A Semiblind Channel Estimation Approach for MIMO–OFDM Systems

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

***Abstract*—In this paper, a very efficient semiblind approach that uses a training-based least square criterion along with a blind constraint is proposed for multiple-input–multiple-output–orthogonal frequency-division multiplexing (MIMO–OFDM) channel estimation. The blind constraint is derived from the linear prediction of the received MIMO–OFDM signal and is used in conjunction with a weighting factor in the semiblind cost function. An appealing scheme for the determination of the weighting factor is presented as a part of the proposed approach. A perturbation analysis of the proposed method is conducted to justify the superiority of the semiblind solution and to obtain a closed-form expression for the mean square error (MSE) of the blind constraint, further facilitating the calculation of the weighting factor. The proposed method is validated through computer simulation-based experimentations, showing a very high estimation accuracy of the semiblind solution in terms of the MSE of the channel estimate.**

***Index Terms*—Channel estimation, multiple-input–multiple-output (MIMO) linear prediction, MIMO–orthogonal frequency-division multiplexing (MIMO–OFDM), perturbation analysis, semiblind.**

I. Introduction

THE multiple-input–multiple-output–orthogonal frequency-division multiplexing (MIMO–OFDM) technology has been considered as a strong candidate for the next generation wireless communication systems. Using multiple transmit as well as receive antennas, a MIMO–OFDM system gives a high data rate without increasing the total transmission power or bandwidth compared to a single antenna system. Further, the frequency-selective problem that exists in a conventional wireless system can be well solved by the OFDM technique in the MIMO–OFDM system. On the other hand, the performance of MIMO–OFDM systems depends largely upon the availability of the knowledge of the channel. It has been proved [1] that when the channel is Rayleigh fading and perfectly known to the receiver, the capacity of a MIMO–OFDM system grows linearly with the number of transmit or receive antennas, whichever is less. Therefore, an accurate estimation of the wireless channel is of crucial importance to MIMO–OFDM systems.

A considerable number of channel estimation methods have already been proposed for MIMO–OFDM systems. They can broadly be categorized into three classes, namely, the training-based method, the blind method, and the semiblind one as a combination of the first two methods. First, the training-based methods employ known training signals to render an accurate channel estimation [2]–[5]. One of the most efficient training-based methods is the least squares (LS) algorithm, for which an optimum pilot design scheme has been given in [2]. When the full or partial information of the channel correlation is known, a better channel estimation performance can be achieved via some minimum mean square error (MMSE) methods [3]. By using decision feedback symbols, the Takagi–Sugeno–Kang (TSK) fuzzy approach proposed in [4] can achieve a performance similar to the MMSE methods while with a low complexity. In contrast to training-based methods, blind MIMO–OFDM channel estimation algorithms, such as those proposed in [6] and [7], often exploit the second-order stationary statistics, correlative coding, or other properties to give a better spectral efficiency. With a small number of training symbols, a semiblind method has been proposed in [8] to estimate the channel ambiguity matrix for space-time coded OFDM systems.

It is worth pointing out that most of the existing blind and semiblind MIMO–OFDM channel estimation methods are based on the second-order statistics of a long vector whose size is equal to or larger than the number of subcarriers. To estimate the correlation matrix reliably, they need a large number of OFDM symbols, which is not suitable for fast time-varying channels. In addition, because the matrices involved in these algorithms are of huge size, their computational complexity is extremely high. In contrast, a linear prediction-based semiblind algorithm that is based on the second-order statistics of a short vector with a size only slightly larger the channel length has been found much more efficient than the conventional LS methods for the estimation of frequency-selective MIMO channels [9]–[11]. In this paper, we will extend the linear prediction-based semiblind approach to the channel estimation of MIMO–OFDM systems.

Linear prediction has been widely used in blind MIMO channel estimation and equalization [9], [12]–[17], where the key idea is to represent the received MIMO signal as a finite-order autoregressive (AR) process under the assumption that the transmitted signals are uncorrelated in time [12]. Based on the AR process, a linear prediction filter can be obtained to solve a second-order deconvolution problem for channel equalization. By combining the linear prediction with a higher order statistics (HOS) or the weighted LS method, some blind channel estimation algorithms have been derived [12], [17]. However, these algorithms require a large number of signal samples and moreover, they are not robust. Medles *et al.* have proposed a semiblind algorithm by incorporating a blind criterion derived from the linear prediction into a training-based LS cost function [9]–[11], leading to a closed-form expression for the estimation of the MIMO channel

Manuscript received June 14, 2007; revised November 27, 2007. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Petr Tichavsky. This work was supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada. Part of this work was presented at the IEEE International Symposium on Circuits and Systems, New Orleans, LA, May 2007.

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

Digital Object Identifier 10.1109/TSP.2008.917354

response. It has been shown in their papers that the semiblind method provides a much better channel estimation performance over the pure training-based LS method. However, the superiority of the proposed semiblind method has not been theoretically justified. Moreover, the weighting factor employed to trade off the LS and the blind criteria has not been appropriately determined. As such, the resulting channel estimation performance, though better than that of the LS method, could be further improved. It should also be noted that this semiblind method for MIMO channel estimation cannot be directly applied to MIMO–OFDM systems due to different signal models in the two systems. In MIMO systems, both information data and pilots are generated and transmitted in the time domain, whereas in MIMO–OFDM systems, signals are first generated in the frequency domain and then converted to the time domain via the inverse discrete Fourier transform (IDFT). As a consequence, the training-based LS cost function for MIMO channel estimation cannot be used for MIMO–OFDM systems. Also, the uncorrelation among the time-domain MIMO signals, which is desirable for linear prediction, may not be available in MIMO–OFDM systems. Therefore, a lot of work on the formulation of MIMO–OFDM signals has to be done first to develop a new semiblind estimation solution.

In this paper, we propose a channel estimation algorithm for MIMO–OFDM systems by employing the aforementioned semiblind strategy. First, a new training-based LS criterion for MIMO–OFDM channel estimation is obtained through proper matrix formulations. Then, by proving that the transmitted time-domain MIMO–OFDM signals are uncorrelated, we validate the use of the linear prediction for the formulation of the blind criterion. By employing the LS and the blind criterion, a semiblind channel estimation solution is obtained. As a part of the semiblind approach, we also propose an appealing scheme for choosing the weighting factor, which is shown to be very efficient according to our extensive computer simulations.

The second part of this paper deals with the analysis of the MIMO linear prediction and the corresponding blind constraint. We apply the perturbation theory, which has been successfully used in analyzing the antenna-array-based signal processing algorithms [18]–[23], to the analysis of the proposed MIMO linear-prediction-based method, justifying the superiority of our semiblind solution over the blind algorithms [12], [24], [25]. The perturbation analysis of the blind constraint also leads to a novel closed-form expression for the mean square error (MSE) of the blind estimation that is essential to the calculation of the weighting factor for the semiblind estimation solution.

The rest of this paper is organized as follows. Section II gives a brief review of the MIMO linear prediction and a related semiblind algorithm for MIMO channel estimation. Section III presents a new semiblind approach for the estimation of MIMO–OFDM channels, including the formulation of the semiblind estimation problem, derivation of the blind constraint from the linear prediction, and development of an appealing scheme for the determination of the weighting factor. Section IV conducts a perturbation analysis of the linear-prediction-based semiblind method, justifying the superior estimation performance of the semiblind solution over a pure blind method. The analysis also yields a closed-form expression for the MSE of the blind estimation part of the proposed approach, facilitating the calculation of the weighting factor. Section V comprises a number of experimentations validating the proposed method, showing significant advantage of the semiblind solution over the LS method in terms of the MSE of the channel estimate. Finally, Section VI highlights some of the distinct features of the proposed approach.

Throughout this paper, we adopt the following notations:

$\dagger$	pseudoinverse;
$\otimes$	Kronecker product;
T	transpose;
H	complex conjugate transpose;
$*$	linear convolution;
$\circledast$	circular convolution;
$\Vert\ \Vert_F$	Frobenius norm;
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. Brief Overview of Linear Prediction for MIMO Channel Estimation

Consider a MIMO system with N_T transmit and N_R ($>N_T$) receive antennas. The MIMO channel can be characterized by an array of L-tap finite-impulse response (FIR) filters given by a number of $N_R \times N_T$ matrices $\mathbf{H}(n)$ ($n = 0, 1, \ldots, L-1$), whose (i_R, i_T)th element $h_{i_T, i_R}(n)$ represents the channel response from the i_Tth transmit antenna to the i_Rth receive antenna. Given the transmitted signal vector $\mathbf{x}(n) \triangleq [x_1(n), \ldots, x_{N_T}(n)]^T$, the received signal vector can be written as $\mathbf{y}(n) \triangleq [y_1(n), \ldots, y_{N_R}(n)]^T$, with its element being given by

$$y_{i_R}(n) = \sum_{i_T=1}^{N_T} h_{i_R, i_T}(n) * x_{i_T}(n) + v_{i_R}(n) \tag{1}$$

where $*$ denotes the linear convolution and $v_{i_R}(n)$ is a spatio–temporally uncorrelated noise with variance δ_v^2.

We now briefly review the MIMO linear prediction and the related semiblind channel estimation method [12]. Define

$$\begin{aligned}\mathbf{y}_P(n-1) &\triangleq \left[\mathbf{y}^T(n-1), \ldots, \mathbf{y}^T(n-P)\right]^T \\ \tilde{\mathbf{R}}_{n-1} &\triangleq \mathrm{E}\left\{\mathbf{y}_P(n-1)\mathbf{y}_P^H(n-1)\right\} \\ \ddot{\mathbf{R}}_n &\triangleq \mathrm{E}\left\{\mathbf{y}(n)\mathbf{y}_P^H(n-1)\right\}.\end{aligned} \tag{2}$$

The MIMO linear predictor can be written as

$$\mathbf{P}_P \triangleq [\mathbf{P}_P(1), \mathbf{P}_P(2), \cdots, \mathbf{P}_P(P)] = \ddot{\mathbf{R}}_n \tilde{\mathbf{R}}_{n-1}^{-1} \tag{3}$$

where $\mathbf{P}_P(n), (n = 1, \ldots, P)$ is an $N_R \times N_R$ matrix representing the nth tap of the prediction filter. The covariance matrix of the prediction error can then be given by

$$\boldsymbol{\delta}_{\tilde{\mathbf{y}},P}^2 = \mathbf{R}(0) - \mathbf{P}_P \ddot{\mathbf{R}}_n^H \tag{4}$$

where

$$\mathbf{R}(0) \triangleq \mathrm{E}[\mathbf{y}(n)\mathbf{y}^H(n)].$$

Further, define

$$\mathbf{P}_P(z) = \mathbf{I} - \sum_{i=1}^{P} \mathbf{P}_P(i) z^{-i}$$
$$\mathbf{H}(z) = \sum_{i=0}^{L-1} \mathbf{H}(i) z^{-i}.$$

It has been shown in [12] and [13] that if the transmitted signals are uncorrelated, and moreover $PN_R \geq (L+P-1)N_T$, one can obtain

$$\mathbf{P}_P(z)\mathbf{H}(z) = \mathbf{H}(0) \tag{5}$$
$$\boldsymbol{\delta}^2_{\tilde{\mathbf{y}},P} = \mathbf{H}(0)\mathbf{H}^H(0). \tag{6}$$

Based on (5) and (6), some blind algorithms have been proposed for MIMO channel estimation [12], [24], [25]. The basic idea is to first acquire an estimate of $\mathbf{H}(0)$ from that of $\boldsymbol{\delta}^2_{\tilde{\mathbf{y}},P}$ according to (6), and then use (5) to obtain an estimate of the channel matrix $\mathbf{H}(z)$.

Using the aforementioned linear prediction along with training data, a semiblind approach for MIMO channel estimation has been developed to achieve a better estimation performance [9]–[11]. This approach is briefly described as follows.

Denote the null column space of $\mathbf{H}(0)$ as an $N_R \times (N_R - N_T)$ matrix $\mathbf{U}_{0\text{null}}$. From (6), one can easily find that $\mathbf{U}_{0\text{null}}$ can be estimated from $\boldsymbol{\delta}^2_{\tilde{\mathbf{y}},P}$. Using $\mathbf{U}_{0\text{null}}$ into (5) gives

$$\mathbf{U}^H_{0\text{null}}\mathbf{H}(0) = \mathbf{U}^H_{0\text{null}}\mathbf{P}_P(z)\mathbf{H}(z) = \mathbf{0} \tag{7}$$

which imposes a blind constraint $\mathbf{U}^H_{0\text{null}}\mathbf{P}_P(z)$ on the channel matrix $\mathbf{H}(z)$. Equation (7) can then be used to derive a blind constraint $\bar{\mathbf{B}}$ for the channel vector defined by

$$\mathbf{h}_v \triangleq \left[\mathbf{h}_1^T(0), \ldots, \mathbf{h}_{N_T}^T(0), \ldots, \mathbf{h}_1^T(L-1), \ldots, \mathbf{h}_{N_T}^T(L-1)\right]^T$$

where $\mathbf{h}_{i_T}(l) = [h_{1,i_T}(l), \ldots, h_{N_R,i_T}(l)]^T$. Using this blind constraint in conjunction with a training-based LS criterion [26], a semiblind cost function is then formulated as

$$\min_{\hat{\mathbf{h}}_v}\{\|\mathbf{Y}_{\text{TS}} - \mathbf{A}_{\text{TS}}\hat{\mathbf{h}}\|^2 + \alpha\|\bar{\mathbf{B}}\hat{\mathbf{h}}\|^2\} \tag{8}$$

where $\mathbf{A}_{\text{TS}}$ is a pilot signal matrix, $\mathbf{Y}_{\text{TS}}$ is the corresponding received signal vector, and $\alpha > 0$ is a weighting factor. The minimization problem (8) can be easily solved under the assumption that α is a fixed constant, giving a semiblind solution for the channel estimate $\hat{\mathbf{h}}$. It should be stressed that the estimation accuracy of the aforementioned semiblind algorithm heavily depends on the choice of the weighting factor. This issue, however, has not been investigated. In what follows, we will extend the semiblind method for the estimation of MIMO–OFDM channels. We will also propose a very efficient scheme for the determination of the value of α according to the MSE of the LS estimation as well as that of the blind constraint.

