_{N_R}(n)]^T$ can be written as

$$\mathbf{y}(n) = \mathbf{H}\mathbf{x}(n) + \mathbf{v}(n) \tag{1}$$

where the noise vector $\mathbf{v}(n) \triangleq [v_1(n), \cdots, v_{N_R}(n)]^T$ is spatio-temporally uncorrelated with variance σ_v^2. Note that, in each block, the first K of the N slots are used for training purpose.

We now briefly review the original WR-based semiblind method studied in [10], [15]–[18]. For the sake of simplicity, this method is referred to as the WR-based method in the rest of the paper. Its idea originates from a decomposition of the channel matrix,

$$\mathbf{H} = \mathbf{W}\mathbf{Q}^H \tag{2}$$

where $\mathbf{W}$ is a whitening matrix and $\mathbf{Q}$ is a unitary rotation matrix. Performing the singular value decomposition (SVD) of $\mathbf{H}$ gives

$$\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H. \tag{3}$$

Obviously, one possible choice of $\mathbf{W}$ and $\mathbf{Q}$ can be $\mathbf{U}\mathbf{\Sigma}$ and $\mathbf{V}$, respectively. In general, the WR-based method can be implemented by the following two steps:

i) estimate the whitening matrix $\mathbf{W}$ in a blind fashion using the autocorrelation matrix of the received signal along with a subspace-based method;
ii) estimate the unitary rotation matrix $\mathbf{Q}$ by utilizing training pilots and a constrained maximum-likelihood (ML) method.

In our previous work [19], the first-order perturbation analysis is used to evaluate the performance of the WR-based method. It has been shown that, by considering the perturbation only due to the finite data length in the computation of the correlation matrices, the whitening matrix would be perturbed even in the absence of noise. Using (1), the autocorrelation matrix of the received signal, with such a perturbation can be estimated as

$$\hat{\mathbf{R}}_{\mathbf{Y}} \triangleq \frac{1}{N}\sum_{n=1}^{N}\mathbf{y}(n)\mathbf{y}^H(n) - \sigma_v^2\mathbf{I} = \mathbf{H}\left[\sigma_x^2\mathbf{I} + \Delta\mathbf{R}_x\right]\mathbf{H}^H + \Delta\mathbf{R}_v \tag{4}$$

where $\Delta\mathbf{R}_x$ denotes the signal perturbation matrix

$$\Delta\mathbf{R}_x \triangleq \frac{1}{N}\sum_{n=1}^{N}\mathbf{x}(n)\mathbf{x}^H(n) - \sigma_x^2\mathbf{I} \tag{5}$$

and $\Delta\mathbf{R}_v$ the perturbation matrix introduced by the noise

$$\Delta\mathbf{R}_v \triangleq \mathbf{H}\Delta\mathbf{R}_{xv} + \Delta\mathbf{R}_{xv}^H\mathbf{H}^H + \Delta\mathbf{R}_{vv} \tag{6}$$

with

$$\Delta\mathbf{R}_{xv} \triangleq \frac{1}{N}\sum_{n=1}^{N}\mathbf{x}(n)\mathbf{v}^H(n), \quad \Delta\mathbf{R}_{vv} \triangleq \frac{1}{N}\sum_{n=1}^{N}\mathbf{v}(n)\mathbf{v}^H(n) - \sigma_v^2\mathbf{I}. \tag{7}$$

In the noise-free case, all the perturbation terms introduced by the noise would disappear. Then, (4) reduces to

$$\hat{\mathbf{R}}_{\mathbf{Y}} = \mathbf{H}[\mathbf{I} + \Delta\mathbf{R}_x]\mathbf{H}^H. \tag{8}$$

Based on the above perturbation expressions, we can derive the estimate of $\mathbf{W}$. Using (3) into (8), one can get

$$\hat{\mathbf{R}}_{\mathbf{Y}} = \mathbf{U}\mathbf{I}_C\mathbf{T}\mathbf{I}_C^H\mathbf{U}^H \tag{9}$$

where

$$\mathbf{I}_C \triangleq \begin{bmatrix}\mathbf{I}_{N_T\times N_T}\\ \mathbf{0}_{(N_R-N_T)\times N_T}\end{bmatrix} \tag{10}$$

$$\mathbf{T} \triangleq \mathbf{\Sigma}_S^2 + \mathbf{\Sigma}_S\mathbf{V}^H\Delta\mathbf{R}_x\mathbf{V}\mathbf{\Sigma}_S. \tag{11}$$

In (11), $\mathbf{\Sigma}_S$ is a diagonal matrix satisfying $\mathbf{\Sigma} = \mathbf{I}_C\mathbf{\Sigma}_S$. Since $\mathbf{T}$ is Hermitian and the signal perturbation matrix $\Delta\mathbf{R}_x$ has a small norm, the singular values of $\mathbf{T}$ would be different from but close to those of $\mathbf{\Sigma}_S^2$. Thus, an SVD of $\mathbf{T}$ can be written as

$$\mathbf{T} = \mathbf{\Pi}_S(\mathbf{\Sigma}_S + \Delta\mathbf{\Sigma}_S)^2\mathbf{\Pi}_S^H \tag{12}$$

where $\mathbf{\Pi}_S$ is a unitary matrix and $\Delta\mathbf{\Sigma}_S$ is considered as the perturbation error of $\mathbf{\Sigma}_S$. Substituting (12) into (9), comparing it with (8) and noting that $\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}$, one can obtain an estimate of the whitening matrix $\mathbf{W}$ as

$$\hat{\mathbf{W}} = \hat{\mathbf{U}}\hat{\mathbf{\Sigma}} = \mathbf{U}\mathbf{I}_C\mathbf{\Pi}_S(\mathbf{\Sigma}_S + \Delta\mathbf{\Sigma}_S). \tag{13}$$

From (13), one can find that, even in the noise-free case, $\hat{\mathbf{W}}$ consists of two perturbation terms $\mathbf{\Pi}_S$ and $\Delta\mathbf{\Sigma}_S$, which are dictated by the signal perturbation matrix $\Delta\mathbf{R}_x$. This could explain why the performance of the WR-based method is very poor in the moderate to high SNR cases.

It is worth noting that in the case of higher SNRs, a nulling-based semiblind approach [11], [20] performs better than the WR-based method. Instead of estimating the whitening matrix $\mathbf{W}$, this approach uses a subspace method to obtain an estimate of the nulling subspace of the channel matrix, $\hat{\mathbf{U}}_{\text{null}}$. By utilizing $\hat{\mathbf{U}}_{\text{null}}$ in conjunction with a training-based least-square (LS) criterion, a semiblind cost function can then be formulated as

$$\min_{\mathbf{H}} \Delta = \|\mathbf{Y}_{\text{P}} - \mathbf{H}\mathbf{X}_{\text{P}}\|_F^2 + \alpha\left\|\hat{\mathbf{U}}_{\text{null}}^H\mathbf{H}\right\|_F^2 \tag{14}$$

where $\mathbf{X}_{\text{P}} \triangleq [\mathbf{x}(1), \cdots, \mathbf{x}(K)]$ is the pilot signal matrix, $\mathbf{Y}_{\text{P}} \triangleq [\mathbf{y}(1), \cdots, \mathbf{y}(K)]$ the corresponding received pilot matrix, and $\alpha > 0$ is a weighting factor. By performing an SVD on $\hat{\mathbf{R}}_{\mathbf{Y}}$ in (8), it can be verified that an ideal nulling constraint $\hat{\mathbf{U}}_{\text{null}}$, namely

$$\hat{\mathbf{U}}_{\text{null}}^H\mathbf{H} = \mathbf{0}$$

is obtained in the absence of noise, even if the signal perturbation exists. This implies that in the noise-free case the blind constraint in the nulling-based method is perfectly satisfied without being affected by the signal perturbation term $\Delta\mathbf{R}_x$, and therefore, the nulling-based method is superior to the WR-based method. It has been shown in our previous work [19] that, by properly choosing the weighting factor α, the nulling-based method can achieve a much better performance than the WR-based method in higher SNRs. In the next section, we will propose a novel transmit scheme to improve the estimation performance of the whitening matrix in the WR-based method. As will be shown, the proposed method significantly outperforms both the WR-based method and the nulling-based approach in all SNR cases.

III. Proposed Signal-Perturbation-Free Transmit Scheme

The new signal-perturbation-free transmit scheme is to send information of the signal perturbation matrix $\Delta\mathbf{R}_x$ to the receiver. The received version of this information will be then exploited to cancel the signal perturbation error.

Our idea begins with the eigenvalue decomposition (EVD) of $\Delta\mathbf{R}_x$, which is re-defined as the scaled version of (5) for notational convenience,

$$\Delta\mathbf{R}_A = \sum_{n=1}^{N}\mathbf{x}(n)\mathbf{x}^H(n) - N\sigma_x^2\mathbf{I}. \tag{15}$$

As $\Delta\mathbf{R}_A$ is a Hermitian matrix, its EVD can be written as [30]

$$\Delta\mathbf{R}_A = [\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_{N_T}]\begin{bmatrix}\sigma_1 & 0 & \cdots & 0\\ 0 & \sigma_2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \sigma_{N_T}\end{bmatrix}\begin{bmatrix}\mathbf{u}_1^H\\ \mathbf{u}_2^H\\ \vdots\\ \mathbf{u}_{N_T}^H\end{bmatrix} \tag{16}$$

where σ_i, $(i = 1, 2, \cdots, N_T)$, are real numbers representing the eigenvalues of $\Delta\mathbf{R}_A$, and $\mathbf{u}_i$, $(i = 1, 2, \cdots, N_T)$, are the corresponding eigenvectors. Using (16), one can separate $\Delta\mathbf{R}_A$ into two parts as

$$\Delta\mathbf{R}_A = \Delta\mathbf{R}_{\text{pos}} - \Delta\mathbf{R}_{\text{neg}} \tag{17}$$

where

$$\Delta\mathbf{R}_{\text{pos}} = \sum_{i=1}^{L_{\text{pos}}} \sigma_{\text{pos},i}\mathbf{u}_{\text{pos},i}\mathbf{u}_{\text{pos},i}^H \tag{18}$$

$$\Delta\mathbf{R}_{\text{neg}} = \sum_{i=1}^{L_{\text{neg}}} (-\sigma_{\text{neg},i})\mathbf{u}_{\text{neg},i}\mathbf{u}_{\text{neg},i}^H. \tag{19}$$

Here, $\sigma_{\text{pos},i}$ represents a positive eigenvalue, $\mathbf{u}_{\text{pos},i}$ is the eigenvector associated with $\sigma_{\text{pos},i}$, and the L_{pos} is the total number of positive eigenvalues. Likewise, $\sigma_{\text{neg},i}$, $\mathbf{u}_{\text{neg},i}$ and L_{neg} refer to similar quantities with respect to the negative eigenvalues. In what follows, we will derive a further decomposed form of $\Delta\mathbf{R}_A$ based on (17) such that the information of $\Delta\mathbf{R}_{\text{pos}}$ and $\Delta\mathbf{R}_{\text{neg}}$ can easily be transmitted to the receiver.

By letting

$$\Delta\mathbf{R}_{\text{pos}} = \eta\mathbf{X}_{\text{pos}}\mathbf{X}_{\text{pos}}^H \tag{20}$$

$$\Delta\mathbf{R}_{\text{neg}} = \eta\mathbf{X}_{\text{neg}}\mathbf{X}_{\text{neg}}^H \tag{21}$$

(17) can be rewritten as

$$\Delta\mathbf{R}_A = \eta\left(\mathbf{X}_{\text{pos}}\mathbf{X}_{\text{pos}}^H - \mathbf{X}_{\text{neg}}\mathbf{X}_{\text{neg}}^H\right) \tag{22}$$

where η is a scaling factor, and $\mathbf{X}_{\text{pos}}$ and $\mathbf{X}_{\text{neg}}$ are two matrices containing the information of $\Delta\mathbf{R}_{\text{pos}}$ and that of $\Delta\mathbf{R}_{\text{neg}}$, respectively. We now first show that as long as $\mathbf{X}_{\text{pos}}$ and $\mathbf{X}_{\text{neg}}$ are transmitted to the receiver, the signal perturbation error can completely be cancelled. We will then show that $\mathbf{X}_{\text{pos}}$ and $\mathbf{X}_{\text{neg}}$ can easily be constructed using the eigenvalues and the eigenvectors of $\Delta\mathbf{R}_A$. As shown later, the size of $\mathbf{X}_{\text{pos}}$ and $\mathbf{X}_{\text{neg}}$ can be made comparable to the dimension of $\Delta\mathbf{R}_A$, namely, the number of the transmit antennas N_T. Therefore, the spectral resource used for transmitting $\mathbf{X}_{\text{pos}}$ and $\mathbf{X}_{\text{neg}}$ is negligible as compared to that of the user data.

