

RESEARCH

Open Access

Semiblind channel estimation for MIMO–OFDM systems

Yi-Sheng Chen* and Jyu-Han Song

Abstract

This article proposes a semiblind channel estimation method for multiple-input multiple-output orthogonal frequency-division multiplexing systems based on circular precoding. Relying on the precoding scheme at the transmitters, the autocorrelation matrix of the received data induces a structure relating the outer product of the channel frequency response matrix and precoding coefficients. This structure makes it possible to extract information about channel product matrices, which can be used to form a Hermitian matrix whose positive eigenvalues and corresponding eigenvectors yield the channel impulse response matrix. This article also tests the resistance of the precoding design to finite-sample estimation errors, and explores the effects of the precoding scheme on channel equalization by performing pairwise error probability analysis. The proposed method is immune to channel zero locations, and is reasonably robust to channel order overestimation. The proposed method is applicable to the scenarios in which the number of transmitters exceeds that of the receivers. Simulation results demonstrate the performance of the proposed method and compare it with some existing methods.

Keywords: MIMO, OFDM, Semiblind estimation, Frequency domain equalization, Circular precoding

1 Introduction

Orthogonal frequency-division multiplexing (OFDM), when combined with cyclic prefix (CP) as the guard intervals, is an effective transmission technique for high-speed broadband communication systems because of its high data rate, high spectral efficiency, and lack of intersymbol interference (ISI) [1,2]. The operational principle of OFDM is to use inverse discrete Fourier transform (IDFT) and CP insertion to divide the original bandwidth into multiple narrow sub-bands, in which the mobile channel can be considered non-dispersive [3]. It is then easy to implement low complexity equalization at the receiver by using a set of complex multipliers, one for each sub-band, provided the channel state information is available [4].

Multiple-input multiple-output (MIMO) technology, which employs multiple antennas at the transmitters and receivers, has received much attention due to its ability to improve the data transmission rate through enormous

channel capacity gains. Hence, an MIMO–OFDM system that combines OFDM and MIMO technologies is a key way for achieving high performance transmission in modern wireless communications [5].

The receivers of MIMO–OFDM systems require channel state information to detect symbols reliably. Blind or semiblind channel estimation is a bandwidth-efficient alternative to the conventional training based approaches [6-8]. Researchers have recently proposed various methods for (semi) blind channel estimation for MIMO–OFDM systems [9-12]. Gao et al. [9] proposed a robust subspace method applicable to MIMO–OFDM systems. Their method exhibits many advantages, including robustness to channel order overestimation and guaranteeing the channel identifiability. However, this method is not suitable for the case of more transmitters than receivers, and it imposes some constraints on channel zero locations. Blind or semiblind estimation using non-redundant precoding [8] can solve these problems since it avoids the catastrophic effects of channel zeros and requires less assumptions on channel. Previous studies present three typically non-redundant precoding methods for (semi)blind channel estimation for MIMO–OFDM systems [10-12]. The method in [10] uses the

*Correspondence: yis Chen@fcu.edu.tw
Department of Communications Engineering, Feng Chia University, Taichung, Taiwan

precoding to spread the symbols of each user over all sub-carriers, thus increasing multipath diversity and reducing bit error rate (BER) at the receivers. Gao and Nallanathan [11] generalized the precoding method in [13] to MIMO-OFDM systems. A distinguishing feature of their method is that it can be applied to the scenarios in which the number of transmitters exceeds that of the receivers. Shin et al. [12] presented a framework for exploiting a general non-redundant precoding method for MIMO-OFDM systems and MIMO single-carrier systems with frequency-domain equalization. Their method is robust to channel order overestimation and incurs a relaxed channel identifiability condition.

This article develops a semiblind channel estimation method for MIMO-OFDM systems based on a specific and non-redundant precoding scheme, say, circular precoding, since the circular precoding allows channel estimation at the receiver and simplifies the encoding scheme at the transmitter [14]. In literature, to the best of our knowledge, only two circular precoding based methods have been proposed for single-input single-output (SISO) OFDM systems [14,15]. Thus the current study focuses on generalizing the methods in the SISO case [14,15] to the MIMO-OFDM systems. The proposed method is based on second-order statistics. With circular precoding at the transmitters, the autocorrelation matrix of the received data is equal to a noise-perturbated matrix involving the outer product of the channel frequency response matrix and the coefficients relating to the precoding. Dividing each submatrix in the autocorrelation matrix by the corresponding coefficient related to the precoding gives a noise-perturbed outer product of the channel frequency response matrix. Then we use the relation of the channel frequency response matrix and the channel impulse response matrix to transform the above noise-perturbed matrix to another noise-perturbed matrix. The resulting noise-perturbed matrix is equal to an outer product of the channel impulse response matrix plus a diagonal matrix due to channel noise. Next, we use a simple method to eliminate the noise components to obtain the outer product of the channel impulse response matrix. Finally, the channel impulse response matrix is obtained by