III. Proposed MIMO–OFDM Semiblind Channel Estimation Algorithm

For the sake of simplicity, only one OFDM symbol with K subcarriers is considered. Let $\mathbf{x}(n)$ be the transmitted time-domain signals at the output of the IDFT module, and $\mathbf{y}(n)$ be the received time-domain signals before the discrete Fourier transform (DFT) operation. If the length of the cyclic is not less than the channel length L, after removing the cyclic prefix at the receiver, the i_Rth received signal can be written as

$$y_{i_R}(n) = \sum_{i_T=1}^{N_T} h_{i_R,i_T}(n) \circledast x_{i_T}(n) + v_{i_R}(n), \quad n = 0, 1, \ldots, K-1. \tag{9}$$

Note that, unlike the MIMO signal model in (1), the MIMO–OFDM signal given by (9) involves the circular convolution with a length K. We would like to develop a semiblind MIMO–OFDM channel estimation approach based on (9) and the semiblind criterion in (8). The idea of this approach has been proposed for the first time in [27] and will now be completed in this paper.

A. Training-Based LS Criterion

In MIMO–OFDM systems, the training signal is transmitted in the frequency domain and thus the LS criterion in (8) should be modified. Following the approach in [2], the received frequency-domain signal at the i_Rth receive antenna can be obtained as

$$\mathbf{Y}_{i_R,\text{pilot}} = \mathbf{A}\mathbf{h}_{i_R} + \boldsymbol{\xi}_{i_R,\text{pilot}} \tag{10}$$

where $\boldsymbol{\xi}_{i_R,\text{pilot}}$ is the noise vector with respect to the i_Rth receive antenna, $\mathbf{h}_{i_R}$ is the one-path channel response vector as given by

$$\mathbf{h}_{i_R} \triangleq [h_{n_R,1}(0), \ldots, h_{n_R,1}(L-1), \ldots, h_{n_R,n_T}(L-1)]^T \tag{11}$$

and $\mathbf{A}$ is a $gK_p \times LN_T$ matrix carrying on the pilot signal, with K_p and g being the number of the pilot carriers and the number of OFDM symbols used, respectively. By defining

$$\bar{\mathbf{Y}}_{\text{pilot}} \triangleq [\mathbf{Y}_{1,\text{pilot}}, \ldots, \mathbf{Y}_{N_R,\text{pilot}}]$$
$$\mathbf{H} \triangleq [\mathbf{h}_1, \ldots, \mathbf{h}_{N_R}]$$
$$\bar{\boldsymbol{\xi}}_{\text{pilot}} \triangleq [\boldsymbol{\xi}_{1,\text{pilot}}, \ldots, \boldsymbol{\xi}_{N_R,\text{pilot}}] \tag{12}$$

we have

$$\bar{\mathbf{Y}}_{\text{pilot}} = \mathbf{A}\mathbf{H} + \bar{\boldsymbol{\xi}}_{\text{pilot}}. \tag{13}$$

Further, letting $\mathbf{Y}_{\text{pilot}} \triangleq \text{vec}(\bar{\mathbf{Y}}_{\text{pilot}})$, $\tilde{\mathbf{A}} \triangleq \mathbf{I} \otimes \mathbf{A}$, and $\mathbf{h} \triangleq \text{vec}(\mathbf{H})$, from (13), one can obtain a new LS criterion

$$\left\|\mathbf{Y}_{\text{pilot}} - \tilde{\mathbf{A}}\hat{\mathbf{h}}\right\|^2 \tag{14}$$

which will be used for the training signal in the proposed semiblind method.

B. MIMO Linear-Prediction-Based Blind Criterion

First, to validate the use of linear prediction for the received MIMO–OFDM signal, we show in Appendix A that the transmitted time-domain MIMO–OFDM signal is uncorrelated. We now use the MIMO linear prediction to obtain a blind constraint for the channel matrix. By following the linear prediction process in Section II, we can obtain a time-domain representation of (5)

$$[\mathbf{I}, -\mathbf{P}_P]\mathbf{H}_D = [\mathbf{H}(0), \mathbf{0}, \ldots, \mathbf{0}] \tag{15}$$

where $\mathbf{P}_P$ is given by (3) and $\mathbf{H}_D$ is a $(P+1)N_R \times (L+P)N_T$ block Toeplitz matrix with the first block row as

$$[\mathbf{H}(0), \ldots, \mathbf{H}(L-1), \mathbf{0}, \ldots, \mathbf{0}].$$

Letting

$$\mathbf{H}_F \triangleq \begin{bmatrix} \mathbf{H}(0) \\ \vdots \\ \mathbf{H}(L-1) \end{bmatrix}$$

$$\mathbf{P}_Q \triangleq \begin{bmatrix} \mathbf{I} & & & \mathbf{0} \\ -\mathbf{P}_P(1) & \ddots & & \\ \vdots & \ddots & \ddots & \\ -\mathbf{P}_P(P) & \vdots & \ddots & \mathbf{I} \\ & \ddots & \vdots & -\mathbf{P}_P(1) \\ & & \ddots & \vdots \\ \mathbf{0} & & & -\mathbf{P}_P(P) \end{bmatrix}$$

(15) can be rewritten as

$$\mathbf{P}_Q\mathbf{H}_F = \begin{bmatrix} \mathbf{H}(0) \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}. \tag{16}$$

Using (6), the null column space of $\mathbf{H}(0)$, $\mathbf{U}_{0\text{null}}$, can easily be obtained, which is then used to form

$$\mathbf{P}_\Sigma \triangleq \left(\mathbf{I}_{L+P} \otimes \mathbf{U}_{0\text{null}}^H\right)\mathbf{P}_Q. \tag{17}$$

From (16) and (17), we have

$$\mathbf{P}_\Sigma\mathbf{H}_F = \mathbf{0} \tag{18}$$

which is equivalent to

$$(\mathbf{I} \otimes \mathbf{P}_\Sigma)\text{vec}(\mathbf{H}_F) = \mathbf{0}. \tag{19}$$

Noting that $\text{vec}(\mathbf{H}_F) = \mathbf{E}_P\mathbf{h}$, where $\mathbf{E}_P$ is a known permutation matrix, (19) can be rewritten as

$$(\mathbf{I} \otimes \mathbf{P}_\Sigma)\mathbf{E}_P\mathbf{h} = \mathbf{B}\mathbf{h} = \mathbf{0} \tag{20}$$

implying that $\mathbf{B} = (\mathbf{I} \otimes \mathbf{P}_\Sigma)\mathbf{E}_P$ is a blind constraint for the channel vector $\mathbf{h}$.

In the computation of the linear predictor $\mathbf{P}_P$ and the covariance matrix $\boldsymbol{\delta}^2_{\tilde{\mathbf{y}},P}$, one has to estimate various correlation matrices $\tilde{\mathbf{R}}_{n-1}$, $\ddot{\mathbf{R}}_n$, and $\mathbf{R}(0)$, as discussed in Section II. Considering that the circular convolution is used in MIMO–OFDM systems, a more accurate estimate of these correlation matrices can be obtained in comparison with that in MIMO systems. For example, the estimate of $\tilde{\mathbf{R}}_{n-1}$ in MIMO systems is computed as

$$\hat{\tilde{\mathbf{R}}}_{n-1} = \frac{1}{K}\sum_{n=P-1}^{K-1} \mathbf{y}_P(n)\mathbf{y}_P^H(n) - \delta_v^2\mathbf{I} \tag{21}$$

where only $K-P+1$ received signal vectors $\mathbf{y}_P(n)$ $(n = P-1, \ldots, K-1)$ are available for estimation. In MIMO–OFDM systems, however, the estimation of $\tilde{\mathbf{R}}_{n-1}$ can be modified as

$$\hat{\tilde{\mathbf{R}}}_{n-1} = \frac{1}{K}\sum_{n=0}^{K-1} \mathbf{y}_P(n)\mathbf{y}_P^H(n) - \delta_v^2\mathbf{I} \tag{22}$$

where $\mathbf{y}_P(n)(n = 0, \ldots, P-2)$ can be obtained using $\mathbf{y}(n-i) \triangleq \mathbf{y}(n-i+K)$ when $n-i < 0$ due to the circular convolution. Because more signal samples are used in the estimate, a better linear prediction result can be expected in MIMO–OFDM systems. Note that, when multiple OFDM symbols are used, $\hat{\tilde{\mathbf{R}}}_{n-1}$ can be easily calculated by averaging the results obtained from each OFDM symbol using (22).

C. Semiblind Solution

Combining (14) and (20), a semiblind cost function for the estimation of the channel vector $\mathbf{h}$ can be formulated as

$$\min_{\mathbf{h}} \Delta = \|\mathbf{Y}_{\text{pilot}} - \tilde{\mathbf{A}}\mathbf{h}\|^2 + \alpha\|\hat{\mathbf{B}}\mathbf{h}\|^2 \tag{23}$$

where $\hat{\mathbf{B}}$ is an estimate of the blind constraint. The solution to this minimization problem can be obtained by letting

$$\frac{\partial\Delta}{\partial\mathbf{h}^H} = -\tilde{\mathbf{A}}^H(\mathbf{Y}_{\text{pilot}} - \tilde{\mathbf{A}}\mathbf{h}) + \alpha\hat{\mathbf{B}}^H\hat{\mathbf{B}}\mathbf{h} = \mathbf{0}$$

which gives

$$\hat{\mathbf{h}} = (\tilde{\mathbf{A}}^H\tilde{\mathbf{A}} + \alpha\hat{\mathbf{B}}^H\hat{\mathbf{B}})^\dagger\tilde{\mathbf{A}}^H\mathbf{Y}_{\text{pilot}}. \tag{24}$$

Clearly, the performance of the semiblind algorithm depends on the choice of α. However, the selection of α has not yet been discussed in the existing semiblind MIMO channel estimation methods such as those in [9], [10], [26], and [28]. In this paper, we propose an explicit formula for the calculation of α in terms of the estimated MSE of the training-based criterion and that of the blind part.

It has been shown in [29] and [30] that the weight for a weighted least squares (WLS) minimization problem can be determined according to the variance of the individual estimation error involved provided that the error is Gaussian distributed. A vector version of the WLS problem can be described as

$$S = \sum_{i=1}^{N} w_i\|\mathbf{e}_i\|^2 = \sum_{i=1}^{N} w_i\|\hat{\mathbf{x}}_i - \mathbf{x}_i\|^2 \tag{25}$$

where $\hat{\mathbf{x}}_i$ denotes the estimate of the true vector $\mathbf{x}_i$ and w_i is the weight for the ith error term, which should be chosen as the reciprocal of the variance δ_i^2 of $\mathbf{e}_i$, i.e.,

$$w_i = \frac{1}{\delta_i^2} = \frac{1}{E\{\|\mathbf{e}_i\|^2\}}.$$

In order to develop a closed-form expression for the weighting factor α in (23), let us consider the following WLS problem:

$$\min \Delta = w_{\mathrm{T}}\|\mathbf{e}_{\mathrm{T}}\|^2 + w_{\mathrm{B}}\|\mathbf{e}_{\mathrm{B}}\|^2 \tag{26}$$

where

$$\mathbf{e}_{\mathrm{T}} \triangleq \hat{\mathbf{h}}_{\mathrm{T}} - \mathbf{h}$$
$$\mathbf{e}_{\mathrm{B}} \triangleq \hat{\mathbf{h}}_{\mathrm{B}} - \mathbf{h}$$

with $\hat{\mathbf{h}}_{\mathrm{T}}$ being the estimate of the channel vector resulting from the training-based LS criterion and $\hat{\mathbf{h}}_{\mathrm{B}}$ that from the blind estimation part. Since the error vectors $\mathbf{e}_{\mathrm{T}}$ and $\mathbf{e}_{\mathrm{B}}$ can be considered Gaussian, the coefficients w_{T} and w_{B} can be determined from the variance of $\mathbf{e}_{\mathrm{T}}$ and that of $\mathbf{e}_{\mathrm{B}}$, respectively, thereby the value of α can be estimated.

We first consider the variance of $\mathbf{e}_{\mathrm{T}}$. From [2], one can easily find that the optimal pilots are given by

$$\tilde{\mathbf{A}}^H\tilde{\mathbf{A}} = \left(gK_p\delta_x^2\right)\mathbf{I}_{LN_RN_T}.$$

Thus, for the optimal pilots, one can obtain

$$\|\tilde{\mathbf{A}}\|_F^2 = gK_pLN_RN_T\delta_x^2 \tag{27}$$

$$\begin{aligned}\mathrm{MSE_T} &\triangleq \mathrm{E}\{\|\mathbf{e}_{\mathrm{T}}\|^2\} = \mathrm{E}\{\|\hat{\mathbf{h}}_{\mathrm{T}} - \mathbf{h}\|^2\}\\ &= \frac{N_RN_TL\delta_v^2}{gK_p\delta_x^2}.\end{aligned} \tag{28}$$

Note that it is not possible to directly compute the variance of $\mathbf{e}_{\mathrm{B}}$, since the blind solution $\hat{\mathbf{h}}_{\mathrm{B}}$ is not available in the proposed method. However, we may use the following variance to replace the variance of $\mathbf{e}_{\mathrm{B}}$:

$$\mathrm{MSE_B} \triangleq \mathrm{E}\{\|\hat{\mathbf{B}}(\hat{\mathbf{h}}_{\mathrm{B}} - \mathbf{h})\|^2\} = \mathrm{E}\{\|\hat{\mathbf{B}}\mathbf{h}\|^2\}. \tag{29}$$

It will be shown through simulation study in Section V that this approximation gives an appropriate choice of α in terms of the channel estimation performance.