Letting $\mathbf{Y}_{\text{pos}}$ and $\mathbf{Y}_{\text{neg}}$ be the received signals corresponding to $\mathbf{X}_{\text{pos}}$ and $\mathbf{X}_{\text{neg}}$, respectively, namely,

$$\mathbf{Y}_{\text{pos}} = \mathbf{H}\mathbf{X}_{\text{pos}} + \mathbf{V}_{\text{pos}} \tag{23}$$

$$\mathbf{Y}_{\text{neg}} = \mathbf{H}\mathbf{X}_{\text{neg}} + \mathbf{V}_{\text{neg}} \tag{24}$$

where $\mathbf{V}_{\text{pos}}$ and $\mathbf{V}_{\text{neg}}$ are the corresponding noise matrices, the received version of the signal perturbation matrix can be defined as

$$\Delta\hat{\mathbf{R}}_{\mathbf{Y}} = \frac{\eta}{N}\left[\left(\mathbf{Y}_{\text{pos}}\mathbf{Y}_{\text{pos}}^H - \mathbf{Y}_{\text{neg}}\mathbf{Y}_{\text{neg}}^H\right) - (N_{\text{pos}} - N_{\text{neg}})\sigma_v^2\mathbf{I}\right] \tag{25}$$

where N_{pos} and N_{neg} denote the number of columns of $\mathbf{Y}_{\text{pos}}$ and that of $\mathbf{Y}_{\text{neg}}$, respectively. Using (23) and (24) into (25) and noting that $\Delta\mathbf{R}_A = N\Delta\mathbf{R}_x$, we obtain

$$\Delta\hat{\mathbf{R}}_{\mathbf{Y}} = \mathbf{H}\Delta\mathbf{R}_x\mathbf{H}^H + \Delta\mathbf{R}_{vp} \tag{26}$$

where $\Delta\mathbf{R}_{vp}$ represents a perturbation term introduced by the noise. By utilizing (4) and (26), the received correlation matrix without the signal perturbation error can be obtained from

$$\hat{\mathbf{R}}'_{\mathbf{Y}} = \hat{\mathbf{R}}_{\mathbf{Y}} - \Delta\hat{\mathbf{R}}_{\mathbf{Y}} = \mathbf{H}\mathbf{H}^H + \Delta\mathbf{R}'_v \tag{27}$$

where

$$\Delta\mathbf{R}'_v = \Delta\mathbf{R}_v - \Delta\mathbf{R}_{vp}. \tag{28}$$

As a result, the signal perturbation error has been completely eliminated through the transmission of $\mathbf{X}_{\text{pos}}$ and $\mathbf{X}_{\text{neg}}$. What remains to be done in the proposed scheme is to determine the matrices $\mathbf{X}_{\text{pos}}$ and $\mathbf{X}_{\text{neg}}$ from the eigenvalues σ_i and the eigenvectors $\mathbf{u}_i$.

Note that the total power of N_T transmit antennas in each time slot can be written as $\sigma_{\text{int}} \triangleq N_T\sigma_x^2$. It is found from a large number of simulation experiments that the value of $\sigma_{\text{pos},i}$ is much larger than σ_{int}. In order to ensure a reliable transmission of $\mathbf{X}_{\text{pos}}$ in noisy conditions, we should allow the use of multiple slots to carry a scaled version of $\sigma_{\text{pos},i}$. To this end, we divide $\sigma_{\text{pos},i}$ by the scaling factor η and then split it into $N_{\text{pos},i}$ terms of σ_{int} plus one fractional term as

$$\frac{\sigma_{\text{pos},i}}{\eta} = (N_{\text{pos},i}\sigma_{\text{int}} + \sigma_{\text{pos-frac},i}) \tag{29}$$

where

$$N_{\text{pos},i} = \left\lfloor \frac{\sigma_{\text{pos},i}}{\eta\sigma_{\text{int}}} \right\rfloor \tag{30}$$

$$\sigma_{\text{pos-frac},i} = \frac{\sigma_{\text{pos},i}}{\eta} - N_{\text{pos},i}\sigma_{\text{int}}. \tag{31}$$

Letting

$$\mathbf{x}_{\text{pos-int},i} = \sqrt{\sigma_{\text{int}}}\mathbf{u}_{\text{pos},i},$$

$$\mathbf{x}_{\text{pos-frac},i} = \sqrt{\sigma_{\text{pos-frac},i}}\mathbf{u}_{\text{pos},i},$$

one can construct an $N_T \times (N_{\text{pos},i} + 1)$ matrix $\mathbf{X}_{\text{pos},i}$ for the ith eigenvalue by stacking $N_{\text{pos},i}$ consecutive vectors $\mathbf{x}_{\text{pos-int},i}$ and one vector $\mathbf{x}_{\text{pos-frac},i}$, which satisfies

$$\eta\mathbf{X}_{\text{pos},i}\mathbf{X}_{\text{pos},i}^H = \sigma_{\text{pos},i}\mathbf{u}_{\text{pos},i}\mathbf{u}_{\text{pos},i}^H.$$

Thus, the complete $\mathbf{X}_{\text{pos}}$ can be formed as

$$\mathbf{X}_{\text{pos}} = \left[\mathbf{X}_{\text{pos},1}, \mathbf{X}_{\text{pos},2}, \cdots, \mathbf{X}_{\text{pos},L_{\text{pos}}}\right]. \tag{32}$$

Obviously, the number of columns of $\mathbf{X}_{\text{pos}}$ is given by $N_{\text{pos}} = L_{\text{pos}} + \sum_{i=1}^{L_{\text{pos}}} N_{\text{pos},i}$.

In a similar manner, $\mathbf{X}_{\text{neg}}$ can be constructed as follows

$$\mathbf{X}_{\text{neg}} = \left[\mathbf{X}_{\text{neg},1}, \mathbf{X}_{\text{neg},2}, \cdots, \mathbf{X}_{\text{neg},L_{\text{neg}}}\right] \tag{33}$$

where $\mathbf{X}_{\text{neg},i}$ consists of $N_{\text{neg},i} = \lfloor -\sigma_{\text{neg},i}/\eta\sigma_{\text{int}} \rfloor$ consecutive vectors $\mathbf{x}_{\text{neg-int},i}$ and one vector $\mathbf{x}_{\text{neg-frac},i}$ as given by

$$\mathbf{x}_{\text{neg-int},i} = \sqrt{\sigma_{\text{int}}}\mathbf{u}_{\text{neg},i} \tag{34}$$

$$\mathbf{x}_{\text{neg-frac},i} = \sqrt{\sigma_{\text{neg-frac},i}}\mathbf{u}_{\text{neg},i} \tag{35}$$

$$\sigma_{\text{neg-frac},i} = \frac{-\sigma_{\text{neg},i}}{\eta} - N_{\text{neg},i}\sigma_{\text{int}}. \tag{36}$$

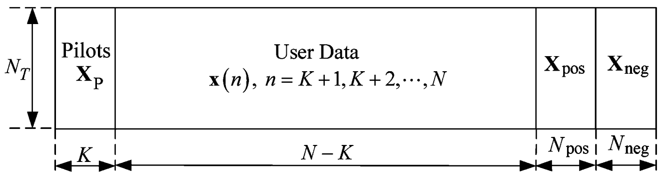


Fig. 1. New transmit structure.

Note that the number of columns of $\mathbf{X}_\mathrm{neg}$ is

$$N_\mathrm{neg} = L_\mathrm{neg} + \sum_{i=1}^{L_\mathrm{neg}} N_{\mathrm{neg},i}. \tag{37}$$

From the above discussion, a new transmit structure, which consists of conventional pilots, user's data and the additional data $\mathbf{X}_\mathrm{pos}$ and $\mathbf{X}_\mathrm{neg}$, referred to as the signal perturbation free (SPF) data for convenience, can be designed as shown in Fig. 1. As usual, the user-data are generated from a finite constellation such as quadrature phase-shift keying (QPSK). In order to achieve a superior channel estimation performance in MIMO systems, however, the pilots are often designed as complex-valued data [31], [32], which are, in general, not from the user-data constellation. In the proposed scheme, the SPF data are also complex-valued and therefore, they can be transmitted in the same manner as the complex-valued pilots.

The advantage of the proposed transmit scheme lies in that the receiver does not need detect or estimate the transmitted SPF data $\mathbf{X}_\mathrm{pos}$ and $\mathbf{X}_\mathrm{neg}$ and it can rely on $\mathbf{Y}_\mathrm{pos}$ and $\mathbf{Y}_\mathrm{neg}$, the received version of $\mathbf{X}_\mathrm{pos}$ and $\mathbf{X}_\mathrm{neg}$, to cancel the signal perturbation error by using (27). It should be noted that the total column size of $\mathbf{X}_\mathrm{pos}$ and $\mathbf{X}_\mathrm{neg}$ is inversely proportional to the scaling factor η. It can be shown that when η is sufficiently large, as few as N_T slots can be used for the transmission of $\mathbf{X}_\mathrm{pos}$ and $\mathbf{X}_\mathrm{neg}$. In general, the choice of η should depend on the number of the transmit antennas as well as the length of the user data. Our extensive simulations show that $\eta = 16$ is a proper choice to achieve a very good channel estimate for a 4×8 MIMO system, in which case, the transmission of $\mathbf{X}_\mathrm{pos}$ and $\mathbf{X}_\mathrm{neg}$ requires only 9 slots when the user data length is about 1000.

In summary, the scheme developed above gives a signal-perturbation-free estimate of the whitening matrix in the noise-free case as seen from (27). It should be pointed out that in the presence of noise, although the WR-based method with the new transmit scheme is subject to the noise perturbation, the proposed method still outperforms the WR-based method, since the perturbation introduced by the noise is, in general, significantly smaller than the signal perturbation.

IV. Analysis of the New WR-Based Method With the Proposed Transmit Scheme

It is known that the solution of subspace based methods is always perturbed by various sources, such as the finite data length and the measurement noise [25]–[28]. The perturbation theory has been employed for the analysis of subspace based methods [21]–[24], [26]. In this section, the first-order perturbation theory is employed to analyze the proposed WR-based method, in which the channel estimate can be denoted as $\hat{\mathbf{H}} = \hat{\mathbf{W}}\hat{\mathbf{Q}}^H$. Here, the symbol "^" means the perturbed version of the quantities involved. We first derive the perturbation errors of the whitening and the rotation matrices, namely, $\Delta\mathbf{W} = \hat{\mathbf{W}} - \mathbf{W}$ and $\Delta\mathbf{Q} = \hat{\mathbf{Q}} - \mathbf{Q}$, and then reveal a closed-form expression for the MSE of the new WR-based method with the proposed transmit scheme.

A. The Perturbation Error of the Whitening Matrix $\mathbf{W}$

First of all, from (27), the ideal correlation matrix $\mathbf{R}_\mathbf{Y}$ without both the signal and the noise perturbation terms can be written as

$$\mathbf{R}_\mathbf{Y} \triangleq \mathbf{H}\mathbf{H}^H = [\mathbf{U}_S, \mathbf{U}_N] \begin{bmatrix} \mathbf{\Sigma}_S^2 & \\ & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{U}_S^H \\ \mathbf{U}_N^H \end{bmatrix}. \tag{38}$$

In obtaining (38), we have employed the SVD of $\mathbf{H}$ as given by

$$\mathbf{H} = [\mathbf{U}_S, \mathbf{U}_N] \begin{bmatrix} \mathbf{\Sigma}_S \\ \mathbf{0} \end{bmatrix} \mathbf{V}_S^H \tag{39}$$

where $\mathbf{V}_S = \mathbf{V}$ under the assumption of the full-column rank of $\mathbf{H}$.

