This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the WCNC 2008 proceedings.

Subspace Blind MIMO-OFDM Channel Estimation with Short Averaging Periods: Performance Analysis

Chao-Cheng Tu and Benoît Champagne
Department of Electrical and Computer Engineering
McGill University
Montréal, Québec, Canada

***Abstract*—Among all blind channel estimation problems, subspace-based algorithms are attractive due to its fast-converging nature. It primarily exploits the orthogonality structure of the noise and signal subspaces by applying a signal-noise space decomposition to the correlation matrix of the received signal. In practice, the correlation matrix is unknown and must be estimated through time averaging over multiple time samples. To this end, the wireless channel must be time-invariant over a sufficient time interval, which may pose a problem for wideband applications. We proposed a novel subspace-based blind channel estimation algorithm with short time averaging periods, as obtained by exploiting the frequency correlation among adjacent OFDM subcarriers. In this paper, asymptotic performance bounds of the proposed algorithm are investigated by using perturbation analysis. We also present numerical results of the proposed as well as referenced subspace-based methods, including Cyclic Prefix and Virtual Carriers approaches. Based on the asymptotic performance bounds, the proposed scheme is justified in obtaining a desired correlation matrix efficiently by reducing the number of the OFDM blocks for time averaging up to 85%.**

I. Introduction

Multiple Input Multiple Output (MIMO) systems are created by deploying multiple antennas at both ends of a wireless link to increase the channel capacity and to mitigate the adverse effects of the wireless channels. Orthogonal Frequency Division Multiplexing (OFDM) modulation is created to provide high spectral efficiency and to eliminate the need for high-complexity equalization algorithms. Therefore, MIMO-OFDM, produced by employing multiple transmit and receive antennas in an OFDM system has becoming a practical alternative to single carrier and Single Input Single Output (SISO) transmission [1].

Among the recent studies of MIMO-OFDM, blind channel estimation has received great attention and has become a very vital area of research in recent years. Under multichannel or multirate models, blind channel estimation by using Second Order Statistics (SOS) potentially has faster convergence rates than that by using higher order statistics [2]. Amid these SOS blind approaches, subspace-based channel estimation is attractive since channel estimates can often be obtained in a closed-form from optimizing a quadratic cost function [3]. Without employing any precoding at the transmitter, a noise subspace-based method is proposed for OFDM systems by utilizing the redundancy introduced by the Cyclic Prefix (CP) [4], and it is further extended for MIMO-OFDM systems in [5]. Virtual Carriers (VC) are subcarriers that are set to zero without any information being transmitted. The presence of VC provides another useful resource that can be used for channel estimation. Such a scheme is proposed for OFDM systems [6], and it is further extended for MIMO-OFDM systems in [7].

The aforementioned approaches primarily exploit the separability of the noise and signal subspaces by applying the Eigenvalue Decomposition (EVD) to the correlation matrix of the received signal. In practice, the correlation matrix can only be estimated by averaging over multiple time samples, given the wireless channel is time-invariant during this averaging period. Since the quadratic cost function is constructed from the eigenvectors of the noise subspace obtained from the EVD, the accuracy of the eigenvectors obtained from the sampled correlation matrix dominates the performance of the estimation. Hence, the more time samples are averaged, the better the estimation performance is. However, how many samples are sufficient to obtain a sampled correlation matrix meeting a certain level of confidence? A basic rule of thumb is: the number of the time samples must be larger than or equal to the order of the correlation matrix so as to make it full rank or invertible [8][9]. Simulation results have shown that at least 250 OFDM blocks (or 500 OFDM symbols) are required in order to achieve a normalized root mean square error (NRMSE) = 10^{-2} when we consider an IFFT size of 16 and SNR = 20dB for the subspace-based approaches [7]. If the size of IFFT were increased to 64, the required OFDM blocks would increase up to thousands for time averaging [10], making these subspace-based blind channel estimation approaches impractical.

We proposed a new subspace-based blind channel estimator for MIMO-OFDM systems with short time averaging periods of the correlation matrix [11]. This is achieved by exploiting the frequency correlation among adjacent OFDM subcarriers through the concept of subcarrier grouping [12][13]. The proposed scheme requires only an upper bound of the channel order, and the ambiguity matrix which is embedded in all the subspace-based estimation problems can be solved by optimization. In this paper, asymptotic performance bounds of the proposed method are derived by using perturbation analysis. Numerical results are also presented to support our claims and designs.

The rest of this paper is organized as follows. In section

1525-3511/08/$25.00 ©2008 IEEE

II, the system model is described. In section III, the subspace blind MIMO-OFDM channel estimator [11] will be briefly reviewed. Asymptotic performance bounds of the proposed estimator will be studied in section IV by the perturbation analysis. Numerical results will then be presented in section V, and conclusions will be drawn in the final section.

The notation used in this paper is as follows: The range of $\mathbf{A}$, denoted by $\Re(\mathbf{A})$, is defined by $\Re(\mathbf{A}) := \{\mathbf{y} : \mathbf{y} = \mathbf{A}\mathbf{x} \text{ for some } \mathbf{x}\}$. $E[x]$ denotes the expected value of the random variable x. $\otimes$ denotes the Kronecker product. $\mathrm{vec}(\cdot)$ is the Vec operator. $\mathrm{tr}(\mathbf{A})$ denotes the trace of a square matrix $\mathbf{A}$. $\|\mathbf{x}\|_1$ represents the L1 norm of a vector $\mathbf{x}$, and $\|\mathbf{X}\|_F$ denotes the Frobenius norm of a matrix $\mathbf{X}$. $\mathbf{X}^\dagger$ denotes the pseudo-inverse of a matrix $\mathbf{X}$. $\bigoplus \sum_i \mathbf{X}_i = \mathrm{diag}(\mathbf{X}_i)$ denotes the direct sum of matrices $\mathbf{X}_i$ [14].