Using (28) and (29), (26) can be rewritten as [29], [30]

$$\min_{\mathbf{h}} \Delta = \frac{1}{\mathrm{MSE_T}}\|\hat{\mathbf{h}}_{\mathrm{T}} - \mathbf{h}\|^2 + \frac{1}{\mathrm{MSE_B}}\|\hat{\mathbf{h}}_{\mathrm{B}} - \mathbf{h}\|^2. \tag{30}$$

The previous minimization problem can be nearly reformulated as the following LS problem:

$$\min_{\mathbf{h}} \Delta = \frac{\|\mathbf{Y}_{\mathrm{pilot}} - \tilde{\mathbf{A}}\mathbf{h}\|^2}{\|\tilde{\mathbf{A}}\|_F^2\mathrm{MSE_T}} + \frac{\|\hat{\mathbf{B}}\mathbf{h}\|^2}{\|\hat{\mathbf{B}}\|_F^2\mathrm{MSE_B}}. \tag{31}$$

Evidently, the solution of (31) is equivalent to that of (23) when α is chosen as

$$\alpha = \frac{\mathrm{MSE_T}}{\mathrm{MSE_B}}\frac{\|\tilde{\mathbf{A}}\|_F^2}{\|\hat{\mathbf{B}}\|_F^2}. \tag{32}$$

What remains in the computation of α is to determine the $\mathrm{MSE_B}$. In the next section, a perturbation analysis of the linear-prediction-based blind algorithm is performed, leading to a novel closed-form expression for $\mathrm{MSE_B}$.

IV. Perturbation Analysis of the Linear-Prediction-Based Blind Channel Estimation

It is known that the solution of linear prediction or subspace-based methods is always perturbed by various sources, such as finite data length, measurement noise, etc. [18], [21]–[23]. Perturbation theory has been successfully applied to the analysis of subspace-based methods [19]–[21], [26]. In this section, the first-order perturbation theory is employed to analyze the MIMO linear prediction as well as the resulting blind constraint.

A. Linear Predictor With Perturbation for MIMO Channel Estimation

Letting $\mathbf{H}_A \triangleq [\mathbf{H}(0), \mathbf{H}(1), \ldots, \mathbf{H}(L-1)]$ and $\mathbf{x}_L(n) \triangleq [\mathbf{x}^T(n)\cdots\mathbf{x}^T(n-L+1)]^T, (n = 0, 1, \ldots, K-1)$, where $\mathbf{x}(n) = \mathbf{x}(K+n)$ for $n < 0$, the circular convolution (9) can be rewritten in the matrix form as

$$\mathbf{y}(n) = \mathbf{H}_A\mathbf{x}_L(n) + \mathbf{v}(n). \tag{33}$$

Similarly, $\mathbf{y}_P(n-1)$ can be expressed as

$$\mathbf{y}_P(n-1) = \mathbf{H}_B\mathbf{x}_{P+L-1}(n-1) + \mathbf{v}_P(n-1) \tag{34}$$

where $\mathbf{v}_P(n-1) \triangleq [\mathbf{v}^T(n-1)\cdots\mathbf{v}^T(n-P)]^T$ and $\mathbf{H}_B$ is a $PN_R \times (P+L-1)N_T$ block Toeplitz matrix with the first block row given by $[\mathbf{H}(0), \ldots, \mathbf{H}(L-1), \mathbf{0}, \ldots, \mathbf{0}]$. In this paper, we consider only the perturbation due to the finite data length in the computation of the correlation matrices. Without loss of generality, we let the variance of the signal δ_x^2 equal 1. Using (34), the correlation matrix $\hat{\hat{\mathbf{R}}}_{n-1}$ with such a perturbation can be written as

$$\begin{aligned}\hat{\hat{\mathbf{R}}}_{n-1} &= \hat{\mathrm{E}}\left[\mathbf{y}_P(n-1)\mathbf{y}_P^H(n-1)\right] - \delta_v^2\mathbf{I}\\ &= \mathbf{H}_B[\mathbf{I} + \Delta\mathbf{R}_{x1}]\mathbf{H}_B^H + \Delta\mathbf{R}_{v1}\end{aligned} \tag{35}$$

where $\Delta\mathbf{R}_{x1}$ denotes the signal perturbation matrix

$$\Delta\mathbf{R}_{x1} = \frac{1}{K}\sum_{n=0}^{K-1}\mathbf{x}_{P+L-1}(n-1)\mathbf{x}_{P+L-1}^H(n-1) - \mathbf{I}$$

and $\Delta\mathbf{R}_{v1}$ is the perturbation matrix introduced by the noise

$$\Delta\mathbf{R}_{v1} \triangleq \mathbf{H}_B\Delta\mathbf{R}_{xv1} + \Delta\mathbf{R}_{xv1}^H\mathbf{H}_B^H + \Delta\mathbf{R}_{vv1} \tag{36}$$

with

$$\Delta\mathbf{R}_{xv1} \triangleq \frac{1}{K}\sum_{n=0}^{K-1}\mathbf{x}_{P+L-1}(n-1)\mathbf{v}_P^H(n-1)$$

$$\Delta\mathbf{R}_{vv1} \triangleq \frac{1}{K}\sum_{n=0}^{K-1}\mathbf{v}_P(n-1)\mathbf{v}_P^H(n-1) - \delta_v^2\mathbf{I}. \tag{37}$$

Similarly, $\hat{\hat{\mathbf{R}}}_n$ is given by

$$\begin{aligned}\hat{\hat{\mathbf{R}}}_n &= \hat{\mathrm{E}}\left[\mathbf{y}(n)\mathbf{y}_P^H(n-1)\right] - \delta_v^2[\mathbf{I}, \mathbf{0}, \ldots, \mathbf{0}]\\ &= \mathbf{H}_A[\mathbf{I}_C + \Delta\mathbf{R}_{x2}]\mathbf{H}_B^H + \Delta\mathbf{R}_{v2}\end{aligned} \tag{38}$$

where

$$\mathbf{I}_C \triangleq \begin{bmatrix} \mathbf{0}_{N_T\times(L-1)N_T} & \mathbf{0}_{N_T\times PN_T} \\ \mathbf{I}_{(L-1)N_T\times(L-1)N_T} & \mathbf{0}_{(L-1)N_T\times PN_T} \end{bmatrix}$$

$$\Delta\mathbf{R}_{x2} \triangleq \frac{1}{K}\sum_{n=0}^{K-1}\mathbf{x}_L(n)\mathbf{x}_{P+L-1}^H(n-1) - \mathbf{I}_C$$

$$\Delta\mathbf{R}_{v2} \triangleq \mathbf{H}_A\Delta\mathbf{R}_{xv21} + \Delta\mathbf{R}_{xv22}^H\mathbf{H}_B^H + \Delta\mathbf{R}_{vv2}$$

with

$$\Delta\mathbf{R}_{xv21} \triangleq \frac{1}{K}\sum_{n=0}^{K-1}\mathbf{x}_L(n)\mathbf{v}_P^H(n-1)$$

$$\Delta\mathbf{R}_{xv22} \triangleq \frac{1}{K}\sum_{n=0}^{K-1}\mathbf{x}_{P+L-1}(n-1)\mathbf{v}^H(n)$$

$$\Delta\mathbf{R}_{vv2} \triangleq \frac{1}{K}\sum_{n=0}^{K-1}\mathbf{v}(n)\mathbf{v}_P^H(n-1) - \delta_v^2[\mathbf{I},\mathbf{0},\ldots,\mathbf{0}].$$

Based on the previous expressions for the correlation matrices with perturbation, we now derive the perturbation form of the channel constraint given in (15) and that of the covariance matrix given in (4). To this end, we first derive the perturbation form of the linear predictor $\hat{\mathbf{P}}_P$. Using the first-order approximation, the pseudoinverse of $\hat{\tilde{\mathbf{R}}}_{n-1}$ given by (35) can be obtained as

$$\hat{\tilde{\mathbf{R}}}_{n-1}^{\dagger} \approx \mathbf{\Pi}_1 - \mathbf{\Pi}_2\Delta\mathbf{R}_{x1}\mathbf{\Pi}_2^H - \mathbf{\Pi}_1\Delta\mathbf{R}_{v1}\mathbf{\Pi}_1^H \tag{39}$$

where $\mathbf{\Pi}_1 \triangleq (\mathbf{H}_B\mathbf{H}_B^H)^{\dagger}$ and $\mathbf{\Pi}_2 \triangleq \mathbf{\Pi}_1\mathbf{H}_B$. It should be noted that the noise subspace of $\hat{\tilde{\mathbf{R}}}_{n-1}$ in (35) has been omitted in obtaining (39), and such an omission would not affect the derivation of the perturbation form of (15) and (4). Using (38) and (39), the linear predictor can be derived as

$$\hat{\mathbf{P}}_P \approx \mathbf{H}_A[\mathbf{I}_C + \Delta\mathbf{R}_{x2}]\mathbf{\Pi}_2^H - \mathbf{H}_A\mathbf{I}_C\mathbf{\Pi}_2^H\mathbf{H}_B\Delta\mathbf{R}_{x1}\mathbf{\Pi}_2^H - \mathbf{H}_A\mathbf{I}_C\mathbf{\Pi}_2^H\Delta\mathbf{R}_{v1}\mathbf{\Pi}_1 + \Delta\mathbf{R}_{v2}\mathbf{\Pi}_1. \tag{40}$$

Noting that $\mathbf{H}_D$ can be partitioned as

$$\mathbf{H}_D \triangleq \begin{bmatrix} \mathbf{H}(0) & \mathbf{H}_E \\ \mathbf{0} & \mathbf{H}_B \end{bmatrix}$$

where $\mathbf{H}_E \triangleq [\mathbf{H}(1),\ldots,\mathbf{H}(L-1),\mathbf{0},\ldots,\mathbf{0}]$, the perturbation form of the left-hand side (LHS) of (15) can be written as

$$[\mathbf{I}, -\hat{\mathbf{P}}_P]\mathbf{H}_D = [\mathbf{H}_0, \mathbf{H}_E] - [\mathbf{0}, \hat{\mathbf{P}}_P\mathbf{H}_B]. \tag{41}$$

Using (40) and $\mathbf{H}_E = \mathbf{H}_A\mathbf{I}_C$, and noting that $\mathbf{H}_B^H(\mathbf{H}_B\mathbf{H}_B^H)^{\dagger}\mathbf{H}_B = \mathbf{I}$ when $\mathbf{H}_B$ has a full column rank and $PN_R \geq (P+L-1)N_T$, one can deduce

$$\hat{\mathbf{P}}_P\mathbf{H}_B = \mathbf{H}_E + \mathbf{\Pi}_P$$

where

$$\mathbf{\Pi}_P \triangleq \mathbf{H}_A\Delta\mathbf{R}_{x2} - \mathbf{H}_A\mathbf{I}_C\Delta\mathbf{R}_{x1} + \Delta\mathbf{R}_{v2}\mathbf{\Pi}_2 - \mathbf{H}_A\mathbf{I}_C\mathbf{\Pi}_2^H\Delta\mathbf{R}_{v1}\mathbf{\Pi}_2. \tag{42}$$

Thus, (41) can be rewritten as

$$[\mathbf{I}, -\hat{\mathbf{P}}_P]\mathbf{H}_D = [\mathbf{H}(0), \mathbf{0},\ldots,\mathbf{0}] - [\mathbf{0}, \mathbf{\Pi}_P]. \tag{43}$$

Clearly, (43) is the perturbation form of (15). By noting that

$$\Delta\mathbf{R}_{x2} - \mathbf{I}_C\Delta\mathbf{R}_{x1} = \frac{1}{K}\sum_{n=0}^{K-1}\begin{bmatrix}\mathbf{x}(n)\\ \mathbf{0}_{(L-1)N_T\times 1}\end{bmatrix}\mathbf{x}_{P+L-1}^H(n-1)$$

(42) can be reduced to

$$\mathbf{\Pi}_P = \mathbf{H}(0)\Delta\mathbf{R}_{x4} + \mathbf{\Pi}'_P \tag{44}$$

where

$$\Delta\mathbf{R}_{x4} \triangleq \frac{1}{K}\sum_{n=0}^{K-1}\mathbf{x}(n)\mathbf{x}_{P+L-1}^H(n-1) \tag{45}$$

$$\mathbf{\Pi}'_P \triangleq \Delta\mathbf{R}_{v2}\mathbf{\Pi}_2 - \mathbf{H}_A\mathbf{I}_C\mathbf{\Pi}_2^H\Delta\mathbf{R}_{v1}\mathbf{\Pi}_2. \tag{46}$$

It is obvious from (44) that the perturbation term $\mathbf{\Pi}_P$ consists of the signal perturbation $\mathbf{H}(0)\Delta\mathbf{R}_{x4}$ and the perturbation introduced by the noise $\mathbf{\Pi}'_P$.