In the case of noise perturbation but without the signal perturbation error, the autocorrelation matrix of the received signal can be written as

$$\hat{\mathbf{R}}'_\mathbf{Y} = \hat{\mathbf{H}}\hat{\mathbf{H}}^H \tag{40}$$

where $\hat{\mathbf{H}}$ denotes the estimate of the channel matrix containing the noise perturbation error. Let the SVD of $\hat{\mathbf{H}}$ be given by

$$\hat{\mathbf{H}} = \hat{\mathbf{U}}\hat{\mathbf{\Sigma}}\hat{\mathbf{V}}^H. \tag{41}$$

As $\mathbf{H}$ has a full column rank, one can partition $\hat{\mathbf{U}}$ and $\hat{\mathbf{\Sigma}}$ according to the signal and the noise subspaces of $\hat{\mathbf{H}}$ as

$$\hat{\mathbf{U}} = [\hat{\mathbf{U}}_S, \hat{\mathbf{U}}_N], \hat{\mathbf{\Sigma}} = \begin{bmatrix} \hat{\mathbf{\Sigma}}_S \\ \mathbf{0} \end{bmatrix}.$$

Using the above partitioned form in (41) and then substituting (41) into (40), we have

$$\hat{\mathbf{R}}'_\mathbf{Y} = [\hat{\mathbf{U}}_S, \hat{\mathbf{U}}_N] \begin{bmatrix} \hat{\mathbf{\Sigma}}_S^2 & \\ & \mathbf{0} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{U}}_S^H \\ \hat{\mathbf{U}}_N^H \end{bmatrix}. \tag{42}$$

Clearly, (42) gives an SVD of $\hat{\mathbf{R}}'_\mathbf{Y}$, implying that $\hat{\mathbf{U}}_S$ and $\hat{\mathbf{\Sigma}}_S$ can be obtained from the SVD of $\hat{\mathbf{R}}'_\mathbf{Y}$. Comparing (40) with (42) and noting that $\hat{\mathbf{H}} = \hat{\mathbf{W}}\hat{\mathbf{Q}}^H$, we can obtain the perturbed version of $\mathbf{W}$,

$$\hat{\mathbf{W}} = \hat{\mathbf{U}}_S\hat{\mathbf{\Sigma}}_S. \tag{43}$$

By defining $\Delta\mathbf{U}_S \triangleq \hat{\mathbf{U}}_S - \mathbf{U}_S$ and $\Delta\mathbf{\Sigma}_S \triangleq \hat{\mathbf{\Sigma}}_S - \mathbf{\Sigma}_S$, the perturbation error of $\mathbf{W}$ can be expressed as

$$\begin{aligned} \Delta\mathbf{W} &= \hat{\mathbf{W}} - \mathbf{W} = \hat{\mathbf{U}}_S\hat{\mathbf{\Sigma}}_S - \mathbf{U}_S\mathbf{\Sigma}_S, \\ &\approx \Delta\mathbf{U}_S\mathbf{\Sigma}_S + \mathbf{U}_S\Delta\mathbf{\Sigma}_S \end{aligned} \tag{44}$$

where the second-order perturbation error term has been neglected. In the following, we simplify the computation of $\Delta\mathbf{W}$ by investigating $\Delta\mathbf{U}_S$ and $\Delta\mathbf{\Sigma}_S$.

In a manner similar to the one used in Subsection II.A of [23], the first-order perturbation error of $\mathbf{U}_S$ can be obtained as

$$\Delta\mathbf{U}_S \approx \mathbf{U}_N\mathbf{U}_N^H\Delta\mathbf{R}'_v\mathbf{U}_S\mathbf{\Sigma}_S^{-2}. \tag{45}$$

One can easily verify that, when $N \gg N_{\text{plus}} + N_{\text{neg}}$, $\Delta\mathbf{R}_{vp}$ in (28) can be neglected and thus $\Delta\mathbf{R}'_v$ approximates to $\Delta\mathbf{R}_v$. Using this result into (6) and neglecting the perturbation matrix $\Delta\mathbf{R}_{vv}$ of the noise autocorrelation for a medium to high SNR, we have

$$\Delta\mathbf{R}'_v \approx \mathbf{H}\Delta\mathbf{R}_{xv} + \Delta\mathbf{R}_{xv}^H\mathbf{H}^H. \quad (46)$$

Substituting (46) into (45) and utilizing (39), one can obtain

$$\Delta\mathbf{U}_S \approx \mathbf{U}_N\mathbf{U}_N^H\Delta\mathbf{R}_{xv}^H\mathbf{V}_S\mathbf{\Sigma}_S^{-1}. \quad (47)$$

On the other hand, using (27), (42) and the first-order approximation [26], one can show that

$$(\mathbf{\Sigma}_S + \Delta\mathbf{\Sigma}_S)^2 - \mathbf{\Sigma}_S^2 \approx \mathbf{U}_S^H\Delta\mathbf{R}'_v\mathbf{U}_S. \quad (48)$$

Using (39) and (46), (48) can be rewritten as

$$(\mathbf{\Sigma}_S + \Delta\mathbf{\Sigma}_S)^2 - \mathbf{\Sigma}_S^2 \approx \mathbf{\Sigma}_S\mathbf{V}_S^H\Delta\mathbf{R}_{xv}\mathbf{U}_S + \mathbf{U}_S^H\Delta\mathbf{R}_{xv}^H\mathbf{V}_S\mathbf{\Sigma}_S^H \quad (49)$$

which, under the assumption that the second-order perturbation error terms can be ignored, leads to

$$\Delta\mathbf{\Sigma}_S \approx \mathbf{U}_S^H\Delta\mathbf{R}_{xv}^H\mathbf{V}_S. \quad (50)$$

Using (47) and (50) into (44), $\Delta\mathbf{W}$ can be rewritten as

$$\Delta\mathbf{W} \approx \mathbf{U}_N\mathbf{U}_N^H\Delta\mathbf{R}_{xv}^H\mathbf{V}_S + \mathbf{U}_S\mathbf{U}_S^H\Delta\mathbf{R}_{xv}^H\mathbf{V}_S. \quad (51)$$

Recalling that $\mathbf{U}_N\mathbf{U}_N^H + \mathbf{U}_S\mathbf{U}_S^H = \mathbf{I}$, (51) reduces to

$$\Delta\mathbf{W} \approx \Delta\mathbf{R}_{xv}^H\mathbf{V}_S. \quad (52)$$

As will be shown in the following subsections, (52) can be used to derive the MSE of the proposed signal-perturbation-free WR approach.

B. The Perturbation Error of the Rotation Matrix $\mathbf{Q}$

Prior to the derivation of the perturbation error of $\mathbf{Q}$, we first conduct an analysis for the training-based estimation of $\mathbf{Q}$ given in [15]. In the noise-free case, the rotation matrix $\mathbf{Q}$ can be calculated from

$$\mathbf{Q} = \mathbf{V}_Q\mathbf{U}_Q^H \quad (53)$$

where $\mathbf{U}_Q$ and $\mathbf{V}_Q$ are obtained from an SVD of the matrix

$$\mathbf{Y}_Q \triangleq \frac{1}{K\sigma_x^2}\mathbf{W}^H\mathbf{Y}_P\mathbf{X}_P^H \quad (54)$$

namely

$$\mathbf{Y}_Q = \mathbf{U}_Q\mathbf{\Sigma}_Q\mathbf{V}_Q^H. \quad (55)$$

By noting that $\mathbf{W} = \mathbf{U}_S\mathbf{\Sigma}_S$ and

$$\frac{1}{K\sigma_x^2}\mathbf{Y}_P\mathbf{X}_P^H = \mathbf{H} \quad (56)$$

(54) can be rewritten as

$$\mathbf{Y}_Q = \mathbf{\Sigma}_S^2\mathbf{V}_S^H. \quad (57)$$

As such, one realization of the SVD of $\mathbf{Y}_Q$ is given by

$$\mathbf{U}_Q = \mathbf{I}, \mathbf{\Sigma}_Q = \mathbf{\Sigma}_S^2, \mathbf{V}_Q = \mathbf{V}_S. \quad (58)$$

Using (58) into (53) yields

$$\mathbf{Q} = \mathbf{V}_S. \quad (59)$$

The above discussion indicates that the method in [15] gives an ideal rotation matrix $\mathbf{V}_S$ in the noise-free case. In the following, we derive an expression of the perturbation error $\Delta\mathbf{Q}$ of $\mathbf{Q}$ in the presence of noise.

In the noisy case, (54) should be modified to

$$\hat{\mathbf{Y}}_Q = \frac{1}{K\sigma_x^2}\hat{\mathbf{W}}^H\mathbf{Y}_P\mathbf{X}_P^H \quad (60)$$

where $\hat{\mathbf{W}} = \mathbf{W} + \Delta\mathbf{W}$ is the perturbed version of $\mathbf{W}$, which has been discussed in the previous subsection. Noting that $\sigma_x^2 = 1$, one can easily verify that

$$\frac{1}{K}\mathbf{Y}_P\mathbf{X}_P^H = \mathbf{H} + \Delta\mathbf{R}_{xv,P}^H \quad (61)$$

where

$$\Delta\mathbf{R}_{xv,P} \triangleq \frac{1}{K}\sum_{n=1}^{K}\mathbf{x}(n)\mathbf{v}^H(n). \quad (62)$$

Using (52) and (61), (60) can be expressed as

$$\hat{\mathbf{Y}}_Q \approx \mathbf{\Sigma}_S^2\mathbf{V}_S^H + \mathbf{V}_S^H\Delta\mathbf{R}_{xv}\mathbf{H} + \mathbf{W}^H\Delta\mathbf{R}_{xv,P}^H \quad (63)$$

which can be rewritten, utilizing (57), as

$$\hat{\mathbf{Y}}_Q = \mathbf{Y}_Q + \Delta\mathbf{Y}_Q \quad (64)$$

where

$$\Delta\mathbf{Y}_Q \approx \mathbf{V}_S^H\Delta\mathbf{R}_{xv}\mathbf{H} + \mathbf{W}^H\Delta\mathbf{R}_{xv,P}^H. \quad (65)$$

Our next goal is to disclose an expression for $\Delta\mathbf{Q}$ in terms of the perturbation error $\Delta\mathbf{Y}_Q$ of $\mathbf{Y}_Q$ that is caused by noise. Let us consider the SVD of $\hat{\mathbf{Y}}_Q$

$$\begin{aligned}\hat{\mathbf{Y}}_Q &= \hat{\mathbf{U}}_Q\hat{\mathbf{\Sigma}}_Q\hat{\mathbf{V}}_Q^H \\ &= [(\mathbf{U}_Q + \Delta\mathbf{U}_Q)\mathbf{P}](\mathbf{\Sigma}_Q + \Delta\mathbf{\Sigma}_Q) \\ &\quad \times [(\mathbf{V}_Q + \Delta\mathbf{V}_Q)\mathbf{P}]^H \end{aligned} \quad (66)$$

where $\mathbf{P}$ is a diagonal unitary matrix used to represent a general form of the SVD, since the SVD of $\hat{\mathbf{Y}}_Q$ is not unique. By utilizing $\hat{\mathbf{U}}_Q = (\mathbf{U}_Q + \Delta\mathbf{U}_Q)\mathbf{P}$ and $\hat{\mathbf{V}}_Q = (\mathbf{V}_Q + \Delta\mathbf{V}_Q)\mathbf{P}$, we have

$$\begin{aligned}\Delta\mathbf{Q} &\triangleq \hat{\mathbf{Q}} - \mathbf{Q} = \hat{\mathbf{V}}_Q\hat{\mathbf{U}}_Q^H - \mathbf{V}_Q\mathbf{U}_Q^H \\ &\approx \Delta\mathbf{V}_Q\mathbf{U}_Q^H + \mathbf{V}_Q\Delta\mathbf{U}_Q^H. \end{aligned} \quad (67)$$

Using (58), (67) reduces to

$$\Delta\mathbf{Q} \approx \Delta\mathbf{V}_Q + \mathbf{V}_S\Delta\mathbf{U}_Q^H. \quad (68)$$

In general, if the noise subspace of $\mathbf{Y}_Q$ exists, the expressions for $\Delta\mathbf{V}_Q$ and $\Delta\mathbf{U}_Q$ can be easily derived in terms of the perturbation error $\Delta\mathbf{Y}_Q$ by utilizing the analysis results of the existing perturbation theory such as that in [9], [21]–[28]. However, these results are not applicable to our case, since only the signal subspace exists in $\mathbf{Y}_Q$. To overcome this difficulty, we propose the following theorem to derive an expression for $\Delta\mathbf{Q}$ in terms of the perturbation error $\Delta\mathbf{Y}_Q$.