II. System Model

We consider a MIMO-OFDM system with N_T transmit and N_R receive antennas, employing N_C subcarriers as shown in Fig. 1. Let the mth OFDM symbol transmitted over the kth subcarrier be denoted as $\mathbf{x}^m[k] := \left[x_1^m[k] \; x_2^m[k] \; \cdots \; x_{N_T}^m[k] \right]^T$, where $x_q^m[k]$ is the symbol transmitted at the qth transmit antenna. Then the mth OFDM symbol transmitted over N_C subcarriers can be written as $\mathbf{x}^m := \left[\mathbf{x}^m[0]^T \; \mathbf{x}^m[1]^T \; \cdots \; \mathbf{x}^m[N_C-1]^T \right]^T$. Assume the incoming symbol streams span over N_F OFDM symbols at each epoch, then $\mathbf{x} = [\mathbf{x}^{1^T} \; \mathbf{x}^{2^T} \; \cdots \; \mathbf{x}^{N_F{}^T}]^T$ represents the complete set of transmitted symbols. At the receiver, let the mth received OFDM symbol over the kth subcarrier be denoted as $\mathbf{y}^m[k] := \left[y_1^m[k] \; y_2^m[k] \; \cdots \; y_{N_R}^m[k] \right]^T$, where $y_p^m[k]$ is the symbol received at the pth receive antenna. Then the mth OFDM symbol received over N_C subcarriers can be written as $\mathbf{y}^m := \left[\mathbf{y}^m[0]^T \; \mathbf{y}^m[1]^T \; \cdots \; \mathbf{y}^m[N_C-1]^T \right]^T$, and $\mathbf{y} = [\mathbf{y}^{1^T} \; \mathbf{y}^{2^T} \; \cdots \; \mathbf{y}^{N_F{}^T}]^T$ represents the complete set of received symbols. In the following, we assume that (1) the length of the OFDM cyclic prefix is greater than the maximum excess delay of the channel, (2) the channel is time-invariant over N_F OFDM symbols, and (3) $E[|x_q^m[k]|^2] = 1$.

For MIMO-OFDM systems, the signal received at the pth receive antenna over the kth subcarrier and the mth OFDM symbol is given by

$$y_p^m[k] = \sqrt{\frac{E_s}{N_T}} \sum_{q=1}^{N_T} h_{p,q}[k] x_q^m[k] + n_p^m[k], \; p = 1, 2, \cdots, N_R, \tag{1}$$

where E_s is the average energy evenly divided across the transmit antennas and allocated to the kth subcarrier, $n_p^m[k]$ represents the zero mean circularly symmetric complex Gaussian (ZMCSCG) noise at the pth receive antenna over the kth subcarrier and the mth OFDM symbol, and $h_{p,q}[k]$ represents the equivalent frequency response between the pth receive antenna and the qth transmit antenna, over the kth subcarrier and the mth OFDM symbol. Let $\mathbf{n}^m[k] := \left[n_1^m[k] \; n_2^m[k] \; \cdots \; n_{N_R}^m[k] \right]^T$, $\mathbf{n}^m := \left[\mathbf{n}^m[0]^T \; \mathbf{n}^m[1]^T \; \cdots \; \mathbf{n}^m[N_C-1]^T \right]^T$, and $\mathbf{n} = [\mathbf{n}^{1^T} \; \mathbf{n}^{2^T} \; \cdots \; \mathbf{n}^{N_F{}^T}]^T$, then the input-output relation of the MIMO-OFDM systems may be expressed by

$$\mathbf{y} = \sqrt{\frac{E_s}{N_T}} \mathcal{H} \cdot \mathbf{x} + \mathbf{n}, \tag{2}$$

where $\mathcal{H} := \mathbf{I}_{N_F} \otimes (\bigoplus \sum_{k=0}^{N_C-1} \mathbf{H}[k])$, with diagonal blocks defined as

$$\mathbf{H}[k] = \begin{bmatrix} h_{1,1}[k] & h_{1,2}[k] & \cdots & h_{1,N_T}[k] \\ h_{2,1}[k] & h_{2,2}[k] & \cdots & h_{2,N_T}[k] \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_R,1}[k] & h_{N_R,2}[k] & \cdots & h_{N_R,N_T}[k] \end{bmatrix} \in \mathbb{C}^{N_R \times N_T}. \tag{3}$$

III. Proposed Approach

By assuming the noise $\mathbf{n}$ is independent of the transmitted symbols $\mathbf{x}$, the identification of the channel $\mathcal{H}$ can be realized by first applying the EVD on the autocorrelation matrix $\mathbf{R_y} = E\left[\mathbf{y}\mathbf{y}^H\right]$, then optimizing a quadratic cost function constructed from the eigenvectors of the noise subspace. An estimate of the correlation matrix can be obtained through time averaging by

$$\hat{\mathbf{R}}_{\mathbf{y}} = \frac{1}{T_{av}} \sum_{j=1}^{T_{av}} \mathbf{y}_{(j)} \mathbf{y}_{(j)}^H, \tag{4}$$

where $\mathbf{y}_{(j)} \in \mathbb{C}^{N_R N_C N_F}$ denotes the jth epoch of the received signal $\mathbf{y}$, consisting of N_F OFDM symbols, and T_{av} is the number of OFDM blocks (or simply the number of time samples). Although a pilot-based (not blind) subspace method in the frequency domain was proposed in [15][16], to our knowledge, a subspace-based blind channel estimation constructed from (2) has never been considered; since there are $N_T N_R N_C \geq N_T N_R L$ unknown channel coefficients to be estimated, where L denotes the channel order. Nevertheless, the number of unknowns can be reduced by exploiting the frequency correlation among adjacent OFDM subcarriers with some loss in the estimation performance; in return, the order of the correlation matrix and hence the number of time samples required for time averaging can be reduced significantly. The details are given below:

Let the frequency span of P adjacent subcarriers reside inside the coherence bandwidth of the wireless channel and let $\mathcal{S} := \{0, 1, \cdots, N_C - 1\}$ be the index set of the N_C subcarriers. $\mathcal{S}$ is partitioned into P disjoint subsets (assuming $(N_C/P) \in \mathbb{Z}^+$) with each subset denoted as $\mathcal{S}_p := \{s_{p,1}, s_{p,2}, \cdots, s_{p,(N_C/P)}\}$, where $s_{p,i} := p - 1 + (i-1)P$, $i = 1, 2, \cdots, (N_C/P)$ for $p = 1, 2, \cdots, P$. Note that $\mathcal{S}_1 \cup \mathcal{S}_2 \cup \cdots \cup \mathcal{S}_P = \mathcal{S}$, and $\mathcal{S}_i \cap \mathcal{S}_j = \emptyset$ for $i \neq j$, where $\emptyset$ denotes the empty set. Define $\mathbf{x}_p = [\mathbf{x}_p^{1^T} \mathbf{x}_p^{2^T} \cdots \mathbf{x}_p^{N_F{}^T}]^T$, $\mathbf{y}_p = [\mathbf{y}_p^{1^T} \mathbf{y}_p^{2^T} \cdots \mathbf{y}_p^{N_F{}^T}]^T$, $\mathbf{n}_p = [\mathbf{n}_p^{1^T} \mathbf{n}_p^{2^T} \cdots \mathbf{n}_p^{N_F{}^T}]^T$, where

$$\begin{aligned} \mathbf{x}_p^m &:= \left[\, \mathbf{x}^m[s_{p,1}]^T \; \mathbf{x}^m[s_{p,2}]^T \; \cdots \; \mathbf{x}^m[s_{p,(N_C/P)}]^T \,\right]^T, \\ \mathbf{y}_p^m &:= \left[\, \mathbf{y}^m[s_{p,1}]^T \; \mathbf{y}^m[s_{p,2}]^T \; \cdots \; \mathbf{y}^m[s_{p,(N_C/P)}]^T \,\right]^T, \\ \mathbf{n}_p^m &:= \left[\, \mathbf{n}^m[s_{p,1}]^T \; \mathbf{n}^m[s_{p,2}]^T \; \cdots \; \mathbf{n}^m[s_{p,(N_C/P)}]^T \,\right]^T. \end{aligned}$$

Then (2) can be re-written for the pth subset as

$$\mathbf{y}_p = \sqrt{\frac{E_s}{N_T}}\,\mathcal{H}_p \cdot \mathbf{x}_p + \mathbf{n}_p, \quad p = 1, 2, \cdots, P, \tag{5}$$

where $\mathcal{H}_p := \mathbf{I}_{N_F} \otimes \left(\bigoplus \sum_{k\in\mathcal{S}_p} \mathcal{H}[k]\right)$ is assumed to be a "tall" matrix by choosing $N_R > N_T$. The identification of $\mathcal{H}_p$ is then based on the autocorrelation matrix $\mathbf{R}_{\mathbf{y}_p} = E\left[\mathbf{y}_p\mathbf{y}_p^H\right]$, and can be written as

$$\mathbf{R}_{\mathbf{y}_p} = \mathcal{H}_p\mathbf{R}_{\mathbf{x}_p}\mathcal{H}_p^H + \mathbf{R}_{\mathbf{n}_p}, \tag{6}$$

where $\mathbf{R}_{\mathbf{x}_p} = (E_s/N_T) \cdot E\left[\mathbf{x}_p\mathbf{x}_p^H\right]$ is assumed to be of full rank, and $\mathbf{R}_{\mathbf{n}_p} = E\left[\mathbf{n}_p\mathbf{n}_p^H\right] = \sigma_n^2\,\mathbf{I}$. Since the channel coefficients of adjacent P subcarriers are strongly correlated, the wireless channel can be approximated by denoting $\bar{\mathcal{H}} = \mathcal{H}_1 = \mathcal{H}_2 = \cdots = \mathcal{H}_P$, and hence an estimate of the correlation matrix can be obtained by

$$\hat{\mathbf{R}}_{\mathbf{y}_p} = \frac{1}{PT_{av}}\sum_{j=1}^{T_{av}}\sum_{p=1}^{P}\mathbf{y}_{p\,(j)}\mathbf{y}_{p\,(j)}^H, \tag{7}$$

where $\mathbf{y}_{p\,(j)} \in \mathbb{C}^{N_RN_CN_F/P}$ denotes the jth epoch of the received signal $\mathbf{y}_p$. Therefore, the number of the time samples required can be significantly reduced since the order of the correlation matrix is reduced by a factor of P, and an averaging over P subsets, which is equivalent to the frequency averaging, is applied at each time averaging interval. The parameter P may be chosen to further reduce the time averaging interval with more loss in the estimation performance.

By applying the EVD on $\hat{\mathbf{R}}_{\mathbf{y}_p}$, (7) can be expressed by: $\hat{\mathbf{R}}_{\mathbf{y}_p} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$, where $\mathbf{U}$ is a matrix whose columns are the orthonormal eigenvectors of $\hat{\mathbf{R}}_{\mathbf{y}_p}$, and can be partitioned as

$$\mathbf{U} = [\,\mathbf{U}_s\,|\,\mathbf{U}_n\,] = [\,\mathbf{u}_1\cdots\mathbf{u}_{d_s}\,|\,\mathbf{u}_{d_s+1}\cdots\mathbf{u}_{d_s+d_n}\,]\,. \tag{8}$$

The signal subspace can be denoted as $\mathrm{span}\,(\mathbf{U}_s)$, while its orthogonal complement, the noise subspace, can be denoted as $\mathrm{span}\,(\mathbf{U}_n)$, with $d_s = \mathrm{rank}\,(\bar{\mathcal{H}}) = N_TN_CN_F/P$ and $d_n = (N_R - N_T)N_CN_F/P$. $\mathbf{\Lambda}$ is a diagonal matrix consisting of the corresponding eigenvalues of $\mathbf{U}$, and is denoted as $\mathbf{\Lambda} = \mathrm{diag}\,(\lambda_1, \lambda_2, \cdots, \lambda_{d_s+d_n})$ with $\lambda_{\max} = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{d_s+d_n} = \lambda_{\min} \geq 0$. Since $\bar{\mathcal{H}}$ and $\mathbf{U}_s$ share the same range space and are orthogonal to the range space of $\mathbf{U}_n$, we can have the following orthogonality relationship

$$\mathbf{u}_j^H\bar{\mathcal{H}} = \mathbf{0}, \quad j = d_s + 1, \cdots, d_s + d_n. \tag{9}$$

In order for $\bar{\mathcal{H}}$ to be identifiable, d_s is chosen so that $d_s \geq N_R$, and the matrix $\bar{\mathcal{H}}$ needs to be of full column rank, or $\mathrm{rank}(\bar{\mathcal{H}}) = N_TN_CN_F/P$.