We now turn to the derivation of the perturbed version of (6). The perturbed form of (4) can be written as

$$\hat{\boldsymbol{\delta}}_{\tilde{\mathbf{y}},P}^2 = \hat{\mathbf{R}}(0) - \hat{\mathbf{P}}_P\hat{\tilde{\mathbf{R}}}_n^H \tag{47}$$

where

$$\hat{\mathbf{R}}(0) = \hat{\mathrm{E}}[\mathbf{y}(n)\mathbf{y}^H(n)] - \delta_v^2\mathbf{I}. \tag{48}$$

It is easy to show that

$$\hat{\mathbf{R}}(0) = \mathbf{H}_A[\mathbf{I} + \Delta\mathbf{R}_{x3}]\mathbf{H}_A^H + \Delta\mathbf{R}_{v3} \tag{49}$$

where

$$\Delta\mathbf{R}_{x3} \triangleq \frac{1}{K}\sum_{n=0}^{K-1}\mathbf{x}_L(n)\mathbf{x}_L^H(n) - \mathbf{I}$$

$$\Delta\mathbf{R}_{v3} \triangleq \mathbf{H}_A\Delta\mathbf{R}_{xv3} + \Delta\mathbf{R}_{xv3}^H\mathbf{H}_A^H + \Delta\mathbf{R}_{vv3}$$

with

$$\Delta\mathbf{R}_{xv3} \triangleq \frac{1}{K}\sum_{n=0}^{K-1}\mathbf{x}_L(n)\mathbf{v}^H(n)$$

$$\Delta\mathbf{R}_{vv3} \triangleq \frac{1}{K}\sum_{n=0}^{K-1}\mathbf{v}(n)\mathbf{v}^H(n) - \delta_v^2\mathbf{I}.$$

Using (40), the second term of the right-hand side (RHS) of (47) can be expressed as

$$\begin{aligned}\hat{\mathbf{P}}_P\hat{\tilde{\mathbf{R}}}_n^H = {} & \mathbf{H}_A\mathbf{I}_C\mathbf{I}_C^H\mathbf{H}_A^H + \mathbf{H}_A\Delta\mathbf{R}_{x2}\mathbf{I}_C^H\mathbf{H}_A^H \\ & + \mathbf{H}_A\mathbf{I}_C\Delta\mathbf{R}_{x2}^H\mathbf{H}_A^H + \mathbf{H}_A\mathbf{I}_C\mathbf{\Pi}_2^H\Delta\mathbf{R}_{v2}^H \\ & + \Delta\mathbf{R}_{v2}\mathbf{\Pi}_2\mathbf{I}_C^H\mathbf{H}_A^H - \mathbf{H}_A\mathbf{I}_C\Delta\mathbf{R}_{x1}\mathbf{I}_C^H\mathbf{H}_A^H \\ & - \mathbf{H}_A\mathbf{I}_C\mathbf{\Pi}_2^H\Delta\mathbf{R}_{v1}\mathbf{\Pi}_2\mathbf{I}_C^H\mathbf{H}_A^H.\end{aligned} \tag{50}$$

From

$$\mathbf{I}_C\mathbf{I}_C^H = \begin{bmatrix} \mathbf{0}_{N_T\times N_T} & \mathbf{0}_{N_T\times(L-1)N_T} \\ \mathbf{0}_{(L-1)N_T\times N_T} & \mathbf{I}_{(L-1)N_T\times(L-1)N_T} \end{bmatrix}$$

one can get

$$\mathbf{H}_A\left[\mathbf{I}-\mathbf{I}_C\mathbf{I}_C^H\right]\mathbf{H}_A^H=\mathbf{H}(0)\mathbf{H}^H(0). \tag{51}$$

Using (49) and (50) along with (51), the covariance matrix given by (6) has the following perturbation form:

$$\hat{\boldsymbol{\delta}}_{\tilde{\mathbf{y}},P}^2=\mathbf{H}(0)\mathbf{H}^H(0)+\Delta\hat{\boldsymbol{\delta}}_{\tilde{\mathbf{y}},P}^2 \tag{52}$$

where

$$\begin{aligned}\Delta\hat{\boldsymbol{\delta}}_{\tilde{\mathbf{y}},P}^2 \triangleq\; & \mathbf{H}_A\Delta\mathbf{R}_{x3}\mathbf{H}_A^H-\mathbf{H}_A\Delta\mathbf{R}_{x2}\mathbf{I}_C^H\mathbf{H}_A^H\\ & -\mathbf{H}_A\mathbf{I}_C\Delta\mathbf{R}_{x2}^H\mathbf{H}_A^H+\mathbf{H}_A\mathbf{I}_C\Delta\mathbf{R}_{x1}\mathbf{I}_C^H\mathbf{H}_A^H\\ & +\Delta\mathbf{R}_{v3}-\mathbf{H}_A\mathbf{I}_C\boldsymbol{\Pi}_2^H\Delta\mathbf{R}_{v2}^H-\Delta\mathbf{R}_{v2}\boldsymbol{\Pi}_2\mathbf{I}_C^H\mathbf{H}_A^H\\ & +\mathbf{H}_A\mathbf{I}_C\boldsymbol{\Pi}_2^H\Delta\mathbf{R}_{v1}\boldsymbol{\Pi}_2\mathbf{I}_C^H\mathbf{H}_A^H.\end{aligned} \tag{53}$$

We now show that the perturbation term $\Delta\hat{\boldsymbol{\delta}}_{\tilde{\mathbf{y}},P}^2$ can be split into the signal perturbation and the perturbation due to the noise. By using

$$\begin{aligned}\Delta\mathbf{R}_{x3}-\Delta\mathbf{R}_{x2}\mathbf{I}_C^H=\;&\frac{1}{K}\sum_{n=0}^{K-1}\mathbf{x}_L(n)\left[\mathbf{x}^H(n),\mathbf{0}_{1\times(L-1)N_T}\right]\\ &-\left(\mathbf{I}-\mathbf{I}_C\mathbf{I}_C^H\right)\end{aligned}$$

$$\begin{aligned}&\mathbf{I}_C\Delta\mathbf{R}_{x1}\mathbf{I}_C^H-\mathbf{I}_C\Delta\mathbf{R}_{x2}\\ &\quad=-\frac{1}{K}\sum_{n=0}^{K-1}\begin{bmatrix}\mathbf{0}(n)\\ \mathbf{x}_{L-1}(n-1)\end{bmatrix}\left[\mathbf{x}^H(n),\mathbf{0}_{1\times(L-1)N_T}\right]\end{aligned}$$

the first four terms of the RHS of (53) equal $\mathbf{H}(0)\Delta\mathbf{R}_{x5}\mathbf{H}^H(0)$ with

$$\Delta\mathbf{R}_{x5}\triangleq\frac{1}{K}\sum_{n=0}^{K-1}\mathbf{x}(n)\mathbf{x}^H(n)-\mathbf{I}. \tag{54}$$

Therefore, (53) reduces

$$\Delta\hat{\boldsymbol{\delta}}_{\tilde{\mathbf{y}},P}^2=\mathbf{H}(0)\Delta\mathbf{R}_{x5}\mathbf{H}^H(0)+\boldsymbol{\Xi} \tag{55}$$

where

$$\begin{aligned}\boldsymbol{\Xi}\triangleq\;&\Delta\mathbf{R}_{v3}-\mathbf{H}_A\mathbf{I}_C\boldsymbol{\Pi}_2^H\Delta\mathbf{R}_{v2}^H-\Delta\mathbf{R}_{v2}\boldsymbol{\Pi}_2\mathbf{I}_C^H\mathbf{H}_A^H\\ &+\mathbf{H}_A\mathbf{I}_C\boldsymbol{\Pi}_2^H\Delta\mathbf{R}_{v1}\boldsymbol{\Pi}_2\mathbf{I}_C^H\mathbf{H}_A^H.\end{aligned} \tag{56}$$

It should be mentioned that in the noise-free case, all the perturbation terms in the previous equations introduced by noise can be ignored. As such, (43) and (52) can be simplified as

$$[\mathbf{I},-\hat{\mathbf{P}}_P]\mathbf{H}_D=[\mathbf{H}(0),\mathbf{0},\ldots,\mathbf{0}]-[\mathbf{0},\mathbf{H}(0)\Delta\mathbf{R}_{x4}] \tag{57}$$

and

$$\hat{\boldsymbol{\delta}}_{\tilde{\mathbf{y}},P}^2=\mathbf{H}(0)\mathbf{H}^H(0)+\mathbf{H}(0)\Delta\mathbf{R}_{x5}\mathbf{H}^H(0) \tag{58}$$

respectively. In conventional blind algorithms [12], [24], [25], an estimate of $\mathbf{H}(0)$ is first acquired from $\hat{\boldsymbol{\delta}}_{\tilde{\mathbf{y}},P}^2$ and is then used to estimate the channel matrix $\mathbf{H}_D$. This approach would give rise to a large estimation error due to the presence of perturbation terms as seen from (57) and (58). In the proposed semiblind approach, however, an ideal nulling constraint on the channel matrix $\mathbf{H}_D$ has been obtained from the ideal null space of $\mathbf{H}(0)$, which is not affected by the perturbation terms. Therefore, the semiblind method using MIMO linear prediction is superior to the blind algorithms [12], [24], [25]. It should also be pointed out that in the presence of noise, although both the semiblind and the blind methods are subject to the noise perturbation terms, the semiblind method still outperforms the blind one, since the perturbation introduced by the noise is, in general, significantly smaller than the signal perturbation.

B. MSE of Blind Channel Estimate

To derive the MSE expression, the perturbed version of (18) needs to be determined. By partitioning $\boldsymbol{\Pi}_P$ as $\boldsymbol{\Pi}_P=[\boldsymbol{\Pi}_{P,1},\ldots,\boldsymbol{\Pi}_{P,P+L-1}]$, with each submatrix $\boldsymbol{\Pi}_{P,i},(i=1,2,\ldots,P+L-1)$ being an $N_R\times N_T$ matrix, we can get the perturbation form of (16)

$$\hat{\mathbf{P}}_Q\mathbf{H}_F=\left[\mathbf{H}^H(0),\mathbf{0},\ldots,\mathbf{0}\right]^H-\mathbf{Q} \tag{59}$$

where

$$\mathbf{Q}\triangleq\left[\mathbf{0},\boldsymbol{\Pi}_{P,1}^H,\ldots,\boldsymbol{\Pi}_{P,P+L-1}^H\right]^H. \tag{60}$$

Then, using (52) and (55) along with the first-order approximation [20], one can get the estimate of $\mathbf{U}_{0\mathrm{null}}$

$$\hat{\mathbf{U}}_{0\mathrm{null}}\approx\mathbf{U}_{0\mathrm{null}}-\left(\mathbf{H}(0)\mathbf{H}^H(0)\right)^{\dagger}\boldsymbol{\Xi}\mathbf{U}_{0\mathrm{null}}. \tag{61}$$

Substituting (61) into (17) and using (59), we obtain

$$\hat{\mathbf{P}}_{\Sigma}\mathbf{H}_F=\mathbf{G}_1-\mathbf{G}_2 \tag{62}$$

where

$$\mathbf{G}_1\triangleq\left[\left(\mathbf{U}_{0\mathrm{null}}^H\boldsymbol{\Xi}^H\left(\mathbf{H}(0)\mathbf{H}^H(0)\right)^{\dagger}\mathbf{H}(0)\right)^H,\mathbf{0},\ldots,\mathbf{0}\right]^H \tag{63}$$

$$\mathbf{G}_2\triangleq\left(\mathbf{I}_{P+L}\otimes\mathbf{U}_{0\mathrm{null}}^H\right)\mathbf{Q}. \tag{64}$$

Clearly, (62) is the perturbed version of (18). From (62), the MSE of the blind criterion can be calculated as

$$\mathrm{MSE}_{\mathrm{B}}=\mathrm{Trace}\left\{\mathrm{E}\left[\mathrm{vec}\left(\mathbf{G}_1-\mathbf{G}_2\right)\mathrm{vec}^H(\mathbf{G}_1-\mathbf{G}_2)\right]\right\}. \tag{65}$$

In Appendix B, we investigate the computation of $\mathrm{vec}(\mathbf{G}_1)$ and $\mathrm{vec}(\mathbf{G}_2)$ to obtain a closed-form expression for $\mathrm{MSE}_{\mathrm{B}}$ in terms of the correlation of the transmitted signal, the correlation of the noise as well as the channel matrix.

V. Simulation Results

We consider a MIMO–OFDM system with two transmit and four receive antennas. The number of subcarriers is set to 512, the length of cyclic prefix is ten, and the length of the linear predictor is $P = 4$. In our simulation, the quaternary phase-shift keying (QPSK) modulation is used and a Rayleigh channel modeled by a three-tap MIMO–FIR filter is assumed, in which each tap corresponds to a 2×4 random matrix whose elements are independent identically distributed (i.i.d.) complex Gaussian variables with zero mean and unit variance. Because the blind constraint of the proposed semiblind algorithm is derived directly from $\mathbf{H}(0)$ as seen from (5) and (6), it is necessary to consider in the simulation study the effect of $\mathbf{H}(0)$ on the channel estimation performance. To this end, we define the following metric:

$$\eta \triangleq \frac{\|\mathbf{H}(0)\|_F^2}{\sum_{n=0}^{2} \|\mathbf{H}(n)\|_F^2}$$

and conduct our investigation with respect to different ranges of η.