Theorem 4.1: Given an SVD of a full rank $M \times M$ matrix $\mathbf{Z}$, $\mathbf{Z} = \mathbf{U}_Z\mathbf{\Sigma}_Z\mathbf{V}_Z^H$, where $\mathbf{\Sigma}_Z = \mathrm{diag}(\sigma_{z_1}, \sigma_{z_2}, \cdots, \sigma_{z_M})$, for a small perturbed term $\Delta\mathbf{Z}$, the SVD of $\hat{\mathbf{Z}} = \mathbf{Z} + \Delta\mathbf{Z}$ is defined as

$$\begin{aligned}\hat{\mathbf{Z}} &= \hat{\mathbf{U}}_Z\hat{\mathbf{\Sigma}}_Z\hat{\mathbf{V}}_Z^H \\ &= [(\mathbf{U}_Z + \Delta\mathbf{U}_Z)\mathbf{P}_Z](\mathbf{\Sigma}_Z + \Delta\mathbf{\Sigma}_Z) \\ &\quad \times [(\mathbf{V}_Z + \Delta\mathbf{V}_Z)\mathbf{P}_Z]^H \end{aligned} \tag{69}$$

where $\mathbf{P}_Z$ is a diagonal unitary matrix whose role is the same as $\mathbf{P}$ in (66). Then, it can be shown that

$$\begin{aligned}\mathbf{\Omega} &\triangleq \Delta\mathbf{U}_Z^H\mathbf{U}_Z + \mathbf{V}_Z^H\Delta\mathbf{V}_Z \\ &\approx \mathbf{\Gamma}_Z \circ \left(\mathbf{V}_Z^H\Delta\mathbf{Z}^H\mathbf{U}_Z - \mathbf{U}_Z^H\Delta\mathbf{Z}\mathbf{V}_Z\right)\end{aligned} \tag{70}$$

where

$$\mathbf{\Gamma}_Z = \begin{bmatrix} \frac{1}{2\sigma_{z_1}} & \frac{1}{\sigma_{z_1}+\sigma_{z_2}} & \cdots & \frac{1}{\sigma_{z_1}+\sigma_{z_M}} \\ \frac{1}{\sigma_{z_2}+\sigma_{z_1}} & \frac{1}{2\sigma_{z_2}} & \cdots & \frac{1}{\sigma_{z_2}+\sigma_{z_M}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{\sigma_{z_M}+\sigma_{z_1}} & \frac{1}{\sigma_{z_M}+\sigma_{z_2}} & \cdots & \frac{1}{2\sigma_{z_M}} \end{bmatrix}. \tag{71}$$

The proof of the theorem is given in Appendix A.

In order to use *Theorem* 4.1, we first rewrite (68) as

$$\Delta\mathbf{Q} \approx \mathbf{V}_S\left(\Delta\mathbf{U}_Q^H\mathbf{I} + \mathbf{V}_S^H\Delta\mathbf{V}_Q\right). \tag{72}$$

With replacements $\Delta\mathbf{U}_Z = \Delta\mathbf{U}_Q$, $\mathbf{U}_Z = \mathbf{I}$, $\mathbf{V}_Z = \mathbf{V}_S$, $\Delta\mathbf{V}_Z = \Delta\mathbf{V}_Q$, $\mathbf{\Sigma}_Z = \mathbf{\Sigma}_S^2$ and $\Delta\mathbf{Z} = \Delta\mathbf{Y}_Q$ in (70), we immediately obtain

$$\Delta\mathbf{Q} \approx \mathbf{V}_S(\mathbf{\Gamma}_Q \circ \mathbf{\Pi}) \tag{73}$$

where

$$\mathbf{\Gamma}_Q = \begin{bmatrix} \frac{1}{2\sigma_{S_1}^2} & \frac{1}{\sigma_{S_1}^2+\sigma_{S_2}^2} & \cdots & \frac{1}{\sigma_{S_1}^2+\sigma_{S_{N_T}}^2} \\ \frac{1}{\sigma_{S_2}^2+\sigma_{S_1}^2} & \frac{1}{2\sigma_{S_2}^2} & \cdots & \frac{1}{\sigma_{S_2}^2+\sigma_{S_{N_T}}^2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{\sigma_{S_{N_T}}^2+\sigma_{S_1}^2} & \frac{1}{\sigma_{S_{N_T}}^2+\sigma_{S_2}^2} & \cdots & \frac{1}{2\sigma_{S_{N_T}}^2} \end{bmatrix} \tag{74}$$

$$\mathbf{\Pi} = \mathbf{V}_S^H\Delta\mathbf{Y}_Q^H - \Delta\mathbf{Y}_Q\mathbf{V}_S. \tag{75}$$

Note that σ_{S_i}, $(i = 1, 2, \cdots, N_T)$ is the diagonal element of $\mathbf{\Sigma}_S$. Substituting (65) into (75), we obtain

$$\begin{aligned}\mathbf{\Pi} \approx\ & \mathbf{\Sigma}_S\mathbf{U}_S^H\Delta\mathbf{R}_{xv}^H\mathbf{V}_S + \mathbf{V}_S^H\Delta\mathbf{R}_{xv,P}\mathbf{U}_S\mathbf{\Sigma}_S \\ & - \mathbf{V}_S^H\Delta\mathbf{R}_{xv}\mathbf{U}_S\mathbf{\Sigma}_S - \mathbf{\Sigma}_S\mathbf{U}_S^H\Delta\mathbf{R}_{xv,P}^H\mathbf{V}_S. \end{aligned} \tag{76}$$

As a result, the perturbation error of $\mathbf{Q}$ has been expressed in terms of the perturbation errors, $\Delta\mathbf{R}_{xv}$ and $\Delta\mathbf{R}_{xv,P}$. In the next subsection, we will derive the MSE of the proposed signal-perturbation-free WR approach by utilizing the expression of $\Delta\mathbf{W}$ and $\Delta\mathbf{Q}$ as given by (52) and (73), respectively.

C. MSE of the Proposed Signal-Perturbation-Free WR-Based Method

The estimation error of the channel matrix $\mathbf{H}$ due to the perturbation error can be written as

$$\Delta\mathbf{H} \triangleq \hat{\mathbf{H}} - \mathbf{H} = \hat{\mathbf{W}}\hat{\mathbf{Q}}^H - \mathbf{W}\mathbf{Q}^H \approx \Delta\mathbf{W}\mathbf{Q}^H + \mathbf{W}\Delta\mathbf{Q}^H. \tag{77}$$

Substituting (52) and (73) into (77) gives

$$\Delta\mathbf{H} \approx \Delta\mathbf{R}_{xv}^H + \mathbf{U}_S\mathbf{\Sigma}_S(\mathbf{\Gamma}_Q \circ \mathbf{\Pi}^H)\mathbf{V}_S^H. \tag{78}$$

Thereby, we have

$$\begin{aligned}\mathrm{vec}(\Delta\mathbf{H}) \approx\ & \mathrm{vec}\left(\Delta\mathbf{R}_{xv}^H\right) + [\mathbf{V}_S^* \otimes (\mathbf{U}_S\mathbf{\Sigma}_S)] \\ & \times \left[\mathrm{vec}(\mathbf{\Gamma}_Q) \circ \mathrm{vec}(\mathbf{\Pi}^H)\right]. \end{aligned} \tag{79}$$

Then, the MSE of the channel estimate can be calculated as

$$\begin{aligned}\mathrm{MSE_{WR}} &\approx \mathrm{Trace}\left\{\mathrm{E}\left[\mathrm{vec}(\Delta\mathbf{H})\mathrm{vec}^H(\Delta\mathbf{H})\right]\right\} \\ &= G_0 + G_1 + G_2 + G_2^* \end{aligned} \tag{80}$$

where

$$G_0 = \mathrm{Trace}\left\{\mathrm{E}\left[\mathrm{vec}\left(\Delta\mathbf{R}_{xv}^H\right)\mathrm{vec}^H\left(\Delta\mathbf{R}_{xv}^H\right)\right]\right\} \tag{81}$$

$$\begin{aligned}G_1 = \mathrm{Trace}\Big\{\mathrm{E}\Big\{&[\mathbf{V}_S^* \otimes (\mathbf{U}_S\mathbf{\Sigma}_S)]\Big\{\left[\mathrm{vec}(\mathbf{\Gamma}_Q)\mathrm{vec}^H(\mathbf{\Gamma}_Q)\right] \\ & \circ\left[\mathrm{vec}(\mathbf{\Pi}^H)\mathrm{vec}^H(\mathbf{\Pi}^H)\right]\Big\} \\ & \times [\mathbf{V}_S^* \otimes (\mathbf{U}_S\mathbf{\Sigma}_S)]^H\Big\}\Big\} \end{aligned} \tag{82}$$

$$\begin{aligned}G_2 = \mathrm{Trace}\Big\{\mathrm{E}\Big\{&\mathrm{vec}\left(\Delta\mathbf{R}_{xv}^H\right)\left[\mathrm{vec}(\mathbf{\Gamma}_Q) \circ \mathrm{vec}(\mathbf{\Pi}^H)\right]^H \\ & \times [\mathbf{V}_S^* \otimes (\mathbf{U}_S\mathbf{\Sigma}_S)]^H\Big\}\Big\}. \end{aligned} \tag{83}$$

In Appendix B, we have proved that

$$G_0 = \frac{N_R N_T}{N}\sigma_v^2 \tag{84}$$

$$G_1 = \frac{N_T^2(N-K)\sigma_v^2}{2KN} \tag{85}$$

$$G_2 = 0. \tag{86}$$

Using (84), (85) and (86) into (80) yields

$$\mathrm{MSE_{WR}} \approx \frac{N_R N_T \sigma_v^2}{N} + \frac{N_T^2(N-K)\sigma_v^2}{2KN}. \tag{87}$$

The above result gives the MSE of the proposed signal-perturbation-free WR-based method in terms of the system configuration and the noise variance, when the signal energy σ_x^2 is normalized to unity. As will be shown in the next section, the theoretical value of the MSE given by (87) is highly consistent with the simulation result. It is to be noted that when $\Delta\mathbf{W} = \mathbf{0}$, the WR-based method turns to the ideal case, where the whitening matrix is obtained directly from the true channel matrix. In this case, the second term on the RHS of (63) vanishes due to (52), and one can verify that the above analysis would lead to the following MSE expression

$$\mathrm{MSE}_{\mathrm{WR,ideal}} \approx \frac{N_T^2 \sigma_v^2}{2K}. \quad (88)$$

It is of interest to note that (88) is the same as the Cramer-Rao bound of the WR semiblind technique derived in [15]. For orthogonal training pilots satisfying $\|\mathbf{X}_\mathrm{P}\|_F^2 = KN_T$, it has been proved in [3] that the expression of the MSE of the training-based LS estimation method is given by

$$\mathrm{MSE}_\mathrm{T} = \frac{N_R N_T \sigma_v^2}{K}. \quad (89)$$

Obviously, one can find that the proposed WR semiblind algorithm can achieve about $2N_R/N_T$ gain over the training-based LS method in terms of MSE of the channel estimate.