Although $\bar{\mathcal{H}}$ can be solved from the set of homogeneous linear equations in (9); due to the limited time averaging interval, only an estimate of the noise subspace $\hat{\mathbf{U}}_n$ is available in practice. In this case, we can obtain the channel estimate $\hat{\bar{\mathcal{H}}}$ by minimizing a quadratic cost function given by

$$C(\bar{\mathcal{H}}) = \sum_{j=d_s+1}^{d_s+d_n}\|\hat{\mathbf{u}}_j^H\bar{\mathcal{H}}\|_2^2. \tag{10}$$

By partitioning $\hat{\mathbf{u}}_j$ into N_F equal segments as $\hat{\mathbf{u}}_j = [\hat{\mathbf{u}}_{j,1}\hat{\mathbf{u}}_{j,2}\cdots\hat{\mathbf{u}}_{j,N_F}]^T$, we can define a new matrix $\hat{\mathbf{V}}_j := [\hat{\mathbf{u}}_{j,1}^T\hat{\mathbf{u}}_{j,2}^T\cdots\hat{\mathbf{u}}_{j,N_F}^T]$, where $\hat{\mathbf{u}}_{j,i} \in \mathbb{C}^{1\times N_RN_C/P}$ for $i = 1, 2, \cdots, N_F$. In addition, let us define $\bar{\mathcal{H}}' = [\mathbf{h}_1^{(p)}\;\mathbf{h}_2^{(p)}\;\cdots\;\mathbf{h}_{N_T}^{(p)}]$, where $\mathbf{h}_q^{(p)}$ is given as

$$\begin{aligned}\mathbf{h}_q^{(p)} = [&h_{1,q}[s_{p,1}]\;h_{2,q}[s_{p,1}]\;\cdots\;h_{N_R,q}[s_{p,1}]\;\cdots\\ &h_{1,q}[s_{p,(N_C/P)}]\;h_{2,q}[s_{p,(N_C/P)}]\;\cdots\;h_{N_R,q}[s_{p,(N_C/P)}]]^T\end{aligned}$$

for $q = 1, 2, \cdots, N_T$. Then minimizing the quadratic cost function in (10) is equivalent to minimizing

$$C(\bar{\mathcal{H}}') = \sum_{j=d_s+1}^{d_s+d_n}\|\bar{\mathcal{H}}'^T\hat{\mathbf{V}}_j^*\|_F^2 = \sum_{j=d_s+1}^{d_s+d_n}\mathrm{tr}\left(\bar{\mathcal{H}}'^T\hat{\mathbf{V}}_j^*\hat{\mathbf{V}}_j^T\bar{\mathcal{H}}'^*\right). \tag{11}$$

Furthermore, let $\mathbf{\Psi} := \sum_{j=d_s+1}^{d_s+d_n}\hat{\mathbf{V}}_j^*\hat{\mathbf{V}}_j^T \in \mathbb{C}^{(N_RN_C/P)\times(N_RN_C/P)}$, and the eigenvalues of $\mathbf{\Psi}$ be ordered as $\gamma_{\min} = \gamma_1 \leq \gamma_2 \leq \cdots \leq \gamma_{(N_RN_C/P)} = \gamma_{\max}$. Then from the Rayleigh-Ritz theory [14], we can have

$$\gamma_{\min} = \gamma_1 = \min_{\mathbf{w}\neq\mathbf{0}}\frac{\mathbf{w}^H\mathbf{\Psi}\mathbf{w}}{\mathbf{w}^H\mathbf{w}} = \min_{\mathbf{w}^H\mathbf{w}=1}\mathbf{w}^H\mathbf{\Psi}\mathbf{w} \tag{12}$$

for any $\mathbf{w} \in \mathbb{C}^{(N_RN_C/P)\times 1}$. If r is a given integer with $1 \leq r \leq N_RN_C/P$, then

$$\gamma_1(\mathbf{\Psi}) + \cdots + \gamma_r(\mathbf{\Psi}) = \min_{\mathbf{Q}^H\mathbf{Q}=\mathbf{I}}\mathrm{tr}\,(\mathbf{Q}^H\mathbf{\Psi}\mathbf{Q}), \tag{13}$$

where $\mathbf{Q} \in \mathbb{C}^{(N_RN_C/P)\times r}$ is a matrix whose columns are the orthonormal eigenvectors corresponding to the r smallest eigenvalues of $\mathbf{\Psi}$.

In our case, $r = \mathrm{rank}\,(\bigoplus\sum_{k\in\mathcal{S}_p}\mathbf{H}[k]) = N_TN_C/P$ and minimizing of (11) can be re-written by

$$\min C(\bar{\mathcal{H}}') = \min\sum_{q=1}^{N_T}\mathrm{tr}\left((\mathbf{h}_q^{(p)})^T\mathbf{\Psi}(\mathbf{h}_q^{(p)})^*\right), \tag{14}$$

and hence the channel estimate $\hat{\bar{\mathcal{H}}}'$ can be obtained from (13) as

$$\hat{\bar{\mathcal{H}}}' = \mathbf{Q}^*\mathbf{A}, \tag{15}$$

where $\mathbf{A} \in \mathbb{C}^{(N_TN_C/P)\times N_T}$ can be seen as an ambiguity matrix. Note that the ambiguity matrix $\mathbf{A}$ can be solved from (6) by employing the following optimization

$$\begin{aligned}\hat{\mathbf{A}} = \arg\min_{\mathbf{A}}\Big\|\mathrm{vec}\Big(&\mathcal{T}^{-1}(\mathbf{Q}^*\mathbf{A})\,\hat{\mathbf{R}}_{\mathbf{x}_p}\,(\mathcal{T}^{-1}(\mathbf{Q}^*\mathbf{A}))^H\\ &- (\hat{\mathbf{R}}_{\mathbf{y}_p} - \hat{\mathbf{R}}_{\mathbf{n}_p})\Big)\Big\|_2^2,\end{aligned} \tag{16}$$

assuming $\hat{\mathbf{R}}_{\mathbf{x}_p}$ and $\hat{\mathbf{R}}_{\mathbf{n}_p}$ are obtained from additional estimations [2]. $\mathcal{T}^{-1}$ is the inverse of the matrix transformation $\mathcal{T}$, which is defined by $\mathcal{T} : \bar{\mathcal{H}} \rightarrow \bar{\mathcal{H}}'$.