It is seen from (32) that one has to choose the value of $\mathrm{MSE_B}$ to determine the parameter α. Although a closed-form expression for $\mathrm{MSE_B}$ has been obtained in Section IV, it requires the true channel matrix, which is, however, unknown in practice. Therefore, we first estimate the channel matrix using the LS method and then utilize the preliminary estimate to compute the $\mathrm{MSE_B}$ and thereby the value of α. To evaluate the proposed algorithm, we also calculate the value of $\mathrm{MSE_B}$ based on the true channel matrix, which gives a reference for the choice of α. The semiblind solution thus obtained is called the reference algorithm. The estimation performance is evaluated in terms of the MSE of the estimate of the channel matrix given by

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

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

Experiment 1: MSE Versus Weight Factor α: In the first experiment, the channel estimation performance in terms of the plot of the MSE versus α is investigated. The simulation is based on four randomly generated channel matrices, each having a value of η falling within the range of $[0.2, 0.3]$, $[0.3, 0.4]$, $[0.4, 0.6]$, or $[0.6, 0.8]$. For each channel matrix, three runs of the transmission of one OFDM symbol at 512 subcarriers, of which 32 are used as pilot for training purpose, are performed. Fig. 1 shows the MSE plots of the semiblind algorithm with respect to different choices of α in the range of $[0, 2]$ for the four channels at a signal-to-noise ratio (SNR) of 15 dB. The two points identified by "+" and "o" in each run indicate the results from the proposed and reference algorithms, respectively. It is observed that the proposed scheme for the determination of α gives a competitive MSE result. It is also seen that the use of the preliminary LS estimate of the channel matrix, instead of the true channel characteristic, would suffice for the calculation of the value of α.

Experiment 2: MSE Versus η: In this experiment, the channel estimation performance in terms of the MSE versus η is investigated. The simulation is undertaken based on 5000 Monte Carlo runs of the transmission of one OFDM symbol on 512 subcarriers at an SNR of 15 dB. Fig. 2 shows the MSE plots resulting

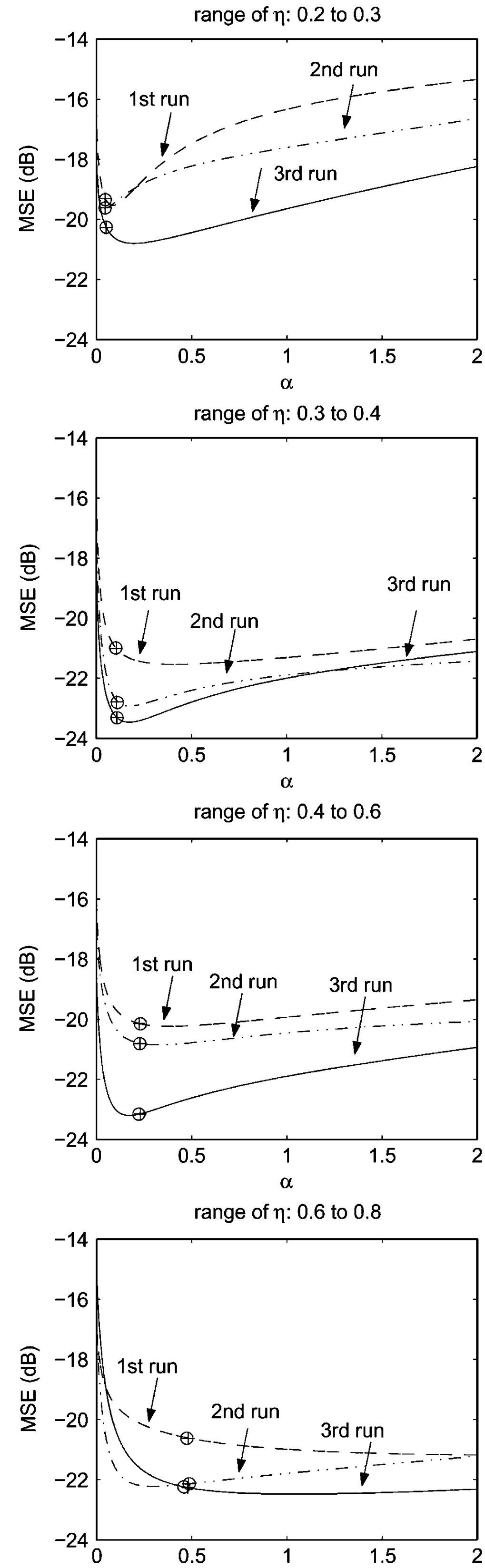

Fig. 1. MSE versus α for different ranges of η in which the two points indicated by "+" and "o" for each run are obtained from the proposed and reference algorithms, respectively.

from the proposed as well as the reference algorithm along with that from the LS estimation, indicating a high consistency of

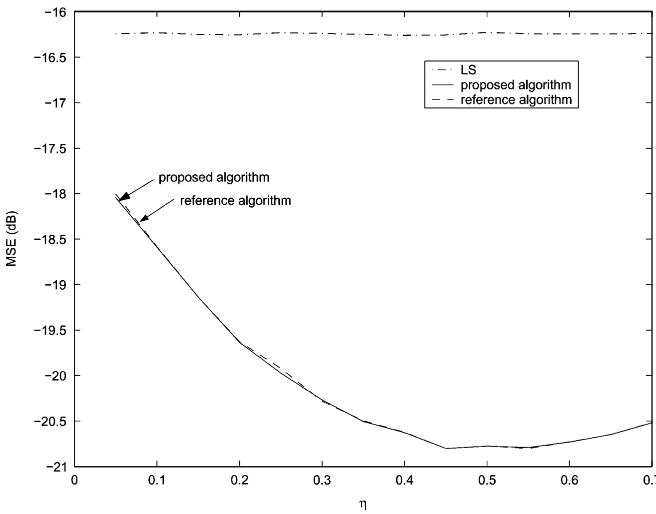

Fig. 2. MSE versus the value of η.

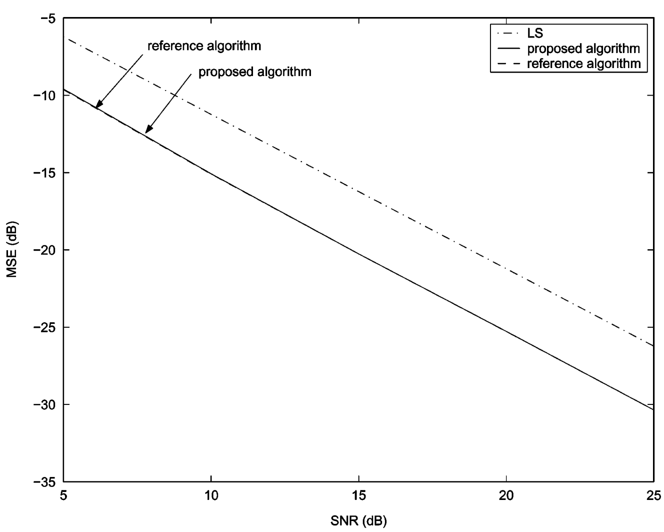

Fig. 3. MSE versus SNR.

the two semiblind methods. It is noted that the proposed semiblind algorithm significantly outperforms the LS method. It is also clear that the MSE performance of the semiblind estimation depends on the value of η. A better estimation performance is reached when η varies from 0.3 to 0.7, which represents a typical mobile communication scenario, where the first arrived path is comparable to or stronger than other paths [31]. From Fig. 2, one can find, when η is larger than 0.55, the MSE becomes a little larger with the increase of η. We believe the possible reason is that the full column rank assumption of $\mathbf{H}_B$ for the proposed semiblind algorithm cannot be always satisfied in the experiment simulations, since the Rayleigh channel is generated for the Monte Carlo iterations. With the increase of the value of η, especially when η is larger than 0.55, $\mathbf{H}(1)$ and $\mathbf{H}(2)$ become more and more insignificant. As a result, the probability for the case $\mathbf{H}_B$ does not have a full column rank becoming larger, increasing the MSE a little.

Experiment 3: MSE Versus SNR: Now, we examine the channel estimation performance as a function of the SNR.

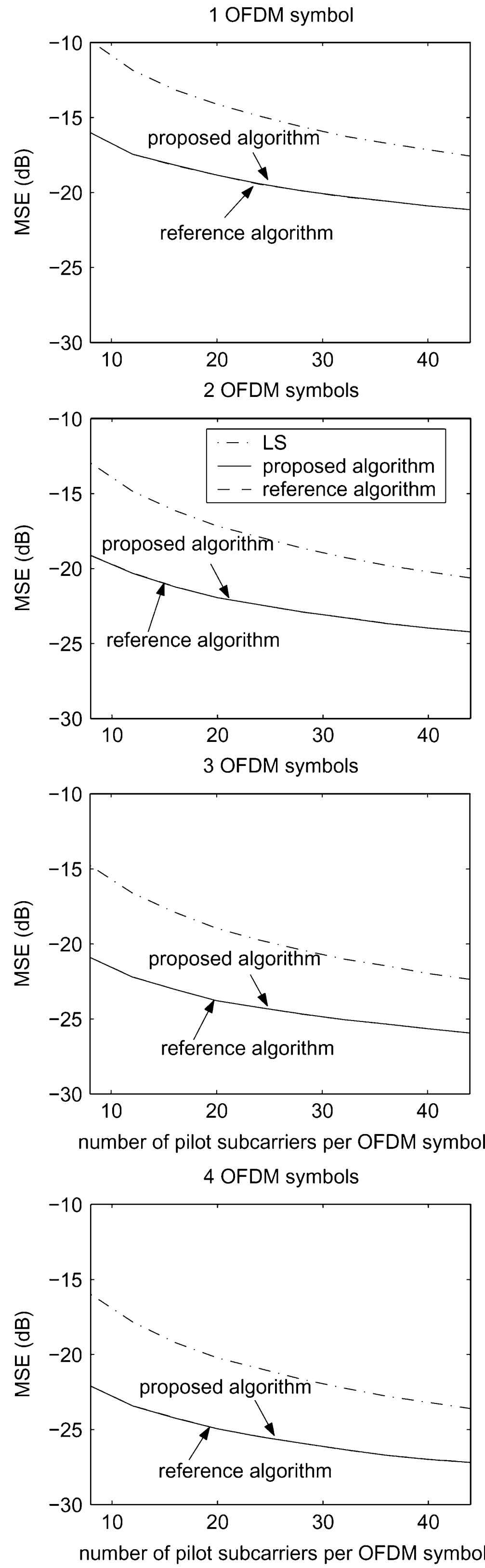

Fig. 4. MSE versus pilot length for different number of OFDM symbols.

Again, the simulation involves 5000 Monte Carlo runs of the transmission of one OFDM symbol. Fig. 3 shows the channel estimation results of the three methods, when $\eta > 0.2$. It is seen that the performances of the two semiblind algorithms are very close and both can achieve a gain of nearly 3.3 dB over the LS method regardless of the level of the SNR.

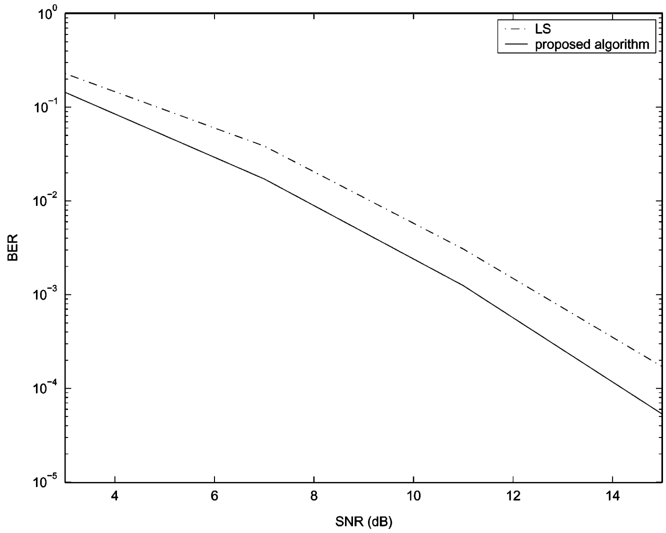

Fig. 5. BER versus SNR.

Experiment 4: MSE Versus Pilot Length: Here, we investigate the channel estimation performance of the proposed semiblind algorithm versus the number of OFDM symbols as well as the number of pilot subcarriers per symbol, in comparison with that of the LS method. The number of OFDM symbols is set to be from 1 to 4, and in each case, the number of pilot subcarriers per OFDM symbol varies from 8 to 48. Note that the proposed method can easily be applied to the case of multiple OFDM symbols, where the calculation of $\mathbf{Y}_{\text{pilot}}$, $\tilde{\mathbf{A}}$, and $\hat{\mathbf{B}}$ in (24) should be based on the multiple OFDM symbols. Fig. 4 shows the MSE plots from 500 Monte Carlo iterations for an SNR of 15 dB, when $\eta > 0.2$. It is seen that for the same number of OFDM symbols, the performance of all the algorithms improved with increasing number of pilot subcarriers. Again, the performance of the proposed method is very close to that of the reference method, both being superior to the LS method by 6 and 4 dB when the number of pilot subcarriers is 8 and 48, respectively. It implies that the proposed semiblind method is more advantageous for pilot signals of a shorter length. Furthermore, it is observed that the performance improvement of the proposed semiblind method over the LS method remains almost the same with the increase of the number of OFDM symbols employed for channel estimation.

Experiment 5: BER Versus SNR: In this experiment, the bit error rate (BER) performance of the MIMO–OFDM system is investigated by using the estimated channel matrix and an ordered vertical-Bell Laboratories layered space time (V-BLAST) decoder. The simulation involves 5000 Monte Carlo runs of the transmission of one OFDM symbol with eight pilot subcarriers. Fig. 5 shows the BER performance versus the SNR for the LS method and the proposed semiblind method, when $\eta > 0.2$. It is seen that the performance of the proposed semiblind algorithm is superior to the LS method by 2–5 dB.