V. Simulation Results

We consider a MIMO system with 4 transmit and 8 receive antennas, in which the QPSK modulation is used and a Rayleigh channel, whose elements are i.i.d. complex Gaussian variables with zero mean and unit variance, is assumed. Here, the orthogonal pilots are generated by using the scheme proposed in [31]. In our simulation, we will evaluate the proposed signal perturbation free WR-based method, abbreviated as SPF-WR method in comparison with some of the existing methods, including the LS, the WR-based and the nulling-based methods, in terms of the MSE of the channel estimate well as the bit error rate (BER) performance of the underlying MIMO system. Here, the MSE of the channel estimate is defined as

$$\mathrm{MSE} = \frac{1}{N_\mathrm{MC}} \sum_{n=1}^{N_\mathrm{MC}} \|\hat{\mathbf{H}}_n - \mathbf{H}_n\|^2$$

where N_MC is the number of Monte Carlo iterations, and $\mathbf{H}_n$ and $\hat{\mathbf{H}}_n$ are the true and estimated channel matrices with respect to the nth Monte Carlo iteration, respectively.

It is noted that a joint optimization approach has been proposed in [15] to improve the estimation of the whitening matrix. Although this approach can improve the estimation accuracy of the whitening matrix slightly, the complexity of this method is extremely high since it involves many iterations in the computation of the "fminunc" function in MATLAB, making its implementation very difficult for real-world applications. Thus, this approach is not considered in our experiments. Instead, an ideal WR-based method, which obtains the whitening matrix directly from the true channel matrix, is simulated as a reference method for comparison.

Experiment 1: MSE Versus SNR: In the first experiment, the channel estimation performance in terms of the MSE versus the SNR is investigated. The simulation is undertaken based on 20 000 Monte Carlo runs of the transmission of one data frame with 1000 slots out of which 100 are used as pilots. Fig. 2 shows the MSE plots of the LS method, the ideal WR-based method, the original WR-based method, the nulling-based method and the proposed SPF-WR method with $\eta = 1$, 4, 9, 16, respectively. It is seen that the MSE of the proposed method is closest to that of the ideal WR-based method in comparison to the other methods irrespective of the choice of η. Interestingly, different values of η only make a little difference on the MSE result. However, the number of slots for the transmission of $\mathbf{X}_\mathrm{pos}$ and $\mathbf{X}_\mathrm{neg}$ depends largely on the value of η. We have found that $\eta = 16$ is a very good choice for the proposed method, since it significantly outperforms the WR, the nulling-based and the LS methods at all SNR levels, while requiring a very small number of slots for the transmission of $\mathbf{X}_\mathrm{pos}$ and $\mathbf{X}_\mathrm{neg}$. It is found that the average number of pilots used for $\mathbf{X}_\mathrm{pos}$ and $\mathbf{X}_\mathrm{neg}$ in 20 000 Monte Carlo runs is 5.3 and the maximum number is 10 only.

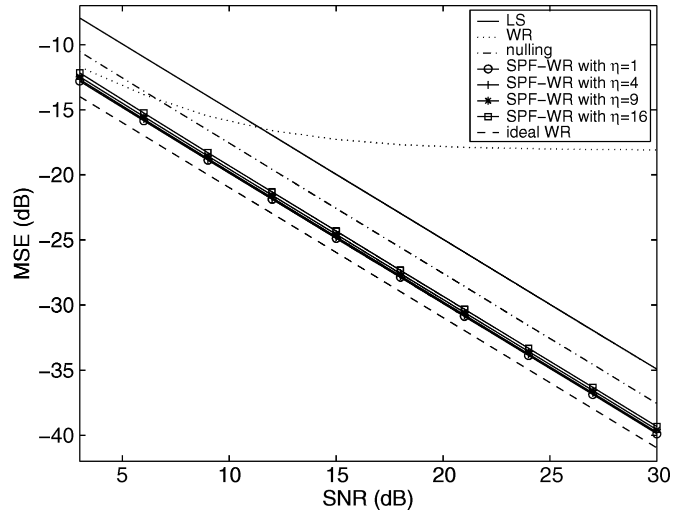


Fig. 2. MSE versus SNR.

Experiment 2: MSE Versus Pilot Length: Here, we investigate the effect of pilot length on the channel estimation performance. Fig. 3 shows the MSE plots of the channel estimate from 20 000 Monte Carlo runs for the transmission of one data frame of 1000 slots for an SNR of 10 dB, indicating a high estimation consistency of the proposed method with different values of η. Clearly, the performance improvement of the proposed method becomes more prominent compared to the WR-based method with the increase of the pilot length. For example, when the pilot length is increased to 100 from 30, the performance gain of the proposed method over the WR-based method is boosted to 3.4 dB from 2dB. On the other hand, the average number of slots for $\mathbf{X}_\mathrm{pos}$ and $\mathbf{X}_\mathrm{neg}$ is 5.3 and the maximum number is 8 in this case.

Considering that the proposed SPF-WR method requires both regular pilot and a certain number of slots for the transmission of $\mathbf{X}_\mathrm{pos}$ and $\mathbf{X}_\mathrm{neg}$ while the LS, the WR and the nulling-based methods use pilot only, we now compare these methods under the same transmission overhead. We reserve a maximum of 10 slots for the user specific data $\mathbf{X}_\mathrm{pos}$ and $\mathbf{X}_\mathrm{neg}$ in the proposed SPF-WR method, which is sufficient from our extensive simulations, even though the actual average number of slots used for $\mathbf{X}_\mathrm{pos}$ and $\mathbf{X}_\mathrm{neg}$ may be significantly less than 10. Then, we set the pilot length of the LS, WR and nulling methods to be equal to that of the SPF-WR method plus 10 slots, such that all the methods in comparison have the same length of nonuser data as overhead. Fig. 4 shows the MSE performance from 20 000 Monte Carlo runs for the transmission of one data frame of 1000

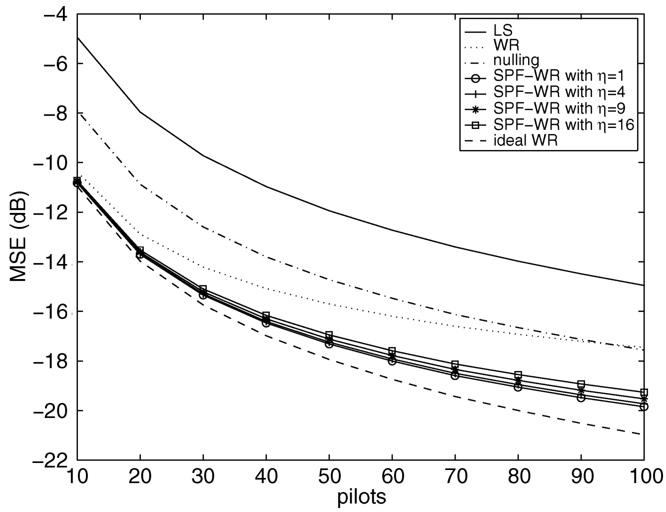


Fig. 3. MSE versus pilot length.

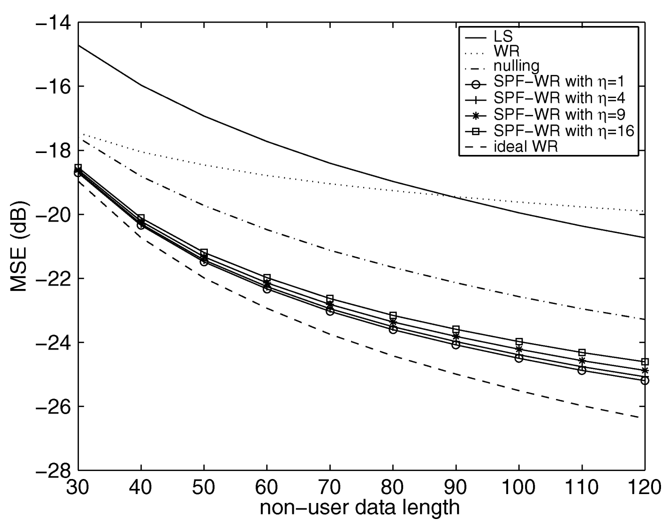


Fig. 4. MSE versus nonuser data length.

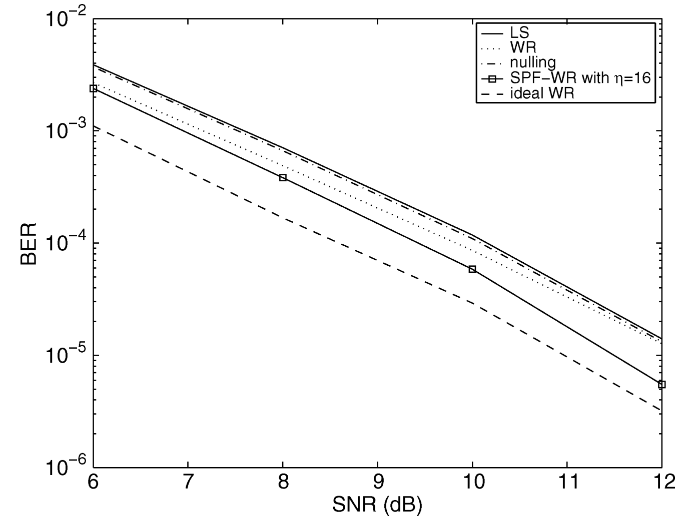


Fig. 5. BER versus SNR.

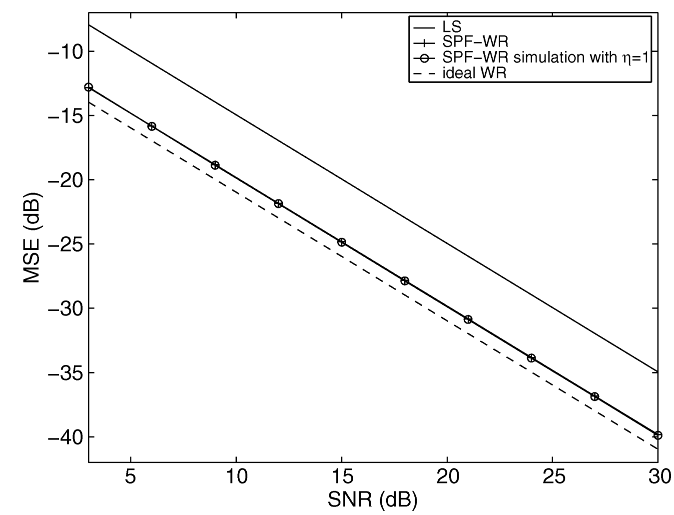


Fig. 6. Theoretical MSE versus SNR.

slots for an SNR of 15 dB. It is seen that, for the same transmission overhead, the proposed method significantly outperforms the three other methods.

Experiment 3: BER Versus SNR: Now, we look into the BER performance of the MIMO system by using the estimated channel matrix and an ordered vertical-Bell laboratories layered space time (V-BLAST) decoder. The simulation involves 50 000 Monte Carlo runs of the transmission of one data frame with 100 slots out of which 10 are used as pilots. Fig. 5 shows the BER performance versus the SNR. It is seen that the BER performance of the proposed method with $\eta = 16$ becomes much better than that of the WR-based method with the increase of SNR. In particular, when SNR is 12 dB, the BER of the proposed method is about 3.3 dB over that of the WR-based method.

Experiment 4: The Theoretical Value of MSE: In this experiment, we compare the proposed method with the ideal WR-based method as well as the LS method in terms of their theoretical MSE expressions. Using the same conditions as in *Experiment* 1 and *Experiment* 2, the three MSE plots calculated from (87), (88) and (89) are shown in Fig. 6 and Fig. 7, respectively. For comparison, the simulation results of the SPF-WR method with $\eta = 1$ obtained from *Experiment* 1 and *Experiment* 2 are also included in Fig. 6 and Fig. 7. Obviously, one can see that the theoretical MSE values are consistent with the simulation results, confirming the high accuracy of the derivation of the MSE expressions in Section IV.