IV. Perturbation Analysis

By denoting $\mathcal{H}_p = \bar{\mathcal{H}} + \Delta\mathcal{H}_p$, the correlation matrix $\mathbf{R}_{\mathbf{y}_p}$ in (6) can be re-written as

$$\mathbf{R}_{\mathbf{y}_p} = \bar{\mathcal{H}}\mathbf{R}_{\mathbf{x}_p}\bar{\mathcal{H}}^H + \Delta\mathcal{H}_p\mathbf{R}_{\mathbf{x}_p}\bar{\mathcal{H}}^H + \bar{\mathcal{H}}\mathbf{R}_{\mathbf{x}_p}\Delta\mathcal{H}_p^H + \Delta\mathcal{H}_p\mathbf{R}_{\mathbf{x}_p}\Delta\mathcal{H}_p^H + \mathbf{R}_{\mathbf{n}_p}. \quad (17)$$

Let $\mathbf{R}_{\Delta\mathcal{H}_p} := \Delta\mathcal{H}_p\mathbf{R}_{\mathbf{x}_p}\bar{\mathcal{H}}^H + \bar{\mathcal{H}}\mathbf{R}_{\mathbf{x}_p}\Delta\mathcal{H}_p^H + \Delta\mathcal{H}_p\mathbf{R}_{\mathbf{x}_p}\Delta\mathcal{H}_p^H$, then the estimate of the correlation matrix in (7) converges to:

$$\mathbf{R}_{\mathbf{y}} = \underbrace{\bar{\mathcal{H}}\mathbf{R}_{\mathbf{x}_p}\bar{\mathcal{H}}^H}_{:=\mathbf{R}_{\bar{\mathcal{H}}}} + \mathbf{R}_{\Delta\mathcal{H}} + \mathbf{R}_{\mathbf{n}} \text{ as } T_{av} \to \infty, \quad (18)$$

where $\mathbf{R}_{\Delta\mathcal{H}} := (1/P)\sum_{p=1}^{P}\mathbf{R}_{\Delta\mathcal{H}_p}$ and $\mathbf{R}_{\mathbf{n}} := (1/P)\sum_{p=1}^{P}\mathbf{R}_{\mathbf{n}_p}$. In section III, an identification of $\bar{\mathcal{H}}$ was obtained by assuming $\|\Delta\mathcal{H}_p\|_F \to 0$, $\forall p$. In the following, the asymptotic root mean square error (ARMSE) of the proposed estimator at high signal-to-noise ratio (SNR) and long time averaging period is studied:

Since the correlation matrix $\mathbf{R}_{\mathbf{y}} = \mathbf{R}_{\bar{\mathcal{H}}} + \mathbf{R}_{\Delta\mathcal{H}} + \mathbf{R}_{\mathbf{n}}$ can be seen as a perturbed data matrix with $\mathbf{R}_{\mathbf{n}}^{'} := \mathbf{R}_{\Delta\mathcal{H}} + \mathbf{R}_{\mathbf{n}}$ being the perturbation source, we have the first order perturbation of the noise subspace [17][18] denoted by

$$\Delta\mathbf{U}_{n,\bar{\mathcal{H}}} = -\mathbf{U}_{s,\bar{\mathcal{H}}}\Sigma_{s,\bar{\mathcal{H}}}^{-1}\mathbf{V}_{s,\bar{\mathcal{H}}}^H(\mathbf{R}_{\mathbf{n}}^{'})^H\mathbf{U}_{n,\bar{\mathcal{H}}}, \quad (19)$$

assuming the Singular Value Decomposition (SVD) of $\mathbf{R}_{\bar{\mathcal{H}}}$ is written as

$$\mathbf{R}_{\bar{\mathcal{H}}} = [\,\mathbf{U}_{s,\bar{\mathcal{H}}} \mid \mathbf{U}_{n,\bar{\mathcal{H}}}\,]\begin{bmatrix}\mathbf{\Sigma}_{\mathbf{s},\bar{\mathcal{H}}} & \\ & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{V}_{s,\bar{\mathcal{H}}}^H \\ \mathbf{V}_{n,\bar{\mathcal{H}}}^H\end{bmatrix}. \quad (20)$$

Furthermore, by partitioning the jth column of $\mathbf{U}_{n,\bar{\mathcal{H}}}$ and $\Delta\mathbf{U}_{n,\bar{\mathcal{H}}}$ into N_F equal segments, we can define new matrices $\mathbf{V}_j := [\mathbf{U}_{n,\bar{\mathcal{H}}}^{j,1}\mathbf{U}_{n,\bar{\mathcal{H}}}^{j,2}\cdots\mathbf{U}_{n,\bar{\mathcal{H}}}^{j,N_F}]$ and $\Delta\mathbf{V}_j := [\Delta\mathbf{U}_{n,\bar{\mathcal{H}}}^{j,1}\Delta\mathbf{U}_{n,\bar{\mathcal{H}}}^{j,2}\cdots\Delta\mathbf{U}_{n,\bar{\mathcal{H}}}^{j,N_F}]$, where $\mathbf{U}_{n,\bar{\mathcal{H}}}^{j,i}$ and $\Delta\mathbf{U}_{n,\bar{\mathcal{H}}}^{j,i}$ denote the ith segment of the jth column of $\mathbf{U}_{n,\bar{\mathcal{H}}}$ and $\Delta\mathbf{U}_{n,\bar{\mathcal{H}}}$, respectively. Similar to (11), $\mathbf{\Psi} := \sum_{j=1}^{d_n}\mathbf{V}_j^*\mathbf{V}_j^T$ is then constructed for an estimate of the Hermitian matrix $\mathbf{Q}$. Let the SVD of $\mathbf{\Psi}$ be written as