VI. Conclusion

A semiblind MIMO–OFDM channel estimation approach that incorporates a linear-prediction-based blind criterion into the LS method was proposed. A practical yet very efficient scheme was presented for the determination of the weighting factor of the semiblind cost function. The perturbation analysis of the MIMO linear prediction justified the advantage of the semiblind method over the pure blind estimation, and led to a closed-form expression for the MSE of the blind criterion. The proposed method was simulated and compared with the training-based LS method, showing a significant improvement in terms of the MSE of the channel estimate.

Appendix A
Proof of Uncorrelation of Transmitted Time-Domain MIMO–OFDM Signal

Generally speaking, the frequency-domain signal in MIMO–OFDM systems before the DFT module can be considered as an i.i.d. Gaussian process with zero mean and variance δ_x^2, implying that the frequency-domain signal at the i_Tth antenna $X_{i_T}(k)$ is uncorrelated in both time and space domains. Given $X_{i_{T1}}(k)$ and $X_{i_{T2}}(k)$ being the frequency-domain signals at the i_{T1}th and i_{T2}th antennas, respectively, the corresponding time-domain signals can be written as

$$x_{i_{T1}}(n_1) = \frac{1}{\sqrt{K}} \sum_{k=0}^{K-1} X_{i_{T1}}(k) \mathrm{e}^{j2\pi(kn_1/K)}$$
$$x_{i_{T2}}(n_2) = \frac{1}{\sqrt{K}} \sum_{k=0}^{K-1} X_{i_{T2}}(k) \mathrm{e}^{j2\pi(kn_2/K)}.$$

Consider the correlation between the two time-domain signals, i.e.,

$$\begin{aligned} &\mathrm{E}\left\{x_{i_{T1}}(n_1)x_{i_{T2}}^*(n_2)\right\} \\ &\quad= \frac{1}{K}\sum_{k_1=0}^{K-1}\sum_{k_2=0}^{K-1} \mathrm{E}[X_{i_{T1}}(k_1)\, X_{i_{T2}}^*(k_2)]\mathrm{e}^{j2\pi(k_1n_1-k_2n_2)/K}. \end{aligned} \tag{A-1}$$

Obviously

$$\mathrm{E}\left\{x_{i_{T1}}(n_1)x_{i_{T2}}^*(n_2)\right\} = 0, \qquad \text{if } i_{T1} \neq i_{T2}. \tag{A-2}$$

If $i_{T1} = i_{T2}$, (A-1) reduces to

$$\begin{aligned} \mathrm{E}\left\{x_{i_{T1}}(n_1)x_{i_{T2}}^*(n_2)\right\} &= \frac{1}{K}\sum_{k=0}^{K-1} \delta_x^2 \mathrm{e}^{j2\pi k(n_1-n_2)/K} \\ &= \begin{cases} \delta_x^2, & \text{if } (n_1 = n_2) \\ 0, & \text{if } (n_1 \neq n_2). \end{cases} \end{aligned} \tag{A-3}$$

It is clear from (A-2) and (A-3) that the transmitted time-domain MIMO–OFDM signal is uncorrelated.

Appendix B
A Closed-Form Expression of $\mathrm{MSE_B}$

We formulate first $\mathrm{vec}(\mathbf{G}_1)$. Defining

$$\mathbf{\Phi}_1 \triangleq \begin{bmatrix} \mathbf{I}_{(N_R-N_T)N_T \times (N_R-N_T)N_T} \\ \mathbf{0}_{(N_R-N_T)(P+L-1)N_T \times (N_R-N_T)N_T} \end{bmatrix}$$

and using (63), we can derive

$$\begin{aligned} \mathrm{vec}(\mathbf{G}_1) &= \mathbf{\Phi}_1 \mathrm{vec}\left\{\mathbf{U}_{0\mathrm{null}}^H \mathbf{\Xi}^H (\mathbf{H}(0)\mathbf{H}^H(0))^{\dagger}\mathbf{H}(0)\right\} \\ &= \mathbf{\Omega}_1 \mathrm{vec}(\mathbf{\Xi}^H) \end{aligned} \tag{B-1}$$

where

$$\Omega_1 \triangleq \mathbf{\Phi}_1 \left\{[\mathbf{H}^T(0)(\mathbf{H}(0)\mathbf{H}^H(0))^{\dagger}] \otimes \mathbf{U}_{0\text{null}}^H\right\} \quad \text{(B-2)}$$

$$\text{vec}(\mathbf{\Xi}^H) = \text{vec}\left(\Delta\mathbf{R}_{v3}^H\right) - \mathbf{\Psi}_1 \text{vec}(\Delta\mathbf{R}_{v2}) - \mathbf{\Psi}_2 \text{vec}\left(\Delta\mathbf{R}_{v2}^H\right) + \mathbf{\Psi}_3 \text{vec}\left(\Delta\mathbf{R}_{v1}^H\right) \quad \text{(B-3)}$$

with

$$\mathbf{\Psi}_1 \triangleq \left[\left(\mathbf{\Pi}_2 \mathbf{I}_C^H \mathbf{H}_A^H\right)^T \otimes \mathbf{I}_{N_R}\right]$$
$$\mathbf{\Psi}_2 \triangleq \left[\mathbf{I}_{N_R}^T \otimes \left(\mathbf{H}_A \mathbf{I}_C \mathbf{\Pi}_2^H\right)\right]$$
$$\mathbf{\Psi}_3 \triangleq \left[\left(\mathbf{\Pi}_2 \mathbf{I}_C^H \mathbf{H}_A^H\right)^T \otimes \left(\mathbf{H}_A \mathbf{I}_C \mathbf{\Pi}_2^H\right)\right].$$

It is clear from (64) that, to obtain $\text{vec}(\mathbf{G}_2)$, one needs to first compute $\text{vec}(\mathbf{Q})$, which can be given by (see Appendix C for details)

$$\text{vec}(\mathbf{Q}) = \left(\mathbf{\Phi}_2^T \otimes \mathbf{I}_{(P+L)N_R}\right) \mathbf{\Phi}_3 \text{vec}\left(\mathbf{\Pi}_P\right) \quad \text{(B-4)}$$

where $\mathbf{\Phi}_2$ and $\mathbf{\Phi}_3$ are two transformation matrices which can be easily constructed as shown in Appendix C. Using (64) and (B-4) gives

$$\text{vec}(\mathbf{G}_2) = \mathbf{\Omega}_2 \text{vec}(\mathbf{\Pi}'_P) \quad \text{(B-5)}$$

where

$$\mathbf{\Omega}_2 \triangleq \left[\mathbf{I}_{(P+L)N_R} \otimes \left(\mathbf{I}_{P+L} \otimes \mathbf{U}_{0\text{null}}^H\right)\right] \times \left(\mathbf{\Phi}_2^T \otimes \mathbf{I}_{(P+L)N_R}\right) \mathbf{\Phi}_3 \quad \text{(B-6)}$$

$$\text{vec}(\mathbf{\Pi}'_P) = \mathbf{\Psi}_4 \text{vec}(\Delta\mathbf{R}_{v2}) - \mathbf{\Psi}_5 \text{vec}(\Delta\mathbf{R}_{v1}) \quad \text{(B-7)}$$

with $\mathbf{\Psi}_4 \triangleq \mathbf{\Pi}_2^T \otimes \mathbf{I}_{N_R}$ and $\mathbf{\Psi}_5 \triangleq \mathbf{\Pi}_2^T \otimes (\mathbf{H}_A \mathbf{I}_C \mathbf{\Pi}_2^H)$. Note that $\mathbf{\Pi}_P$ has been replaced by $\mathbf{\Pi}'_P$ in (B-5), since the signal perturbation term has been eliminated.

Finally, by using (B-1) and (B-5) in (65) and noting that $\Delta\mathbf{R}_{v1} = \Delta\mathbf{R}_{v1}^H$ and $\Delta\mathbf{R}_{v3} = \Delta\mathbf{R}_{v3}^H$, the MSE of the blind criterion can be derived as

$$\text{MSE}_\text{B} = \text{Trace}\left\{\mathbf{\Omega}_1 \left(\mathbf{\Gamma}_1 + \mathbf{\Gamma}_2 + \mathbf{\Gamma}_2^H\right) \mathbf{\Omega}_1^H + \mathbf{\Omega}_2 \mathbf{\Gamma}_3 \mathbf{\Omega}_2^H - \mathbf{\Omega}_1 \mathbf{\Gamma}_4 \mathbf{\Omega}_2^H - \mathbf{\Omega}_2 \mathbf{\Gamma}_4^H \mathbf{\Omega}_1^H\right\} \quad \text{(B-8)}$$

where

$$\mathbf{\Gamma}_1 = \mathbf{R}_{\Delta v4} + \mathbf{\Psi}_1 \mathbf{R}_{\Delta v2} \mathbf{\Psi}_1^H + \mathbf{\Psi}_2 \mathbf{R}_{\Delta v3} \mathbf{\Psi}_2^H + \mathbf{\Psi}_3 \mathbf{R}_{\Delta v1} \mathbf{\Psi}_3^H$$
$$\mathbf{\Gamma}_2 = -\mathbf{R}_{\Delta v9}^H \mathbf{\Psi}_1^H - \mathbf{R}_{\Delta v10}^H \mathbf{\Psi}_2^H + \mathbf{R}_{\Delta v7}^H \mathbf{\Psi}_3^H + \mathbf{\Psi}_1 \mathbf{R}_{\Delta v8} \mathbf{\Psi}_2^H - \mathbf{\Psi}_1 \mathbf{R}_{\Delta v5}^H \mathbf{\Psi}_3^H - \mathbf{\Psi}_2 \mathbf{R}_{\Delta v6}^H \mathbf{\Psi}_3^H$$
$$\mathbf{\Gamma}_3 = \mathbf{\Psi}_4 \mathbf{R}_{\Delta v2} \mathbf{\Psi}_4^H - \mathbf{\Psi}_4 \mathbf{R}_{\Delta v5}^H \mathbf{\Psi}_5^H - \mathbf{\Psi}_5 \mathbf{R}_{\Delta v5} \mathbf{\Psi}_4^H + \mathbf{\Psi}_5 \mathbf{R}_{\Delta v1} \mathbf{\Psi}_5^H$$
$$\mathbf{\Gamma}_4 = \mathbf{R}_{\Delta v9}^H \mathbf{\Psi}_4^H - \mathbf{R}_{\Delta v7}^H \mathbf{\Psi}_5^H - \mathbf{\Psi}_1 \mathbf{R}_{\Delta v2} \mathbf{\Psi}_4^H + \mathbf{\Psi}_1 \mathbf{R}_{\Delta v5}^H \mathbf{\Psi}_5^H - \mathbf{\Psi}_2 \mathbf{R}_{\Delta v8}^H \mathbf{\Psi}_4^H + \mathbf{\Psi}_2 \mathbf{R}_{\Delta v6}^H \mathbf{\Psi}_5^H + \mathbf{\Psi}_3 \mathbf{R}_{\Delta v5} \mathbf{\Psi}_4^H - \mathbf{\Psi}_3 \mathbf{R}_{\Delta v1} \mathbf{\Psi}_5^H$$

with $\mathbf{R}_{\Delta vi}, (i = 1, \ldots, 10)$ being given by

$$\mathbf{R}_{\Delta v1} = \text{E}\left\{\text{vec}\left(\Delta\mathbf{R}_{v1}\right) \text{vec}^H\left(\Delta\mathbf{R}_{v1}\right)\right\}$$
$$\mathbf{R}_{\Delta v2} = \text{E}\left\{\text{vec}\left(\Delta\mathbf{R}_{v2}\right) \text{vec}^H\left(\Delta\mathbf{R}_{v2}\right)\right\}$$
$$\mathbf{R}_{\Delta v3} = \text{E}\left\{\text{vec}\left(\Delta\mathbf{R}_{v2}^H\right) \text{vec}^H\left(\Delta\mathbf{R}_{v2}^H\right)\right\}$$
$$\mathbf{R}_{\Delta v4} = \text{E}\left\{\text{vec}\left(\Delta\mathbf{R}_{v3}\right) \text{vec}^H\left(\Delta\mathbf{R}_{v3}\right)\right\}$$
$$\mathbf{R}_{\Delta v5} = \text{E}\left\{\text{vec}\left(\Delta\mathbf{R}_{v1}\right) \text{vec}^H\left(\Delta\mathbf{R}_{v2}\right)\right\}$$
$$\mathbf{R}_{\Delta v6} = \text{E}\left\{\text{vec}\left(\Delta\mathbf{R}_{v1}\right) \text{vec}^H\left(\Delta\mathbf{R}_{v2}^H\right)\right\}$$
$$\mathbf{R}_{\Delta v7} = \text{E}\left\{\text{vec}\left(\Delta\mathbf{R}_{v1}\right) \text{vec}^H\left(\Delta\mathbf{R}_{v3}\right)\right\}$$
$$\mathbf{R}_{\Delta v8} = \text{E}\left\{\text{vec}\left(\Delta\mathbf{R}_{v2}\right) \text{vec}^H\left(\Delta\mathbf{R}_{v2}^H\right)\right\}$$
$$\mathbf{R}_{\Delta v9} = \text{E}\left\{\text{vec}\left(\Delta\mathbf{R}_{v2}\right) \text{vec}^H\left(\Delta\mathbf{R}_{v3}\right)\right\}$$
$$\mathbf{R}_{\Delta v10} = \text{E}\left\{\text{vec}\left(\Delta\mathbf{R}_{v2}^H\right) \text{vec}^H\left(\Delta\mathbf{R}_{v3}\right)\right\}.$$

The closed-form expression of $\mathbf{R}_{\Delta vi}, (i = 1, \ldots, 10)$ is derived in Appendix D.