VI. Conclusion

A new signal-perturbation-free WR-based semiblind approach has been proposed for MIMO channel estimation. To improve the performance of the WR-based method in the high SNR case, a new transmit structure, which contains user-specific data bearing the information of the signal perturbation matrix, has been proposed for the cancellation of the signal perturbation error at the receiver. Furthermore, a novel perturbation analysis theory has been developed for the SVD of the square matrix without noise subspace, facilitating the derivation

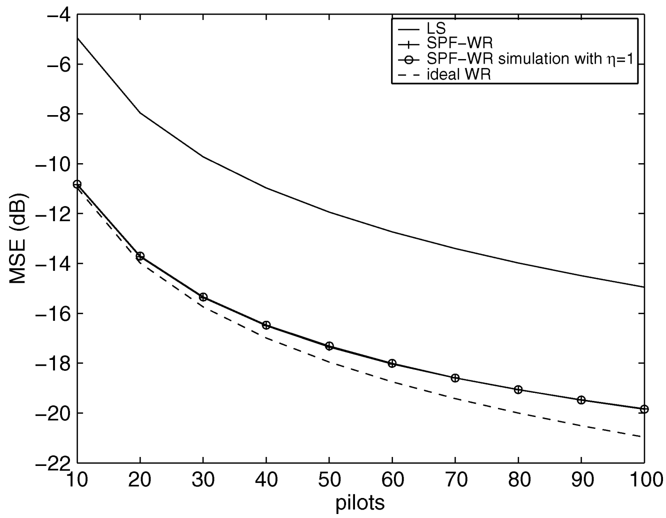


Fig. 7. Theoretical MSE versus pilot length.

of the closed-form expression for the MSE of the new SPF-WR method. It has been proved that a performance gain of nearly $2N_R/N_T$ is achieved by the new method over the LS method in terms of the MSE of the channel estimate. Simulation results have confirmed that, by using a small number of additional slots bearing the information of the autocorrelation matrix of the transmitted signal, a significant improvement in both the MSE of the channel estimate and the BER of the MIMO system is achieved by the proposed method over the WR-based method as well as the nulling-based approach at all SNR levels.

APPENDIX A
PROOF OF THEOREM 4.1

Since $\mathbf{Z}$ is a full rank square matrix, we have $\sigma_{z_1} \geq \sigma_{z_2} \geq \cdots \geq \sigma_{z_M} > 0$. By letting $\mathbf{\Psi}_1 \triangleq \Delta\mathbf{U}_Z^H\mathbf{U}_Z$ and $\mathbf{\Psi}_2 \triangleq \mathbf{V}_Z^H\Delta\mathbf{V}_Z$, $\mathbf{\Omega}$ can be rewritten as

$$\mathbf{\Omega} = \mathbf{\Psi}_1 + \mathbf{\Psi}_2. \tag{A1}$$

In what follows, we compute the elements of $\mathbf{\Omega}$, i.e., those of $\mathbf{\Psi}_1$ and $\mathbf{\Psi}_2$. Let us consider the nondiagonal elements first. Noting that $\hat{\mathbf{U}}_Z^H\hat{\mathbf{U}}_Z \approx \mathbf{U}_Z^H\mathbf{U}_Z + \mathbf{\Psi}_1^H + \mathbf{\Psi}_1$ and $\hat{\mathbf{U}}_Z^H\hat{\mathbf{U}}_Z = \mathbf{U}_Z^H\mathbf{U}_Z = \mathbf{I}$, one can get

$$\mathbf{\Psi}_1 + \mathbf{\Psi}_1^H \approx \mathbf{0} \tag{A2}$$

implying that $\mathbf{\Psi}_1$ is a skew-Hermitian matrix. Using (69) and neglecting the second-order error terms, one can prove that

$$\mathbf{U}_Z^H(\hat{\mathbf{Z}}\hat{\mathbf{Z}}^H - \mathbf{Z}\mathbf{Z}^H)\mathbf{U}_Z \approx \mathbf{\Psi}_1^H\mathbf{\Sigma}_Z^2 + \mathbf{\Sigma}_Z^2\mathbf{\Psi}_1 + 2\mathbf{\Sigma}_Z\Delta\mathbf{\Sigma}_Z. \tag{A3}$$

Using the SVD of $\mathbf{Z}$ and that of $\hat{\mathbf{Z}} = \mathbf{Z} + \Delta\mathbf{Z}$ into (A3) gives

$$\mathbf{\Psi}_1^H\mathbf{\Sigma}_Z^2 + \mathbf{\Sigma}_Z^2\mathbf{\Psi}_1 \approx \left(\mathbf{U}_Z^H\Delta\mathbf{Z}\mathbf{V}_Z - \Delta\mathbf{\Sigma}_Z\right)\mathbf{\Sigma}_Z + \mathbf{\Sigma}_Z\left(\mathbf{U}_Z^H\Delta\mathbf{Z}\mathbf{V}_Z - \Delta\mathbf{\Sigma}_Z\right)^H. \tag{A4}$$

By pre- and postmultiplying both sides of (A4) by $\mathbf{\Sigma}_Z^{-1}$, one can get

$$\mathbf{\Sigma}_Z^{-1}\mathbf{\Psi}_1^H\mathbf{\Sigma}_Z + \mathbf{\Sigma}_Z\mathbf{\Psi}_1\mathbf{\Sigma}_Z^{-1} \approx \mathbf{\Sigma}_Z^{-1}\mathbf{A} + \mathbf{A}^H\mathbf{\Sigma}_Z^{-1} \tag{A5}$$

where

$$\mathbf{A} \triangleq \mathbf{U}_Z^H\Delta\mathbf{Z}\mathbf{V}_Z - \Delta\mathbf{\Sigma}_Z. \tag{A6}$$

Using (A2) as well as (A5), one can obtain

$$\mathbf{\Psi}_1(i,j) \approx \frac{\sigma_{z_i}^{-1}\mathbf{A}(i,j) + \sigma_{z_j}^{-1}\mathbf{A}^*(j,i)}{\sigma_{z_i}\sigma_{z_j}^{-1} - \sigma_{z_j}\sigma_{z_i}^{-1}}, (i \neq j). \tag{A7}$$

We now consider the nondiagonal elements of $\mathbf{\Psi}_2$. Utilizing $\hat{\mathbf{V}}_Z^H\hat{\mathbf{V}}_Z \approx \mathbf{V}_Z^H\mathbf{V}_Z + \mathbf{\Psi}_2^H + \mathbf{\Psi}_2$ and noting that $\hat{\mathbf{V}}_Z^H\hat{\mathbf{V}}_Z = \mathbf{V}_Z^H\mathbf{V}_Z = \mathbf{I}$, one can show that $\mathbf{\Psi}_2$ is also skew-Hermitian. In order to determine the nondiagonal elements of $\mathbf{\Psi}_2$, we compute from (69)

$$\mathbf{V}_Z^H(\hat{\mathbf{Z}}^H\hat{\mathbf{Z}} - \mathbf{Z}^H\mathbf{Z})\mathbf{V}_Z \approx \mathbf{\Psi}_2\mathbf{\Sigma}_Z^2 + \mathbf{\Sigma}_Z^2\mathbf{\Psi}_2^H + 2\mathbf{\Sigma}_Z\Delta\mathbf{\Sigma}_Z. \tag{A8}$$

In a manner similar to the derivation of (A5), one can obtain

$$\mathbf{\Sigma}_Z^{-1}\mathbf{\Psi}_2\mathbf{\Sigma}_Z + \mathbf{\Sigma}_Z\mathbf{\Psi}_2^H\mathbf{\Sigma}_Z^{-1} \approx \mathbf{A}\mathbf{\Sigma}_Z^{-1} + \mathbf{\Sigma}_Z^{-1}\mathbf{A}^H \tag{A9}$$

which leads the nondiagonal elements of $\mathbf{\Psi}_2$ to

$$\mathbf{\Psi}_2(i,j) \approx \frac{\sigma_{z_j}^{-1}\mathbf{A}(i,j) + \sigma_{z_i}^{-1}\mathbf{A}^*(j,i)}{\sigma_{z_j}\sigma_{z_i}^{-1} - \sigma_{z_i}\sigma_{z_j}^{-1}}, (i \neq j). \tag{A10}$$

Using (A7) and (A10) into (A1), one can eventually obtain the nondiagonal elements of $\mathbf{\Omega}$ as

$$\mathbf{\Omega}(i,j) \approx \frac{\mathbf{A}^*(j,i) - \mathbf{A}(i,j)}{\sigma_{z_i} + \sigma_{z_j}}, (i \neq j). \tag{A11}$$

We now determine the diagonal elements of $\mathbf{\Omega}$. From (69), one can have

$$\Delta\mathbf{Z} - \mathbf{U}_Z\Delta\mathbf{\Sigma}_Z\mathbf{V}_Z^H \approx \mathbf{U}_Z\mathbf{\Sigma}_Z\Delta\mathbf{V}_Z^H + \Delta\mathbf{U}_Z\mathbf{\Sigma}_Z\mathbf{V}_Z^H. \tag{A12}$$

Premultiplying $\mathbf{U}_Z^H$ and postmultiplying $\mathbf{V}_Z$ on both sides of (A12) give

$$\mathbf{U}_Z^H\Delta\mathbf{Z}\mathbf{V}_Z - \Delta\mathbf{\Sigma}_Z \approx \mathbf{\Sigma}_Z\Delta\mathbf{V}_Z^H\mathbf{V}_Z + \mathbf{U}_Z^H\Delta\mathbf{U}_Z\mathbf{\Sigma}_Z. \tag{A13}$$

Using (A6) and (A13) and noting that $\mathbf{\Psi}_1 = \Delta\mathbf{U}_Z^H\mathbf{U}_Z$ and $\mathbf{\Psi}_2 = \mathbf{V}_Z^H\Delta\mathbf{V}_Z$, one can verify

$$\mathbf{\Psi}_2\mathbf{\Sigma}_Z + \mathbf{\Sigma}_Z\mathbf{\Psi}_1 \approx \mathbf{A}^H. \tag{A14}$$

Recall that $\mathbf{\Sigma}_Z$ is a real diagonal matrix and both $\mathbf{\Psi}_1$ and $\mathbf{\Psi}_2$ are skew-Hermitian which implies that their diagonal elements are imaginary. That is to say the diagonal elements of $\mathbf{\Psi}_2\mathbf{\Sigma}_Z + \mathbf{\Sigma}_Z\mathbf{\Psi}_1$ are imaginary. As a result, we have

$$\mathrm{diag}(\mathbf{\Psi}_2\mathbf{\Sigma}_Z + \mathbf{\Sigma}_Z\mathbf{\Psi}_1) \approx \mathrm{diag}\left(\frac{\mathbf{A}^H - \mathbf{A}}{2}\right). \tag{A15}$$

Using (A1) and (A15), the diagonal elements of $\mathbf{\Omega}$ are readily given by

$$\mathbf{\Omega}(i,i) \approx \frac{\mathbf{A}^*(i,i) - \mathbf{A}(i,i)}{2\sigma_{z_i}}. \tag{A16}$$

From (A6), (A11), and (A16), the theorem is proved.