$$\mathbf{\Psi} = [\,\mathbf{U}_{s,\Psi} \mid \mathbf{U}_{n,\Psi}\,]\begin{bmatrix}\mathbf{\Sigma}_{\mathbf{s},\mathbf{\Psi}} & \\ & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{V}_{s,\Psi}^H \\ \mathbf{V}_{n,\Psi}^H\end{bmatrix}, \quad (21)$$

then we immediately have $\mathbf{Q} = \mathbf{U}_{n,\Psi}$ [14] and hence

$$\Delta\mathbf{Q} = -\mathbf{U}_{s,\Psi}\Sigma_{s,\Psi}^{-1}\mathbf{V}_{s,\Psi}^H(\Delta\mathbf{\Psi})^H\mathbf{Q}, \quad (22)$$

where $\Delta\mathbf{\Psi} := \sum_{j=1}^{d_n}\mathbf{V}_j^*\Delta\mathbf{V}_j^T + \Delta\mathbf{V}_j^*\mathbf{V}_j^T + \Delta\mathbf{V}_j^*\Delta\mathbf{V}_j^T$. In addition, by assuming that $\bar{\mathcal{H}}_p^{'} = [\mathbf{h}_1^{(p)}\ \mathbf{h}_2^{(p)}\ \cdots\ \mathbf{h}_{N_T}^{(p)}]$ is known, the ambiguity matrix $\mathbf{A}$ can be denoted by $\mathbf{A} = (\mathbf{Q}^* + \Delta\mathbf{Q}^*)^\dagger\bar{\mathcal{H}}_p^{'}$. Therefore, the channel estimate and its error can be written as

$$\hat{\bar{\mathcal{H}}}^{'} = (\mathbf{Q}^* + \Delta\mathbf{Q}^*)(\mathbf{Q}^* + \Delta\mathbf{Q}^*)^\dagger\bar{\mathcal{H}}_p^{'}, \quad (23)$$

and

$$\Delta\bar{\mathcal{H}}^{'} = \left[\mathbf{I} - (\mathbf{Q}^* + \Delta\mathbf{Q}^*)(\mathbf{Q}^* + \Delta\mathbf{Q}^*)^\dagger\right]\bar{\mathcal{H}}_p^{'}. \quad (24)$$

Note that since $(\mathbf{Q}^* + \Delta\mathbf{Q}^*)(\mathbf{Q}^* + \Delta\mathbf{Q}^*)^\dagger = \mathbf{P}_{\Re(\mathbf{Q}^*+\Delta\mathbf{Q}^*)}$ and $\mathbf{I} - (\mathbf{Q}^* + \Delta\mathbf{Q}^*)(\mathbf{Q}^* + \Delta\mathbf{Q}^*)^\dagger = \mathbf{P}_{\Re(\mathbf{Q}^*+\Delta\mathbf{Q}^*)^\perp}$, each column of the estimated channel matrix $\hat{\bar{\mathcal{H}}}^{'}$ can be seen as the projection of the corresponding column of the channel matrix $\bar{\mathcal{H}}^{'}$ on the range of $(\mathbf{Q}^* + \Delta\mathbf{Q}^*)$ [19][20].

The root mean square error (RMSE) of the estimated channel coefficients, concerning all the P subsets, can then be written as

$$\mathbf{RMSE} = \sqrt{\frac{1}{N_T N_R N_C}E\left\{\sum_{p'=1}^{P}\left\|\bar{\mathcal{H}}_{p'}^{'} - \hat{\bar{\mathcal{H}}}^{'}\right\|_F^2\right\}}, \quad (25)$$

and the channel average bias (CAB) of the estimated channel coefficients, concerning all the P subsets, can then be written as

$$\mathbf{CAB} = \frac{1}{N_T N_R N_C}E\left\{\sum_{p'=1}^{P}\left\|\mathrm{vec}\left(\bar{\mathcal{H}}_{p'}^{'} - \hat{\bar{\mathcal{H}}}^{'}\right)\right\|_1\right\}. \quad (26)$$

V. Numerical Results

Numerical results of the proposed as well as the referenced subspace-based methods, including CP and VC approaches for MIMO-OFDM systems are presented in this section. We consider MIMO-OFDM systems with 2 transmit ($N_T = 2$) and 3 receive antennas ($N_R = 3$). The number of subcarriers used in the OFDM systems is 256 (N_C = 256). For each time epoch, the incoming symbol streams, which are independent and identically distributed (i.i.d) QPSK symbols, span over 2 OFDM symbols (N_F = 2). Note that part of the results are drawn from [11].

In order to include the subspace-based methods [5][7] for comparisons, we consider a 2-tap delay-line scenario where both the real and imaginary parts of the tap coefficients are i.i.d. Gaussian random variables with zero mean and unit variance. Under these circumstances, there are 10 subcarriers residing inside the coherence bandwidth (P = 10) if the coherence bandwidth is defined as the bandwidth over which the frequency correlation function is above 0.9, while there are 100 subcarriers residing inside the coherence bandwidth (P = 100) if the definition is relaxed so that the frequency correlation function is above 0.5. For the purpose of evaluating the estimation performance, the ambiguity matrices for all the methods are resolved by assuming the channel responses are known.

Fig. 2 and 3 show the RMSE measure of the proposed and referenced subspace-based methods, which is a function of the number of the OFDM blocks (each OFDM block is consitituted of 2 OFDM symbols) employed for obtaining a sampled correlation matrix when the SNR = 20dB. As expected, the estimation performance of all the methods is improved when the number of the OFDM blocks is increased for time averaging. For the referenced methods, we consider a fixed degree of freedom equals to 8 and 16, respectively. For the proposed methods, we consider P = 2, 8, 32, and 64. We note that the proposed methods outperform the referenced methods with any degrees of freedom, within the given time averaging period. From the comparison of the orders of correlation matrices listed in Table I, we also find that the minimum number of the time samples required (i.e., the order of the correlation matrix) is reduced when P is large, while the estimation performance is also reduced due to the channel approximations over P adjacent subcarriers. On the contrary, the number of the time samples required is increased when P is small, while the estimation performace is also improved. To achieve the best tradeoff, we note that P should be chosen to have the frequency spans of these P subcarriers residing around the coherence bandwidth, with the definition of the correlation function above 0.9.