APPENDIX C
DERIVATION OF (B-4)

By defining a transformation matrix $\mathbf{\Phi}_2 = [\mathbf{0}, \mathbf{\Phi}_{2,1}^H, \ldots, \mathbf{\Phi}_{2,P+L-1}^H]^H$ such that its submatrices, being of the same size, satisfy $[\mathbf{\Phi}_{2,1}, \ldots, \mathbf{\Phi}_{2,P+L-1}] = \mathbf{I}_{(P+L-1)N_T}$, (60) can be rewritten as

$$\mathbf{Q} = (\mathbf{I}_{P+L} \otimes \mathbf{\Pi}_P)\mathbf{\Phi}_2 \quad \text{(C-1)}$$

from which we have

$$\text{vec}(\mathbf{Q}) = \left(\mathbf{\Phi}_2^T \otimes \mathbf{I}_{(P+L)N_R}\right) \text{vec}(\mathbf{I}_{P+L} \otimes \mathbf{\Pi}_P). \quad \text{(C-2)}$$

We now simplify $\text{vec}(\mathbf{Q})$ by introducing another transformation matrix $\mathbf{\Phi}_3$. Our idea is to construct $\mathbf{\Phi}_3$ such that

$$\text{vec}(\mathbf{I}_{P+L} \otimes \mathbf{\Pi}_P) = \mathbf{\Phi}_3 \text{vec}(\mathbf{\Pi}_P)$$

which leads $\text{vec}(\mathbf{Q})$ to

$$\text{vec}(\mathbf{Q}) = \left(\mathbf{\Phi}_2^T \otimes \mathbf{I}_{(P+L)N_R}\right) \mathbf{\Phi}_3 \text{vec}(\mathbf{\Pi}_P). \quad \text{(C-3)}$$

To obtain such a $\mathbf{\Phi}_3$, we define a $[(P+L)^2(P+L-1)N_R N_T]$-dimensional vector, denoted as

$$\mathbf{i}_{F2} = \text{vec}(\mathbf{I}_{P+L} \otimes \mathbf{I}_{F1})$$

where $\mathbf{I}_{F1}$ is an $N_R \times (P+L-1)N_T$ matrix whose elements are all unity. The matrix $\mathbf{\Phi}_3$ of size $[(P+L)^2(P+L-1)N_R N_T] \times [(P+L-1)N_R N_T]$ can then be constructed according to the following rules.

1) If the nth element of vector $\mathbf{i}_{F2}$ is zero, the nth row of $\mathbf{\Phi}_3$ is set to a zero vector.
2) If the nth element of vector $\mathbf{i}_{F2}$ is 1, then only one element in the nth row of $\mathbf{\Phi}_3$ is 1 and all the other elements are zero. The position of the element 1 shifts to the right by one entry with respect to the previous row with an element of 1. If the element 1 in the previous reference row happens at the last column, the next element 1 returns to the first column. Note that the first element in the first row of $\mathbf{\Phi}_3$ is always

$$\mathbf{i}_{F2} = \text{vec}(\mathbf{I}_3 \otimes [1,1]) = [1 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 1]^T.$$

1. As an example, considering $P+L=3,$, $N_R=1$, and $N_T=1$, we have the equation shown at the bottom of the previous page. Then, the corresponding $\mathbf{\Phi}_3$ is readily given by the equation shown at the bottom of the page.

Appendix D
Derivation of $\mathbf{R}_{\Delta vi}, (i=1,2,\ldots,10)$

Here, we derive the autocorrelation and cross-correlation matrices of $\mathrm{vec}(\Delta\mathbf{R}_{vi})$ and $\mathrm{vec}(\Delta\mathbf{R}_{vi}^H)$, $(i=1,2,3)$. First, we consider $\mathbf{R}_{\Delta v1}$. For a medium to high SNR, the perturbation matrix of the noise autocorrelation matrix can be neglected. Thus, from (36), one can get

$$\mathrm{vec}(\Delta\mathbf{R}_{v1}) \approx (\mathbf{I}_{PN_R}\otimes\mathbf{H}_B)\,\mathrm{vec}(\Delta\mathbf{R}_{xv1}) + (\mathbf{H}_B^*\otimes\mathbf{I}_{PN_R})\,\mathrm{vec}\left(\Delta\mathbf{R}_{xv1}^H\right).$$

Since the noise element $v_{i_R}(n)$ is circularly symmetric with equal variance $\delta_v^2/2$ for the real and imaginary part, one can verify that $\mathrm{E}\{v_{i_R}(n)v_{i_R}(n)\}=(1/2)\delta_v^2-(1/2)\delta_v^2=0$ and $\mathrm{E}\{v_{i_R}(n)v_{i_R}^*(n)\}=(1/2)\delta_v^2+(1/2)\delta_v^2=\delta_v^2$. In addition, it can be easily verified that $\mathrm{E}[\mathrm{vec}(\Delta\mathbf{R}_{xv1})\mathrm{vec}^H(\Delta\mathbf{R}_{xv1}^H)]=\mathbf{0}$. Therefore, $\mathbf{R}_{\Delta v1}$ can be calculated by

$$\begin{aligned}\mathbf{R}_{\Delta v1} \approx{}& (\mathbf{I}_{PN_R}\otimes\mathbf{H}_B)\,\mathrm{E}[\mathrm{vec}\,(\Delta\mathbf{R}_{xv1})\,\mathrm{vec}^H\,(\Delta\mathbf{R}_{xv1})]\\ &\times\left(\mathbf{I}_{PN_R}\otimes\mathbf{H}_B^H\right)+(\mathbf{H}_B^*\otimes\mathbf{I}_{PN_R})\,\mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv1}^H\right)\right.\\ &\times\left.\mathrm{vec}^H\left(\Delta\mathbf{R}_{xv1}^H\right)\right\}\left(\mathbf{H}_B^T\otimes\mathbf{I}_{PN_R}\right).\end{aligned}\tag{D-1}$$

To this end, the following two correlation matrices are to be determined:

$$\mathbf{\Upsilon}_1 \triangleq \mathrm{E}\{\mathrm{vec}(\Delta\mathbf{R}_{xv1})\mathrm{vec}^H(\Delta\mathbf{R}_{xv1})\}\tag{D-2}$$

$$\mathbf{\Upsilon}_1' \triangleq \mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv1}^H\right)\mathrm{vec}^H\left(\Delta\mathbf{R}_{xv1}^H\right)\right\}.\tag{D-3}$$

To compute $\mathbf{\Upsilon}_1$, one can obtain from (37)

$$\mathrm{vec}(\Delta\mathbf{R}_{xv1}) = \frac{1}{K}\sum_{n=0}^{K-1}(\mathbf{X}_{\mathbf{I}1}(n)\mathbf{v}_P(n-1))\tag{D-4}$$

where $\mathbf{X}_{\mathbf{I}1}(n)\triangleq\mathbf{I}_{PN_R}\otimes\mathbf{x}_{P+L-1}(n-1)$. Substituting (D-4) into (D-2) gives

$$\mathbf{\Upsilon}_1 \triangleq \frac{1}{K^2}\mathrm{E}\left\{\sum_{n_1=0}^{K-1}\sum_{n_2=0}^{K-1}\mathbf{X}_{\mathbf{I}1}(n_1)\mathbf{R}_{\mathbf{IV}1}(n_1-n_2)\mathbf{X}_{\mathbf{I}1}^H(n_2)\right\}\tag{D-5}$$

where $\mathbf{R}_{\mathbf{IV}1}(n_1-n_2)\triangleq\mathrm{E}(\mathbf{v}_P(n_1-1)\mathbf{v}_P^H(n_2-1))$. It can be easily shown that $\mathbf{R}_{\mathbf{IV}1}(n_1-n_2)$ is a partial identity matrix. Further, it can be proved that

$$\begin{aligned}&\mathbf{X}_{\mathbf{I}1}(n_1)\mathbf{R}_{\mathbf{IV}1}(n_1-n_2)\mathbf{X}_{\mathbf{I}1}^H(n_2)\\ &\quad=\mathbf{R}_{\mathbf{IV}1}(n_1-n_2)\otimes\mathbf{R}_{\mathbf{IX}1}(n_1-n_2)\end{aligned}\tag{D-6}$$

where

$$\mathbf{R}_{\mathbf{IX}1}(n_1-n_2)\triangleq\mathrm{E}(\mathbf{x}_{P+L-1}(n_1-1)\mathbf{x}_{P+L-1}^H(n_2-1))$$

which is also a partial identity matrix. Thus, substituting (D-6) into (D-5) gives

$$\mathbf{\Upsilon}_1 = \frac{1}{K}\sum_{i=1-P}^{P-1}\mathbf{R}_{\mathbf{IV}1}(i)\otimes\mathbf{R}_{\mathbf{IX}1}(i).\tag{D-7}$$

Similarly, one can derive

$$\mathbf{\Upsilon}_1' = \frac{1}{K}\sum_{i=1-P}^{P-1}\mathbf{R}_{\mathbf{IX}1}(i)\otimes\mathbf{R}_{\mathbf{IV}1}(i).\tag{D-8}$$

Thus, by using (D-7) and (D-8) into (D-1), one can obtain

$$\begin{aligned}\mathbf{R}_{\Delta v1}\approx{}&(\mathbf{I}_{PN_R}\otimes\mathbf{H}_B)\,\mathbf{\Upsilon}_1\left(\mathbf{I}_{PN_R}\otimes\mathbf{H}_B^H\right)\\ &+(\mathbf{H}_B^*\otimes\mathbf{I}_{PN_R})\,\mathbf{\Upsilon}_1'\left(\mathbf{H}_B^T\otimes\mathbf{I}_{PN_R}\right).\end{aligned}\tag{D-9}$$

As a result, $\mathbf{R}_{\Delta v1}$ can now be explicitly computed in terms of $\mathbf{\Upsilon}_1$, $\mathbf{\Upsilon}_1'$, and the channel matrix $\mathbf{H}_B$.

In a similar manner, one can derive the expressions for $\mathbf{R}_{\Delta vi}, (i=2,3,\ldots,10)$

$$\begin{aligned}
\mathbf{R}_{\Delta v2}\approx{}&(\mathbf{I}_{PN_R}\otimes\mathbf{H}_A)\,\mathbf{\Upsilon}_2\left(\mathbf{I}_{PN_R}\otimes\mathbf{H}_A^H\right)\\ &+(\mathbf{H}_B^*\otimes\mathbf{I}_{N_R})\,\mathbf{\Upsilon}_2'\left(\mathbf{H}_B^T\otimes\mathbf{I}_{N_R}\right)\\
\mathbf{R}_{\Delta v3}\approx{}&(\mathbf{H}_A^*\otimes\mathbf{I}_{PN_R})\,\mathbf{\Upsilon}_3\left(\mathbf{H}_A^T\otimes\mathbf{I}_{PN_R}\right)\\ &+(\mathbf{I}_{N_R}\otimes\mathbf{H}_B)\,\mathbf{\Upsilon}_3'\left(\mathbf{I}_{N_R}\otimes\mathbf{H}_B^H\right)\\
\mathbf{R}_{\Delta v4}\approx{}&(\mathbf{I}_{N_R}\otimes\mathbf{H}_A)\,\mathbf{\Upsilon}_4\left(\mathbf{I}_{N_R}\otimes\mathbf{H}_A^H\right)\\ &+(\mathbf{H}_A^*\otimes\mathbf{I}_{N_R})\,\mathbf{\Upsilon}_4\left(\mathbf{H}_A^T\otimes\mathbf{I}_{N_R}\right)\\
\mathbf{R}_{\Delta v5}\approx{}&(\mathbf{I}_{PN_R}\otimes\mathbf{H}_B)\,\mathbf{\Upsilon}_5\left(\mathbf{I}_{PN_R}\otimes\mathbf{H}_A^H\right)\\ &+(\mathbf{H}_B^*\otimes\mathbf{I}_{PN_R})\,\mathbf{\Upsilon}_5'\left(\mathbf{H}_B^T\otimes\mathbf{I}_{N_R}\right)\\
\mathbf{R}_{\Delta v6}\approx{}&(\mathbf{I}_{PN_R}\otimes\mathbf{H}_B)\,\mathbf{\Upsilon}_6\left(\mathbf{I}_{N_R}\otimes\mathbf{H}_B^H\right)\\ &+(\mathbf{H}_B^*\otimes\mathbf{I}_{PN_R})\,\mathbf{\Upsilon}_6'\left(\mathbf{H}_A^T\otimes\mathbf{I}_{PN_R}\right)\\
\mathbf{R}_{\Delta v7}\approx{}&(\mathbf{I}_{PN_R}\otimes\mathbf{H}_B)\,\mathbf{\Upsilon}_7\left(\mathbf{I}_{N_R}\otimes\mathbf{H}_A^H\right)\\ &+(\mathbf{H}_B^*\otimes\mathbf{I}_{PN_R})\,\mathbf{\Upsilon}_7'\left(\mathbf{H}_A^T\otimes\mathbf{I}_{N_R}\right)\\
\mathbf{R}_{\Delta v8}\approx{}&(\mathbf{I}_{PN_R}\otimes\mathbf{H}_A)\,\mathbf{\Upsilon}_8\left(\mathbf{I}_{N_R}\otimes\mathbf{H}_B^H\right)\\ &+(\mathbf{H}_B^*\otimes\mathbf{I}_{N_R})\,\mathbf{\Upsilon}_8'\left(\mathbf{H}_A^T\otimes\mathbf{I}_{PN_R}\right)\\
\mathbf{R}_{\Delta v9}\approx{}&(\mathbf{I}_{PN_R}\otimes\mathbf{H}_A)\,\mathbf{\Upsilon}_9\left(\mathbf{I}_{N_R}\otimes\mathbf{H}_A^H\right)\\ &+(\mathbf{H}_B^*\otimes\mathbf{I}_{N_R})\,\mathbf{\Upsilon}_9'\left(\mathbf{H}_A^T\otimes\mathbf{I}_{N_R}\right)\\
\mathbf{R}_{\Delta v10}\approx{}&(\mathbf{H}_A^*\otimes\mathbf{I}_{PN_R})\,\mathbf{\Upsilon}_{10}\left(\mathbf{H}_A^T\otimes\mathbf{I}_{N_R}\right)\\ &+(\mathbf{I}_{N_R}\otimes\mathbf{H}_B)\,\mathbf{\Upsilon}_{10}'\left(\mathbf{I}_{N_R}\otimes\mathbf{H}_A^H\right)
\end{aligned}$$