APPENDIX B
DERIVATION OF $\mathbf{G}_0$, $\mathbf{G}_1$, AND $\mathbf{G}_2$

To derive the expression of G_0 given by (81), we first simplify the calculation of $\mathrm{E}\{\mathrm{vec}(\Delta\mathbf{R}_{xv}^H)\mathrm{vec}^H(\Delta\mathbf{R}_{xv}^H)\}$ by using the statistical property of the signal as well as the noise. Using (7), one can get

$$\mathrm{vec}\left(\Delta\mathbf{R}_{xv}^H\right) = \frac{1}{N}\sum_{n=1}^{N}\{[\mathbf{x}^*(n)\otimes\mathbf{I}_{N_R}]\,\mathrm{vec}\,[\mathbf{v}(n)]\} \quad \text{(B1)}$$

which leads to

$$\begin{aligned}&\mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv}^H\right)\mathrm{vec}^H\left(\Delta\mathbf{R}_{xv}^H\right)\right\}\\ &= \frac{1}{N^2}\mathrm{E}\left\{\sum_{n_1=1}^{N}\sum_{n_2=1}^{N}[\mathbf{x}^*(n_1)\otimes\mathbf{I}_{N_R}]\,\mathrm{E}\left[\mathbf{v}(n_1)\mathbf{v}^H(n_2)\right]\right.\\ &\qquad \left.\times\left[\mathbf{x}^T(n_2)\otimes\mathbf{I}_{N_R}\right]\right\}. \quad \text{(B2)}\end{aligned}$$

Noting that $\mathrm{E}[\mathbf{v}(n_1)\mathbf{v}^H(n_2)] = \delta(n_1-n_2)\sigma_v^2\mathbf{I}$ and $\mathrm{E}[\mathbf{x}(n)\mathbf{x}^H(n)] = \mathbf{I}$, we can obtain

$$\begin{aligned}&\mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv}^H\right)\mathrm{vec}^H\left(\Delta\mathbf{R}_{xv}^H\right)\right\}\\ &= \frac{\delta_v^2}{N}\left\{\mathrm{E}\left[\frac{1}{N}\sum_{n=1}^{N}\mathbf{x}^*(n)\mathbf{x}^T(n)\right]\otimes\mathbf{I}_{N_R}\right\}\\ &= \frac{1}{N}\delta_x^2\delta_v^2\mathbf{I}_{N_RN_T}. \quad \text{(B3)}\end{aligned}$$

Substituting (B3) into (81) gives

$$G_0 = \frac{N_RN_T}{N}\sigma_v^2 \quad \text{(B4)}$$

We now turn to the derivation of G_1. Letting

$$\boldsymbol{\Upsilon} \triangleq \mathrm{E}\left\{\left[\mathrm{vec}(\boldsymbol{\Gamma}_Q)\mathrm{vec}^H(\boldsymbol{\Gamma}_Q)\right]\circ\left[\mathrm{vec}(\boldsymbol{\Pi}^H)\mathrm{vec}^H(\boldsymbol{\Pi}^H)\right]\right\} \quad \text{(B5)}$$

G_1 given by (82) can be rewritten as

$$G_1 = \mathrm{Trace}\left\{\left[\mathbf{V}_S^*\otimes(\mathbf{U}_S\boldsymbol{\Sigma}_S)\right]\boldsymbol{\Upsilon}\left[\mathbf{V}_S^*\otimes(\mathbf{U}_S\boldsymbol{\Sigma}_S)\right]^H\right\}. \quad \text{(B6)}$$

In order to determine G_1, we first compute $\mathrm{E}[\mathrm{vec}(\boldsymbol{\Pi}^H)\mathrm{vec}^H(\boldsymbol{\Pi}^H)]$. From (76), one can get

$$\begin{aligned}\boldsymbol{\Pi} &= \mathbf{V}_S^H\Delta\mathbf{R}_{xv}\mathbf{U}_S\boldsymbol{\Sigma}_S + \boldsymbol{\Sigma}_S\mathbf{U}_S^H\Delta\mathbf{R}_{xv,P}^H\mathbf{V}_S\\ &\quad -\boldsymbol{\Sigma}_S\mathbf{U}_S^H\Delta\mathbf{R}_{xv}^H\mathbf{V}_S - \mathbf{V}_S^H\Delta\mathbf{R}_{xv,P}\mathbf{U}_S\boldsymbol{\Sigma}_S. \quad \text{(B7)}\end{aligned}$$

Obviously, the computation of $\mathrm{E}[\mathrm{vec}(\boldsymbol{\Pi}^H)\mathrm{vec}^H(\boldsymbol{\Pi}^H)]$ involves a total of 16 terms. It can be shown that computing these terms requires the auto-correlation as well as the cross-correlation matrices of $\mathrm{vec}(\Delta\mathbf{R}_{xv})$, $\mathrm{vec}(\Delta\mathbf{R}_{xv}^H)$, $\mathrm{vec}(\Delta\mathbf{R}_{xv,P})$ and $\mathrm{vec}(\Delta\mathbf{R}_{xv,P}^H)$. Following a manner similar to the derivation (B3), one can derive

$$\begin{aligned}&\mathrm{E}\left\{\mathrm{vec}(\Delta\mathbf{R}_{xv})\mathrm{vec}^H(\Delta\mathbf{R}_{xv})\right\}\\ &= \mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv}^H\right)\mathrm{vec}^H\left(\Delta\mathbf{R}_{xv}^H\right)\right\}\\ &= \frac{\sigma_v^2}{N}\mathbf{I}_{N_RN_T} \quad \text{(B8)}\\ &\mathrm{E}\left\{\mathrm{vec}(\Delta\mathbf{R}_{xv,P})\mathrm{vec}^H(\Delta\mathbf{R}_{xv,P})\right\}\\ &= \mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv,P}^H\right)\mathrm{vec}^H\left(\Delta\mathbf{R}_{xv,P}^H\right)\right\}\\ &= \frac{\sigma_v^2}{K}\mathbf{I}_{N_RN_T} \quad \text{(B9)}\\ &\mathrm{E}\left\{\mathrm{vec}(\Delta\mathbf{R}_{xv})\mathrm{vec}^H(\Delta\mathbf{R}_{xv,P})\right\}\\ &= \mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv}^H\right)\mathrm{vec}^H\left(\Delta\mathbf{R}_{xv,P}^H\right)\right\}\\ &= \frac{\sigma_v^2}{N}\mathbf{I}_{N_RN_T} \quad \text{(B10)}\\ &\mathrm{E}\left\{\mathrm{vec}(\Delta\mathbf{R}_{xv,P})\mathrm{vec}^H\left(\Delta\mathbf{R}_{xv,P}^H\right)\right\}\\ &= \mathrm{E}\left\{\mathrm{vec}(\Delta\mathbf{R}_{xv})\mathrm{vec}^H\left(\Delta\mathbf{R}_{xv}^H\right)\right\} = \mathbf{0}, \quad \text{(B11)}\\ &\mathrm{E}\left\{\mathrm{vec}(\Delta\mathbf{R}_{xv,P})\mathrm{vec}^H\left(\Delta\mathbf{R}_{xv}^H\right)\right\}\\ &= \mathrm{E}\left\{\mathrm{vec}(\Delta\mathbf{R}_{xv})\mathrm{vec}^H\left(\Delta\mathbf{R}_{xv,P}^H\right)\right\} = \mathbf{0}. \quad \text{(B12)}\end{aligned}$$

Then, the first term in the computation of $\mathrm{E}[\mathrm{vec}(\boldsymbol{\Pi}^H)\mathrm{vec}^H(\boldsymbol{\Pi}^H)]$ can be calculated using (B8) as

$$\begin{aligned}&\mathrm{E}\left[\mathrm{vec}\left(\mathbf{V}_S^H\Delta\mathbf{R}_{xv}\mathbf{U}_S\boldsymbol{\Sigma}_S\right)\mathrm{vec}^H\left(\mathbf{V}_S^H\Delta\mathbf{R}_{xv}\mathbf{U}_S\boldsymbol{\Sigma}_S\right)\right]\\ &= \frac{\sigma_v^2}{N}\left[\left(\boldsymbol{\Sigma}_S\mathbf{U}_S^T\right)\otimes\mathbf{V}_S^H\right]\left[(\mathbf{U}_S^*\boldsymbol{\Sigma}_S)\otimes\mathbf{V}_S\right]\\ &= \frac{\sigma_v^2}{N}\left(\boldsymbol{\Sigma}_S^2\otimes\mathbf{I}_{N_T}\right). \quad \text{(B13)}\end{aligned}$$

In a similar manner, all the other terms can be computed using (B8) to (B12). By adding all the 16 terms together, we obtain

$$\begin{aligned}&\mathrm{E}\left[\mathrm{vec}(\boldsymbol{\Pi}^H)\mathrm{vec}^H(\boldsymbol{\Pi}^H)\right]\\ &= \frac{(N-K)\sigma_v^2}{KN}\left(\boldsymbol{\Sigma}_S^2\otimes\mathbf{I}_{N_T} + \mathbf{I}_{N_T}\otimes\boldsymbol{\Sigma}_S^2\right). \quad \text{(B14)}\end{aligned}$$

Utilizing (74) and (B14), one can easily verify that $\boldsymbol{\Upsilon}$ is a diagonal matrix with diagonal elements being given by

$$\begin{aligned}\boldsymbol{\rho}(l = i + N_T(j-1)) &= \frac{(N-K)\sigma_v^2}{KN}\frac{1}{\sigma_{S_i}^2+\sigma_{S_j}^2}\\ &\qquad (i,j = 1,2,\cdots,N_T). \quad \text{(B15)}\end{aligned}$$

Letting $\mathbf{u}_{S_i}$ and $\mathbf{v}_{S_j}$ be the ith and jth column vectors of $\mathbf{U}_S$ and $\mathbf{V}_S$, respectively, one can find the lth column vector of $\mathbf{V}_S^*\otimes(\mathbf{U}_S\boldsymbol{\Sigma}_S)$ as $\sigma_{S_i}\mathbf{v}_{S_j}^*\otimes\mathbf{u}_{S_i}$. Therefore, (B6) can be expressed as

$$\begin{aligned}G_1 &= \frac{(N-K)\sigma_v^2}{KN}\sum_{j=1}^{N_T}\sum_{i=1}^{N_T}\frac{\sigma_{S_i}^2}{\sigma_{S_i}^2+\sigma_{S_j}^2}\\ &\qquad\times\mathrm{Trace}\left[\left(\mathbf{v}_{S_j}^*\otimes\mathbf{u}_{S_i}\right)\left(\mathbf{v}_{S_j}^*\otimes\mathbf{u}_{S_i}\right)^H\right]. \quad \text{(B16)}\end{aligned}$$

Since $\mathrm{Trace}[(\mathbf{v}_{S_j}^*\otimes\mathbf{u}_{S_i})(\mathbf{v}_{S_j}^*\otimes\mathbf{u}_{S_i})^H] = \|\mathbf{v}_{S_j}^*\otimes\mathbf{u}_{S_i}\|_F^2 = 1$, from (B16), the value of G_1 can be found to be

$$G_1 = \frac{(N-K)\sigma_v^2}{KN}\sum_{j=1}^{N_T}\sum_{i=1}^{N_T}\frac{\sigma_{S_i}^2}{\sigma_{S_i}^2+\sigma_{S_j}^2} = \frac{N_T^2(N-K)\sigma_v^2}{2KN}. \quad \text{(B17)}$$

Now we determine the value of G_2. The first two terms of the RHS of (83) can be rewritten as

$$\begin{aligned}&\mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv}^H\right)\left[\mathrm{vec}(\boldsymbol{\Gamma}_Q)\circ\mathrm{vec}(\boldsymbol{\Pi}^H)\right]^H\right\}\\ &= \left[\mathbf{1}\cdot\mathrm{vec}(\boldsymbol{\Gamma}_Q)^H\right]\circ\mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv}^H\right)\mathrm{vec}^H(\boldsymbol{\Pi}^H)\right\}. \quad \text{(B18)}\end{aligned}$$

Using (B7), (B11) and (B12), we have

$$\mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv}^{H}\right)\mathrm{vec}^{H}(\mathbf{\Pi}^{H})\right\} = \mathrm{E}\left[\mathrm{vec}\left(\Delta\mathbf{R}_{xv}^{H}\right)\mathrm{vec}^{H}\left(\Delta\mathbf{R}_{xv,P}^{H}\right)\right]\left[\mathbf{V}_{S}^{T}\otimes\left(\mathbf{\Sigma}_{S}\mathbf{U}_{S}^{H}\right)\right]^{H} - \mathrm{E}\left[\mathrm{vec}\left(\Delta\mathbf{R}_{xv}^{H}\right)\mathrm{vec}^{H}\left(\Delta\mathbf{R}_{xv}^{H}\right)\right] \times\left[\mathbf{V}_{S}^{T}\otimes\left(\mathbf{\Sigma}_{S}\mathbf{U}_{S}^{H}\right)\right]^{H}. \quad \text{(B19)}$$

Using (B8) and (B10) into (B19) gives

$$\mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv}^{H}\right)\mathrm{vec}^{H}(\mathbf{\Pi}^{H})\right\} = \frac{\sigma_v^2}{N}\left[\mathbf{V}_{S}^{T}\otimes\left(\mathbf{\Sigma}_{S}\mathbf{U}_{S}^{H}\right)-\mathbf{V}_{S}^{T}\otimes\left(\mathbf{\Sigma}_{S}\mathbf{U}_{S}^{H}\right)\right]^{H}=\mathbf{0}. \quad \text{(B20)}$$

Therefore, from (83), (B18), and (B20), we have $G_2 = 0$.