By letting T_{av} = 3000 and SNR = 40dB in (18), we calculate the asymptotic performance bounds of the proposed scheme for different values of P. When SNR = 20dB, the required OFDM blocks for P = 8 and 2 are T_{av} = 50 and $T_{av} > 210$ (not shown in the figure), respectively, in order to reach their asymptotic performance. Since the accuracy of the asymptotic performance bounds relies on $\|\Delta\mathcal{H}_p\|_F \rightarrow 0$, $\forall p$, bounds of P = 32 and 64 can only serve as an indication showing that whether the proposed scheme has met a certain level of confidence for a given time averaging interval or SNR. Therefore, we can conclude that when SNR = 20dB, the required OFDM blocks for P = 32 and 64 are T_{av} = 10 in order to reach their asymptotic performance. Note that because frequency averaging is also included in (7), the required time samples of the proposed scheme can be less than the order of the corresponding correlation matrix.

Fig. 4 and 5 show the corresponding CAB measure, and we can reach the same conlusions from similar observations. It should be noted that the error floors of the estimation performance come from the fact that $\|\Delta\mathcal{H}_p\|_F \neq 0$, $\forall p$. However, the effects can be alleviated by increasing the value of N_C (i.e., the number of FFT/IFFT size) or simply letting P = 1.

VI. Conclusion

Traditional subspace-based blind methods suffer from an extremely slow convergence rate, making them impractical [11]. Therefore, we proposed a novel subspace-based estimation method with a faster convergence rate, mainly by exploiting the frequency correlation among adjacent subcarriers through the concept of subcarrier grouping. In this paper, we investigated the asymptotic performance bounds of the proposed scheme by perturbation analysis. Within reasonable time averaging periods, the numerical results were shown to support the proposed method by achieving both a higher convergence rate and a better estimation accuracy.

TABLE I
Comparison of the orders of correlation matrices

Blind Estimator	Order of the Correlation Matrix
cyclic prefix (CP=8)	1581
cyclic prefix (CP=16)	1629
virtual carrier (CP=4, Null=4)	1557
virtual carrier (CP=8, Null=8)	1581
proposed scheme (P=64)	24
proposed scheme (P=32)	48
proposed scheme (P=8)	192
proposed scheme (P=2)	768

References

[1] G. L. Stuber, J. R. Barry, S. W. Mclaughlin, Y. Li, M. A. Ingram, and T. G. Pratt, "Broadband MIMO-OFDM wireless communications," *Proceesings of the IEEE*, vol. 92, no. 2, pp. 271–294, Feb. 2004.

[2] L. Tong and S. Perreau, "Multichannel blind identification: from subspace to maximum likelihood methods," *IEEE Proc.*, vol. 86, no. 10, pp. 1951–1968, Oct. 1998.

[3] E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Processing*, vol. 43, no. 2, pp. 516–525, Feb. 1995.

[4] X. Cai and A. N. Akansu, "A subspace method for blind channel identification in OFDM systems," in *Proc. IEEE ICC*, vol. 2, no. 18-22, June 2000, pp. 929–933.

[5] W. Bai, C. He, L.-G. Jiang, and H.-W. Zhu, "Blind channel estimation in MIMO-OFDM systems," in *Proc. IEEE Globecom*, vol. 1, no. 17-21, Nov. 2002, pp. 317–321.

[6] C. Li and S. Roy, "Subspace-based blind channel estimation for OFDM by exploiting virtual carriers," *IEEE Trans. Wireless Commun.*, vol. 2, no. 1, pp. 141–150, Jan. 2003.

[7] C. Shin and E. J. Powers, "Blind channel estimation for MIMO-OFDM systems using virtual carriers," in *Proc. IEEE Globecom*, vol. 4, no. 29, Nov. 2004, pp. 2465–2469.

[8] J. W. Silverstein and P. L. Combettes, "Large dimensional random matrix theory for signal detection and estimation in array processing," in *Proc. IEEE Statistical Signal and Array Processing*, vol. 7-9, Oct. 1992, pp. 276–279.

[9] X. Mestre and M. A. Lagunas, "Finite sample size effect on minimum variance beamformers: Optimum diagonal loading factor for large arrays," *IEEE Trans. Signal Processing*, vol. 54, no. 1, pp. 69–82, Jan. 2006.

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

[11] C.-C. Tu and B. Champagne, "Subspace-based blind channel estimation for MIMO-OFDM systems: Reducing the time averaging interval of the correlation matrix," in *Proc. IEEE PIMRC*, Sept. 2007.

[12] Z. Liu, Y. Xin, and G. B. Giannakis, "Space-time-frequency coded OFDM over frequency-selective fading channels," *IEEE Trans. Signal Processing*, vol. 50, no. 10, pp. 2465–2476, Oct. 2002.

[13] A. F. Molisch, M. Z. Win, and J. H. Winters, "Space-time-frequency (STF) coding for MIMO-OFDM systems," *IEEE Commun. Lett.*, vol. 6, no. 9, pp. 370–372, Sept. 2002.

[14] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.

[15] O. Simeone, Y. Bar-Ness, and U. Spagnolini, "Pilot-based channel estimation for OFDM systems by tracking the delay subspace," *IEEE Trans. Wireless Commun.*, vol. 3, no. 1, pp. 315–325, Jan. 2004.

[16] M. Cicerone, O. Simeone, and U. Spagnolini, "Channel estimation for MIMO-OFDM systems by using modal analysis/filtering," *IEEE Trans. Commun.*, vol. 54, no. 11, pp. 2062–2074, Nov. 2006.

[17] 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, Oct. 1993.

[18] S. Roy and C. Li, "A subspace blind channel estimation method for OFDM systems without cyclic prefix," *IEEE Trans. Wireless Commun.*, vol. 1, no. 4, pp. 572–579, Oct. 2002.

[19] A. Albert, *Regression and the Moore-Penrose Pseudoinverse*. Academic Press, Inc., 1972.

[20] A. J. Laub, *Matrix Analysis for Scientists and Engineers*. Society for Industrial and Applied Mathematics, 2005.