computing the positive eigenvalues and the corresponding eigenvectors of this outer-product matrix. This study also tests the resistance of the precoding design to finite-sample estimation errors, and explores the effects of the precoding scheme on channel equalization through pairwise error probability (PEP) analysis. Simulation results demonstrate the performance of the proposed method and compare it with previous methods.

This article is organized as follows. Section 2 presents the system model and problem statement. Section 3 derives the estimation method, studies the precoding design, evaluates the equalization performance, and provides some further discussion about the proposed algorithm. Section 4 shows simulation results. Finally, Section 5 concludes this article.

The notations used in this article are quite standard: bold uppercase is used for matrices, and bold lowercase is used for vectors. \mathbf{A}^T represents the transpose of the matrix \mathbf{A} , and \mathbf{A}^* represents the conjugate transpose of the matrix \mathbf{A} . \mathbf{I}_M is the identity matrix of dimension $M \times M$, and $\mathbf{A} \otimes \mathbf{B}$ is the Kronecker product of matrices \mathbf{A} and \mathbf{B} . The symbols \mathbb{R} and \mathbb{C} represent the set of real numbers and the set of complex numbers, respectively.

2 System model and basic assumptions

Consider the K -input J -output discrete time OFDM baseband system shown in Figure 1. Let $\{\mathbf{H}(0) \mathbf{H}(1) \dots \mathbf{H}(L)\}$ be the MIMO channel impulse response from the transmitters to the receivers with order L , where $\mathbf{H}(l) \in \mathbb{C}^{J \times K}$ is the l th lag of the MIMO channel, $\forall l = 0, 1, \dots, L$. The matrix element $[\mathbf{H}(l)]_{jk} = h_{jk}(l)$ is the l th lag of the channel between the k th transmitter and the j th receiver, $\forall k = 1, 2, \dots, K, \forall j = 1, 2, \dots, J$. At the k th transmitter, the n th information vector $\mathbf{x}^{(k)}(n) = [x_1^{(k)}(n) \ x_2^{(k)}(n) \ \dots \ x_M^{(k)}(n)]^T \in \mathbb{C}^M$ is first precoded by a real circular precoder $\mathbf{P} \in \mathbb{R}^{M \times M}$, followed by an IDFT matrix \mathbf{F}^{-1} , to obtain

$$\mathbf{s}^{(k)}(n) = \mathbf{F}^{-1} \mathbf{P} \mathbf{x}^{(k)}(n), \quad (2.1)$$

where $\mathbf{s}^{(k)}(n) = [s_1^{(k)}(n) \ s_2^{(k)}(n) \ \dots \ s_M^{(k)}(n)]^T \in \mathbb{C}^M$, and \mathbf{P} is a circulant matrix with $[p_1 \ p_2 \ \dots \ p_M]^T$ as its

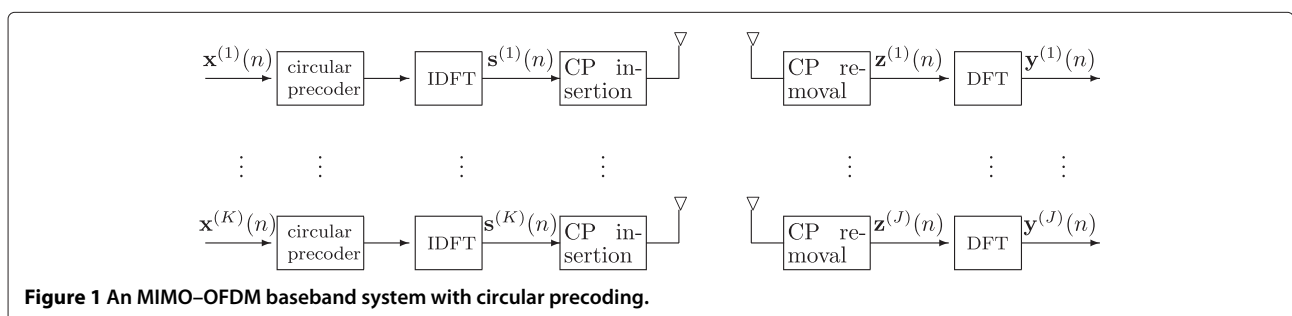


Figure 1 An MIMO-OFDM baseband system with circular precoding.