ACKNOWLEDGMENT

The authors would like to thank the Associate Editor and the anonymous reviewers for their careful review of our paper. Their valuable comments and suggestions have greatly helped improve the presentation of our paper.

REFERENCES

[1] A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bolcskei, "An overview of MIMO communications-a key to gigabit wireless," *Proc. IEEE*, vol. 92, no. 2, pp. 198–218, 2004.
[2] G. Stuber, J. R. Barry, S. W. Mclaughlin, Y. Li, M. A. Ingram, and T. G. Pratt, "Broadband MIMO-OFDM wireless communications," *Proc. IEEE*, vol. 92, no. 2, pp. 271–294, 2004.
[3] M. Biguesh and A. B. Gershman, "MIMO channel estimation: Optimal training and tradeoffs between estimation techniques," in *Proc. IEEE Int. Conf. Commun.*, 2004, vol. 5, pp. 2658–2662.
[4] V. Barroso and J. Xavier, "Blind identification of MIMO channels: A closed form solution based on second order statistics," in *Proc. 33rd Asilomar Conf. Signals, Syst., Comput.*, 1999, vol. 1, pp. 70–74.
[5] J. K. Tugnait, "On linear predictors for MIMO channels and related blind identification and equalization," *IEEE Signal Process. Lett.*, vol. 5, no. 11, pp. 289–291, 1998.
[6] J. K. Tugnait, "Blind estimation and equalization of MIMO channels via multidelay whitening," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 8, pp. 1507–1519, 2001.
[7] Z. Ding and L. Qiu, "Blind MIMO channel identification from second order statistics using rank deficient channel convolution matrix," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 535–544, 2003.
[8] Z. Ding and D. B. Ward, "Subspace approach to blind and semiblind channel estimation for space-time block codes," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 357–362, 2005.
[9] V. Buchoux, O. Cappe, E. Moulines, and A. Gorokhov, "On the performance of semiblind subspace-based channel estimation," *IEEE Trans. Signal Process.*, vol. 48, no. 6, pp. 1750–1759, 2000.
[10] A. Medles, D. T. M. Slock, and E. De Carvalho, "Linear prediction based semiblind estimation of MIMO FIR channels," in *Proc. IEEE Third Workshop on Signal Process. Adv. Wireless Commun.*, 2001, pp. 58–61.
[11] A. Medles and D. T. M. Slock, "Augmenting the training sequence part in semiblind estimation for MIMO channels," in *Proc. 37th Asilomar Conf. Signals, Syst. Comput.*, 2003, vol. 2, pp. 1825–1829.
[12] Y. Zeng and T. Ng, "A semiblind channel estimation method for multiuser multiantenna OFDM systems," *IEEE Trans. Signal Process.*, vol. 52, no. 5, pp. 1419–1429, 2004.
[13] J. Choi, "Equalization and semiblind channel estimation for space-time block coded signals over a frequency-selective fading channel," *IEEE Trans. Signal Process.*, vol. 52, no. 3, pp. 774–785, 2004.
[14] Y. H. Zeng, S. D. Ma, and T. S. Ng, "Semiblind estimation of channels and symbols for asynchronous MIMO systems," *Inst. Elect. Eng. Proc.-Commun.*, vol. 152, no. 6, pp. 883–889, 2005.
[15] A. K. Jagannatham and B. D. Rao, "Whitening-rotation-based semiblind MIMO channel estimation," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 861–869, 2006.
[16] A. K. Jagannatham and B. D. Rao, "Constrained ML algorithms for semiblind MIMO channel estimation," in *Proc. IEEE Global Telecommun. Conf.*, 2004, vol. 4, pp. 2475–2479.
[17] A. K. Jagannatham and B. D. Rao, "Cramer-Rao lower bound for constrained complex parameters," *IEEE Signal Process. Lett.*, vol. 11, no. 11, pp. 875–878, 2004.
[18] C. R. Murthy, A. K. Jagannatham, and B. D. Rao, "Training-based and semiblind channel estimation for MIMO systems with maximum ratio transmission," *IEEE Trans. Signal Process.*, vol. 54, no. 7, pp. 2546–2558, 2006.
[19] F. Wan, W.-P. Zhu, and M. N. S. Swamy, "Perturbation analysis of subspace-based semiblind MIMO channel estimation approaches," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, 2008, pp. 129–132.
[20] A. Medles and D. T. M. Slock, "Semiblind channel estimation for MIMO spatial multiplexing systems," in *Proc. IEEE Veh. Technol. Conf.*, 2001, vol. 2, pp. 1240–1244.
[21] F. Li and R. J. Vaccaro, "Analysis of min-norm and MUSIC with arbitrary array geometry," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 26, no. 6, pp. 976–985, 1990.
[22] F. Li and R. J. Vaccaro, "Analytical performance prediction of subspace-based algorithms for DOA estimation," in *SVD and Signal Processing, II: Algorithms, Analysis and Applications*, R. J. Vaccaro, Ed. New York: Elsevier, 1991, pp. 243–260.
[23] F. Li, H. Liu, and R. J. Vaccaro, "Performance analysis for DOA estimation algorithms: Unification, simplification, and observations," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 29, no. 4, pp. 1170–1184, 1993.
[24] G. W. Stewart, "Perturbation theory for the singular value decomposition," in *SVD and Signal Processing, II: Algorithms, Analysis and Applications*, R. J. Vaccaro, Ed. New York: Elsevier, 1991, pp. 99–109.
[25] B. Friedlander and A. J. Weiss, "On the second-order statistics of the eigenvectors of sample covariance matrices," *IEEE Trans. Signal Process.*, vol. 46, no. 11, pp. 3136–3139, 1998.
[26] Z. Xu, "Perturbation analysis for subspace decomposition with applications in subspace-based algorithms," *IEEE Trans. Signal Process.*, vol. 50, no. 11, pp. 2820–2830, 2002.
[27] Z. Xu, "On the second-order statistics of the weighted sample covariance matrix," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 527–534, 2003.
[28] Z. Xu, "Statistical performance of a data-based covariance estimator," *IEEE Trans. Veh. Technol.*, vol. 53, no. 3, pp. 939–943, 2004.
[29] F. Wan, W.-P. Zhu, and M. N. S. Swamy, "A semiblind channel estimation approach for MIMO-OFDM systems," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 2821–2834, 2008.
[30] K.-B. Yu, "Recursive updating the eigenvalue decomposition of a covariance matrix," *IEEE Trans. Signal Process.*, vol. 39, no. 5, pp. 1136–1145, May 1991.
[31] I. Barhumi, G. Leus, and M. Moonen, "Optimal training design for MIMO OFDM systems in mobile wireless channels," *IEEE Trans. Signal Process.*, vol. 51, no. 6, pp. 1615–1624, 2003.
[32] X. Ma, L. Yang, and G. B. Giannakis, "Optimal training for MIMO frequency-selective fading channels," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 453–466, 2005.

Feng Wan received the B.E. and M.E. degrees from Wuhan University and Institute of Electronics, Chinese Academy of Sciences, in 2000 and 2003, respectively, all in electrical engineering.

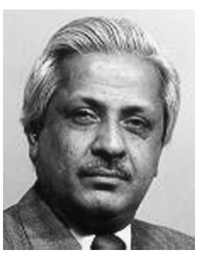
He is currently pursuing the doctoral degree at Concordia University, Montreal, Canada. Since 2004, he has been a research assistant with Concordia University. His research interests are in the area of digital signal processing, statistical estimation theory and its applications in MIMO, OFDM, and smart antenna wireless communications.

Mr. Wan was awarded the Doctoral Thesis Completion Award, the Graduate Fellowship, and the International Tuition Fee Remission Award from Concordia University.

Wei-Ping Zhu (SM'97) received the B.E. and M.E. degrees from Nanjing University of Posts and Telecommunications, and the Ph.D. degree from Southeast University, Nanjing, China, in 1982, 1985, and 1991, respectively, all in electrical engineering.

He was a Postdoctoral Fellow from 1991 to 1992 and a Research Associate from 1996 to 1998 with the Department of Electrical and Computer Engineering, Concordia University, Montreal, Canada. During 1993-1996, he was an Associate Professor with the Department of Information Engineering, Nanjing University of Posts and Telecommunications. From 1998 to 2001, he worked in hi-tech companies in Ottawa, Canada, including Nortel Networks and SR Telecom Inc. Since July 2001, he has been with Concordia's Electrical and Computer Engineering Department as an Associate Professor. His research interests include the areas of digital filters, filter banks, speech and image processing, and signal processing for wireless communication.

Dr. Zhu was an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS PART I: FUNDAMENTAL THEORY AND APPLICATIONS from October 2001 to December 2003. He currently serves as an Associate Editor of *Circuits, Systems and Signal Processing*.



M. N. S. Swamy (S'59–M'62–SM'74–F'80) received the B.Sc. (Hons.) degree in mathematics from Mysore University, India, in 1954, the Diploma degree in electrical communication engineering from the Indian Institute of Science, Bangalore, in 1957, and the M.Sc. and Ph.D. degrees in electrical engineering from the University of Saskatchewan, Saskatoon, Canada, in 1960 and 1963, respectively. In August 2001, he was awarded a Doctor of Science (*Honoris Causa*) degree in engineering by Ansted University "in recognition of his exemplary contributions to the research in Electrical and Computer Engineering and to Engineering Education, as well as his dedication to the promotion of Signal Processing and Communications Applications."

He is presently a Research Professor and the Director of the Center for Signal Processing and Communications in the Department of Electrical and Computer Engineering, Concordia University, Montreal, QC, Canada, where he served as the Founding Chair of the Department of Electrical Engineering from 1970 to 1977, and Dean of Engineering and Computer Science from 1977 to 1993, during which time he developed the Faculty into a research-oriented Faculty from what was primarily an undergraduate one. Since July 2001, he holds the Concordia Chair (Tier I) in Signal Processing. He has also taught in the Electrical Engineering Department, Technical University of Nova Scotia, Halifax, and the University of Calgary, Calgary, as well as in the Department of Mathematics, University of Saskatchewan. He has published extensively in the areas of number theory, circuits, systems and signal processing, and holds five patents. He is the coauthor of two book chapters and four books: *Graphs, Networks and Algorithms* (New York: Wiley, 1981; Russian translation: Moscow, Russia: Mir, 1984; Chinese translation: Beijing, China: Education, 1987), *Graphs: Theory and Algorithms* (New York, Wiley, 1992), *Switched Capacitor Filters: Theory, Analysis and Design* (Prentice-Hall Int., U.K. Ltd., 1995) and *Neural Networks in a Softcomputing Framework* (New York: Springer, 2006). He was a founding member of Micronet from its inception in 1999 as a Canadian Network of Centers of Excellence until its expiration in 2004, and also its coordinator for Concordia University. Recently, Concordia University has instituted the *M.N.S. Swamy Research Chair in Electrical Engineering* as a recognition of his research contributions.

Dr. Swamy is a Fellow of a number of professional societies including the Institute of Electrical Engineers (United Kingdom), the Engineering Institute of Canada, the Institution of Engineers (India), and the Institution of Electronic and Telecommunication Engineers (India). He has served the IEEE in various capacities such as the President-Elect in 2003, President in 2004, Past-President in 2005, Vice President (Publications) during 2001-2002, Vice-President in 1976, Editor-in-Chief of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS-I from June 1999 to December 2001, Associate Editor of the TRANSACTIONS ON CIRCUITS AND SYSTEMS from June 1985 to May 1987, Program Chair for the 1973 IEEE CAS Symposium, General Chair for the 1984 IEEE CAS Symposium, Vice-Chair for the 1999 IEEE Circuits and Systems (CAS) Symposium, and a member of the Board of Governors of the CAS Society. He is the recipient of many IEEE-CAS Society awards, including the Education Award in 2000, Golden Jubilee Medal in 2000, and the 1986 Guillemin-Cauer Best Paper Award.