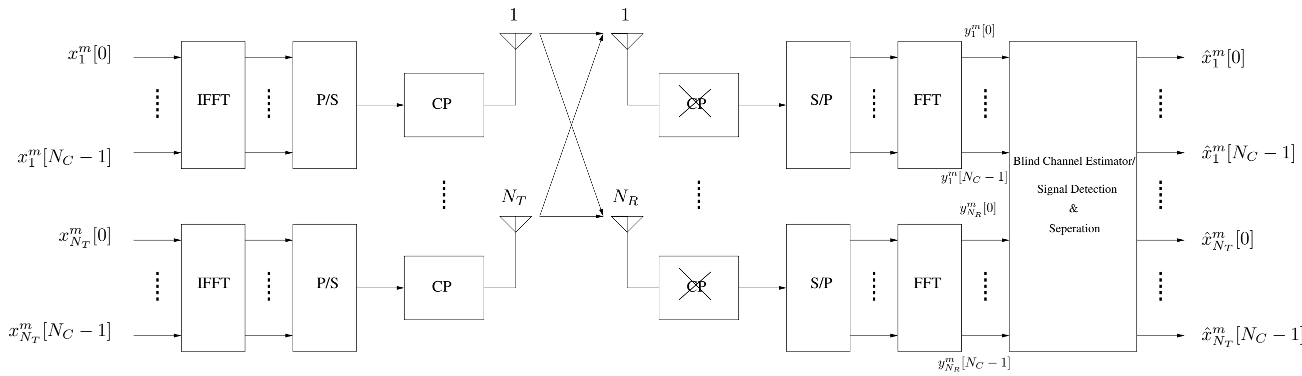

Fig. 1. Proposed subspace-based blind channel estimator for MIMO-OFDM systems.

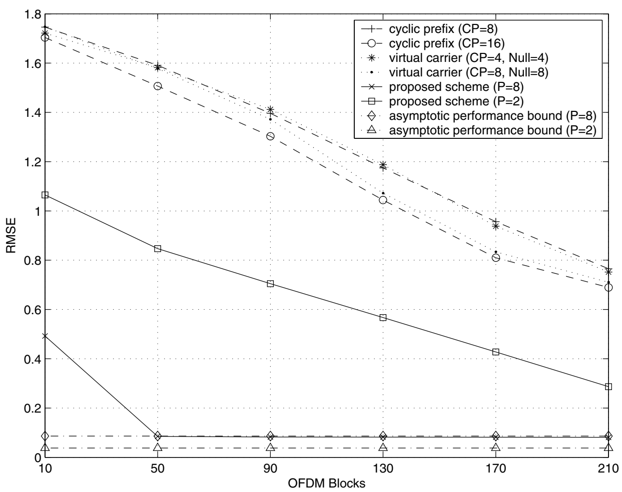

Fig. 2. RMSE versus number of OFDM blocks (SNR=20dB).

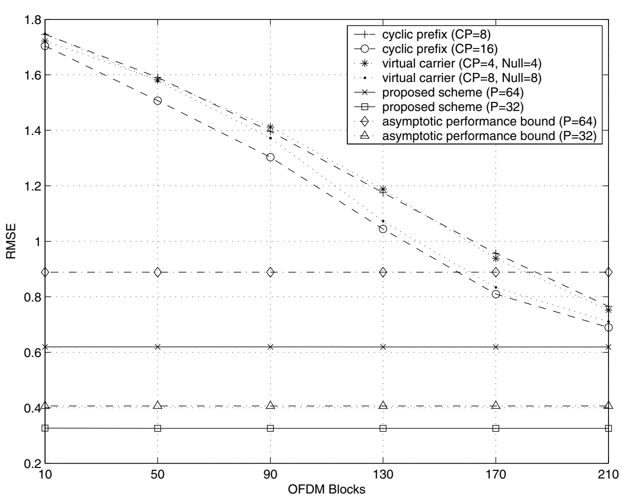

Fig. 3. RMSE versus number of OFDM blocks (SNR=20dB).

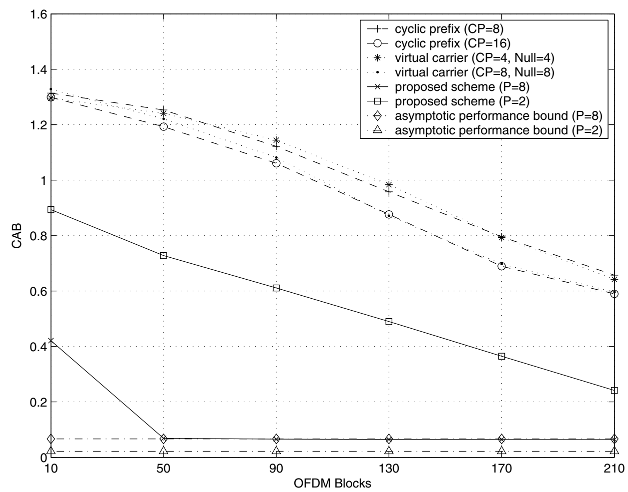

Fig. 4. CAB versus number of OFDM blocks (SNR=20dB).

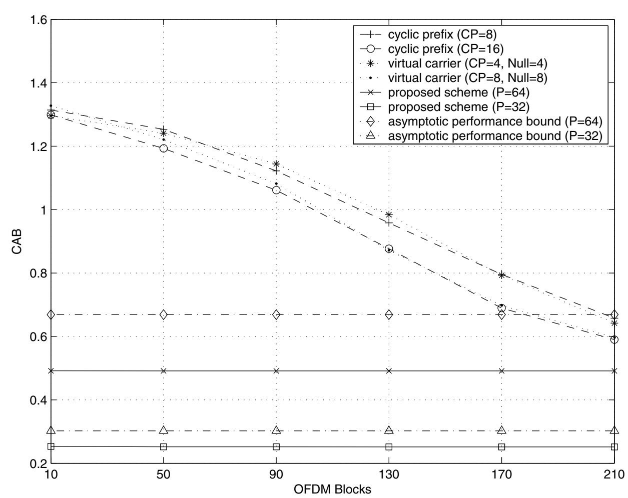

Fig. 5. CAB versus number of OFDM blocks (SNR=20dB).