$$\begin{bmatrix}1&0&0&0&0&0&0&1&0&0&0&0&0&1&0&0&0\\0&0&0&1&0&0&0&0&0&0&1&0&0&0&0&0&1\end{bmatrix}^T.$$

where

$$\Upsilon_2 = \frac{1}{K}\sum_{i=1-L}^{L-1} \mathbf{R_{IV1}}(i) \otimes \mathbf{R_{IX2}}(i)$$
$$\Upsilon'_2 = \frac{\delta_x^2\delta_v^2}{K}\mathbf{I}_{N_R N_T (P+L-1)}$$
$$\Upsilon_3 = \frac{1}{K}\sum_{i=1-L}^{L-1} \mathbf{R_{IX2}}(i) \otimes \mathbf{R_{IV1}}(i)$$
$$\Upsilon'_3 = \frac{\delta_x^2\delta_v^2}{K}\mathbf{I}_{N_R N_T (P+L-1)}$$
$$\Upsilon_4 = \frac{\delta_x^2\delta_v^2}{K}\mathbf{I}_{N_R N_T L}$$
$$\Upsilon_5 = \frac{1}{K}\sum_{i=1-P}^{P-1} \mathbf{R_{IV1}}(i) \otimes \mathbf{R_{IX3}}(i)$$
$$\Upsilon'_5 = \frac{1}{K}\sum_{i=1}^{P} \mathbf{R_{IX1}}(i) \otimes \mathbf{R_{IV2}}(i)$$
$$\Upsilon_6 = \frac{1}{K}\sum_{i=1}^{P} \mathbf{R_{IV2}}(i) \otimes \mathbf{R_{IX1}}(i)$$
$$\Upsilon'_6 = \frac{1}{K}\sum_{i=1-P}^{P-1} \mathbf{R_{IX3}}(i) \otimes (i)\mathbf{R_{IV1}}(i)$$
$$\Upsilon_7 = \frac{1}{K}\sum_{i=1}^{P} \mathbf{R_{IV2}}(i) \otimes \mathbf{R_{IX3}}(i)$$
$$\Upsilon'_7 = \frac{1}{K}\sum_{i=1}^{P} \mathbf{R_{IX3}}(i) \otimes \mathbf{R_{IV2}}(i)$$
$$\Upsilon_8 = \frac{1}{K}\sum_{i=1}^{L-2} \mathbf{R_{IV2}}(i) \otimes \mathbf{R_{IX4}}(i)$$
$$\Upsilon'_8 = \frac{1}{K}\sum_{i=1-L}^{-1} \mathbf{R_{IX3}}(i) \otimes \mathbf{R_{IV3}}(i)$$
$$\Upsilon_9 = \frac{1}{K}\sum_{i=1}^{L-1} \mathbf{R_{IV2}}(i) \otimes \mathbf{R_{IX2}}(i)$$
$$\Upsilon'_9 = \frac{\delta_v^2}{K}\mathbf{R_{IX3}}(0) \otimes \mathbf{I}_{N_R}$$
$$\Upsilon_{10} = \frac{1}{K}\sum_{i=1}^{L-1} \mathbf{R_{IX2}}(i) \otimes \mathbf{R_{IV2}}(i)$$
$$\Upsilon'_{10} = \frac{\delta_v^2}{K}\mathbf{I}_{N_R} \otimes \mathbf{R_{IX3}}(0)$$

with

$$\mathbf{R_{IX2}}(i = n_1 - n_2) = \mathrm{E}\left(\mathbf{x}_L(n_1)\mathbf{x}_L^H(n_2)\right)$$
$$\mathbf{R_{IX3}}(i = n_1 - n_2) = \mathrm{E}\left(\mathbf{x}_{P+L-1}(n_1 - 1)\mathbf{x}_L^H(n_2)\right)$$
$$\mathbf{R_{IX4}}(i = n_1 - n_2) = \mathrm{E}\left(\mathbf{x}_L(n_1)\mathbf{x}_{P+L-1}^H(n_2 - 1)\right)$$
$$\mathbf{R_{IV2}}(i = n_1 - n_2) = \mathrm{E}\left(\mathbf{v}_P(n_1 - 1)\mathbf{v}^H(n_2)\right)$$
$$\mathbf{R_{IV3}}(i = n_1 - n_2) = \mathrm{E}\left(\mathbf{v}(n_1)\mathbf{v}_P^H(n_2 - 1)\right).$$

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, Feb. 2004.

[2] 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, Jun. 2003.

[3] 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, Mar. 2005.

[4] J. Zhang, Z. He, and Y. Huang, "TSK fuzzy approach to channel estimation for MIMO-OFDM systems," *IEEE Signal Process. Lett.*, vol. 14, no. 6, pp. 381–384, Jun. 2007.

[5] R. Chen, H. Zhang, Y. Xu, and H. Luo, "On MM-type channel estimation for MIMO OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 6, no. 3, pp. 1046–1055, May 2007.

[6] C. Shin, R. W. Heath, and E. J. Powers, "Blind channel estimation for MIMO-OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 56, no. 2, pp. 670–685, Mar. 2007.

[7] F. Gao and A. Nallanathan, "Blind channel estimation for MIMO OFDM systems via nonredundant linear precoding," *IEEE Trans. Signal Process.*, vol. 55, no. 2, pp. 784–789, Feb. 2007.

[8] Y. Zeng, W. H. Lam, and T. Ng, "Semiblind channel estimation and equalization for MIMO space-time coded OFDM," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 53, no. 2, pp. 463–474, Feb. 2006.

[9] A. Medles, D. T. M. Slock, and E. De Carvalho, "Linear prediction based semi-blind estimation of MIMO FIR channels," in *Proc. IEEE 3rd Workshop Signal Process. Adv. Wireless Commun.*, 2001, pp. 58–61.

[10] 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.

[11] A. Medles and D. T. M. Slock, "Linear precoding for spatial multiplexing MIMO systems: Blind channel estimation aspects," in *Proc. IEEE Int. Conf. Commun.*, 2002, vol. 1, pp. 401–405.

[12] A. Gorokhov and P. Loubaton, "Blind identification of MIMO-FIR systems: A generalized linear prediction approach," *Signal Process.*, vol. 73, no. 1–2, pp. 105–124, 1999.

[13] Y. Inouye and R. Liu, "A system-theoretic foundation for blind equalization of an FIR MIMO channel system," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 49, no. 4, pp. 425–436, Apr. 2002.

[14] R. Lopez-Valcarce and S. Dasgupta, "Blind channel equalization with colored sources based on second-order statistics: A linear prediction approach," *IEEE Trans. Signal Process.*, vol. 49, no. 9, pp. 2050–2059, Sep. 2001.

[15] C. B. Papadias and D. T. M. Slock, "Fractionally spaced equalization of linear polyphase channels and related blind techniques based on multichannel linear prediction," *IEEE Trans. Signal Process.*, vol. 47, no. 3, pp. 641–654, Mar. 1999.

[16] 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, Nov. 1998.

[17] J. K. Tugnait and B. Huang, "Multistep linear predictors-based blind identification and equalization of multiple-input multiple-output channels," *IEEE Trans. Signal Process.*, vol. 48, no. 1, pp. 26–38, Jan. 2000.

[18] 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, Nov. 1998.

[19] 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, Nov. 1990.

[20] 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, Jul. 1993.

[21] Z. Xu, "Perturbation analysis for subspace decomposition with applications in subspace-based algorithms," *IEEE Trans. Signal Process.*, vol. 50, no. 11, pp. 2820–2830, Nov. 2002.

[22] Z. Xu, "On the second-order statistics of the weighted sample covariance matrix," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 527–534, Feb. 2003.

[23] Z. Xu, "Statistical performance of a data-based covariance estimator," *IEEE Trans. Veh. Technol.*, vol. 53, no. 3, pp. 939–943, May 2004.

[24] S. Houcke and P. Loubaton, "A weighted linear prediction approach for the blind CDMA forward link channel estimation," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2000, vol. 5, pp. 2937–2940.

[25] T. W. S. Chow, B. Wang, and K. T. Ng, "Linear prediction based multipath channel identification algorithm," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 50, no. 6, pp. 769–774, Jun. 2003.

[26] V. Buchoux, O. Cappe, E. Moulines, and A. Gorokhov, "On the performance of semi-blind subspace-based channel estimation," *IEEE Trans. Signal Process.*, vol. 48, no. 6, pp. 1750–1759, Jun. 2000.

[27] F. Wan, W.-P. Zhu, and M. N. S. Swamy, "Linear prediction based semi-blind channel estimation for MIMO-OFDM system," in *Proc. IEEE Int. Symp. Circuits Syst.*, New Orleans, LA, May 2007, pp. 3239–3242.

[28] J. Choi, "Equalization and semi-blind 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, Mar. 2004.

[29] C. Moreno, "A least-squares-based method for determining the ratio between two measured quantities," *Meas. Sci. Technol.*, vol. 7, pp. 137–141, 1996.

[30] J. Valyon and G. Horvath, "Extended least squares LS-SVM," *Int. J. Comput. Intell.*, vol. 3, no. 3, pp. 1304–2386, 2006.

[31] P. Tsai, H. Kang, and T. Chiueh, "Joint weighted least-squares estimation of carrier-frequency offset and timing offset for OFDM systems over multipath fading channels," *IEEE Trans. Veh. Technol.*, vol. 54, no. 1, pp. 211–223, Jan. 2005.

Feng Wan received the B.E. degree from Wuhan University, China, in 2000 and the M.E. degree from Institute of Electronics, Chinese Academy of Sciences, China, in 2003, both in electrical engineering. Currently, he is working towards the Ph.D. degree at Concordia University, Montreal, QC, Canada.

Since 2004, he has been a Research Assistant in Concordia University. His research interests are in the area of digital signal processing, statistical estimation theory and its applications in multiple-input–multiple-output (MIMO), orthogonal frequency-division multiplexing (OFDM), and smart antenna wireless communications.

Mr. Wan was awarded 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, Nanjing, China, in 1982 and 1985, respectively, and the Ph.D. degree from Southeast University, Nanjing, China, in 1991, all in electrical engineering.

He was a Postdoctoral Fellow from 1991 to 1992 and a Research Associate from 1996 to 1998 at the Department of Electrical and Computer Engineering, Concordia University, Montreal, QC, Canada. During 1993–1996, he was an Associate Professor at the Department of Information Engineering, Nanjing University of Posts and Telecommunications. From 1998 to 2001, he worked in hi-tech companies in Ottawa, ON, Canada, including Nortel Networks and SR Telecom, Inc. Since July 2001, he has been an Associate Professor at the Electrical and Computer Engineering Department, Concordia University. 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. Currently, he 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. (honors) degree in mathematics from Mysore University, Mysore, India, in 1954, the Diploma in electrical communication engineering from the Indian Institute of Science, Bangalore, India, in 1957, and the M.Sc. and Ph.D. degrees in electrical engineering from the University of Saskatchewan, Saskatoon, SK, Canada, in 1960 and 1963, respectively. In August 2001, he was awarded a Doctor of Science in Engineering (Honoris Causa) by Ansted University, Tortola, British Virgin Islands, "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."

Currently, he is 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 Chair of the Department of Electrical Engineering from 1970 to 1977 and Dean of Engineering and Computer Science from 1977 to 1993. He has also taught at the Electrical Engineering Department, Technical University of Nova Scotia, Halifax, NS, Canada, and the University of Calgary, Calgary, AB, Canada, as well as at 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), *Graphs: Theory and Algorithms* (New York: Wiley, 1992), *Switched Capacitor Filters: Theory, Analysis and Design* (London, U.K.: Prentice-Hall, 1995) and *Neural Networks in a Softcomputing Framework* (New York: Springer-Verlag, 2006). A Russian Translation of the first book was published by Mir Publishers, Moscow, Russia, in 1984, while a Chinese version was published by the Education Press, Beijing, China, in 1987. 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.

Dr. Swamy is a Fellow of the Institute of Electrical Engineers (U.K.), 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—PART I: FUNDAMENTAL THEORY AND APPLICATIONS from June 1999 to December 2001, Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS during June 1985–May 1987, Program Chair for the 1973 IEEE Circuits and Systems Symposium, General Chair for the 1984 IEEE Circuits and Systems Symposium, Vice-Chair for the 1999 IEEE Circuits and Systems Symposium, and a member of the Board of Governors of the CAS Society. He is the recipient of many IEEE Circuits and Systems Society awards including the Education Award in 2000, Golden Jubilee Medal in 2000, and the 1986 Guillemin–Cauer Best Paper Award.