first column with $p_i > 0, \forall i$. After CP insertion for each transmitted vector $\mathbf{s}^{(k)}(n)$ and CP removal at the receiver, as long as the length of CP is longer than or equal to L , the input-output relation of the system can be described as follows [11]:

$$\begin{cases} \mathbf{z}^{(1)}(n) = \sum_{k=1}^K \mathbf{H}_{1k} \mathbf{s}^{(k)}(n) + \mathbf{w}^{(1)}(n) \\ \mathbf{z}^{(2)}(n) = \sum_{k=1}^K \mathbf{H}_{2k} \mathbf{s}^{(k)}(n) + \mathbf{w}^{(2)}(n) \\ \vdots \\ \mathbf{z}^{(J)}(n) = \sum_{k=1}^K \mathbf{H}_{Jk} \mathbf{s}^{(k)}(n) + \mathbf{w}^{(J)}(n), \end{cases} \quad (2.2)$$

where $\mathbf{H}_{jk} \in \mathbb{C}^{M \times M}$ is a circulant matrix with $[h_{jk}(0) \ h_{jk}(1) \ \dots \ h_{jk}(L) \ 0 \ \dots \ 0]^T \in \mathbb{C}^M$ being its first column, and $\mathbf{z}^{(j)}(n) = [z_1^{(j)}(n) \ z_2^{(j)}(n) \ \dots \ z_M^{(j)}(n)]^T \in \mathbb{C}^M$ and $\mathbf{w}^{(j)}(n) = [w_1^{(j)}(n) \ w_2^{(j)}(n) \ \dots \ w_M^{(j)}(n)]^T \in \mathbb{C}^M$ are the received signal vector and the additive white Gaussian noise (AWGN) vector, respectively, $\forall j = 1, 2, \dots, J$. Taking DFT operation on the received signal $\mathbf{z}^{(j)}(n)$ in (2.2) and using (2.1) lead to the following equations:

$$\begin{cases} \mathbf{y}^{(1)}(n) = \mathbf{Fz}^{(1)}(n) = \sum_{k=1}^K \mathbf{D}_{1k} \mathbf{Px}^{(k)}(n) + \mathbf{v}^{(1)}(n) \\ \mathbf{y}^{(2)}(n) = \mathbf{Fz}^{(2)}(n) = \sum_{k=1}^K \mathbf{D}_{2k} \mathbf{Px}^{(k)}(n) + \mathbf{v}^{(2)}(n) \\ \vdots \\ \mathbf{y}^{(J)}(n) = \mathbf{Fz}^{(J)}(n) = \sum_{k=1}^K \mathbf{D}_{Jk} \mathbf{Px}^{(k)}(n) + \mathbf{v}^{(J)}(n), \end{cases} \quad (2.3)$$

where $\mathbf{v}^{(j)}(n) = \mathbf{Fw}^{(j)}(n) = [v_1^{(j)}(n) \ v_2^{(j)}(n) \ \dots \ v_M^{(j)}(n)]^T \in \mathbb{C}^M$ is the noise vector, $\mathbf{y}^{(j)}(n) = [y_1^{(j)}(n) \ y_2^{(j)}(n) \ \dots \ y_M^{(j)}(n)]^T \in \mathbb{C}^M$ is the received vector, and $\mathbf{D}_{jk} = \mathbf{FH}_{jk}\mathbf{F}^{-1} = \text{diag}[d_{jk}(1) \ d_{jk}(2) \ \dots \ d_{jk}(M)] \in \mathbb{C}^{M \times M}$ is a diagonal matrix, $\forall 1 \leq j \leq J, 1 \leq k \leq K$. Note that $d_{jk}(m) = \sum_{l=0}^L h_{jk}(l) \omega^{-l(m-1)}$ is the channel frequency response between the k th transmitter and the j th receiver at the m th subcarrier for $m = 1, 2, \dots, M$ and $\omega = \exp(i2\pi/M)$.

To further simplify the system model, we regroup the transmitted symbols, received signals, and noise signals on the same time slot as follows:

$$\begin{aligned} \mathbf{x}_i(n) &= [x_i^{(1)}(n) \ x_i^{(2)}(n) \ \dots \ x_i^{(K)}(n)]^T \in \mathbb{C}^K, \\ i &= 1, 2, \dots, M, \\ \mathbf{y}_i(n) &= [y_i^{(1)}(n) \ y_i^{(2)}(n) \ \dots \ y_i^{(J)}(n)]^T \in \mathbb{C}^J, \\ i &= 1, 2, \dots, M, \end{aligned}$$

$$\begin{aligned} \mathbf{v}_i(n) &= [v_i^{(1)}(n) \ v_i^{(2)}(n) \ \dots \ v_i^{(J)}(n)]^T \in \mathbb{C}^J, \\ i &= 1, 2, \dots, M. \end{aligned}$$

Then, after some proper entry permutations, (2.3) can be rewritten as

$$\underbrace{\begin{bmatrix} \mathbf{y}_1(n) \\ \mathbf{y}_2(n) \\ \vdots \\ \mathbf{y}_M(n) \end{bmatrix}}_{\mathbf{y}(n)} = \underbrace{\begin{bmatrix} \mathbf{D}(1) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{D}(2) & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{D}(M) \end{bmatrix}}_{\mathbf{D}} (\mathbf{P} \otimes \mathbf{I}_K) \times \underbrace{\begin{bmatrix} \mathbf{x}_1(n) \\ \mathbf{x}_2(n) \\ \vdots \\ \mathbf{x}_M(n) \end{bmatrix}}_{\mathbf{x}(n)} + \underbrace{\begin{bmatrix} \mathbf{v}_1(n) \\ \mathbf{v}_2(n) \\ \vdots \\ \mathbf{v}_M(n) \end{bmatrix}}_{\mathbf{v}(n)}, \quad (2.4)$$

where $\mathbf{D}(m) = [d_{jk}(m)] \in \mathbb{C}^{J \times K}$ is the channel frequency response matrix from the transmitters to the receivers at the m th subcarrier.

The purpose of this article is to develop a method of semiblindly identifying the MIMO channel impulse response $\{\mathbf{H}(0), \mathbf{H}(1), \dots, \mathbf{H}(L)\}$, using second-order statistics of the received data based on the following assumptions:

- (i) The source signal $\mathbf{x}^{(k)}(n)$ is a zero-mean and white vector sequence with $E[\mathbf{x}^{(k)}(m)\mathbf{x}^{(l)}(n)^*] = \delta(k-l)\delta(m-n)\mathbf{I}_M$, where $\delta(\cdot)$ is the Kronecker delta function. The AWGN is zero-mean with $E[\mathbf{w}^{(k)}(m)\mathbf{w}^{(l)}(n)^*] = \delta(k-l)\delta(m-n)\sigma_w^2\mathbf{I}_M$. In addition, the source signal is uncorrelated with the AWGN $\mathbf{w}^{(j)}(n)$, i.e., $E[\mathbf{x}^{(k)}(m)\mathbf{w}^{(j)}(n)^*] = \mathbf{0}_{M \times M}, \forall m, n, j, k$.
- (ii) The channel impulse response matrix $\mathbf{H} = [\mathbf{H}(0)^T \ \mathbf{H}(1)^T \ \dots \ \mathbf{H}(L)^T]^T \in \mathbb{C}^{J(L+1) \times K}$ is full column rank, i.e., $\text{rank}(\mathbf{H}) = K$.

3 Semiblind channel estimation and equalization

This section develops the proposed method under assumptions (i) and (ii). Section 3.1 first derives the estimation method. Section 3.2 then discusses the precoding design to combat the effect of finite-sample estimation errors. Section 3.3 investigates the equalization performance of the precoding method using PEP analysis. Section 3.4 provides further discussion about the proposed method.

3.1 The estimation method

Under assumption (i), the autocorrelation matrix of $\mathbf{y}(n)$ in (2.4) is shown as follows:

$$\begin{aligned} \mathbf{R} &= E[\mathbf{y}(n)\mathbf{y}(n)^*] \\ &= \mathbf{D}(\mathbf{P} \otimes \mathbf{I}_K)(\mathbf{P}^* \otimes \mathbf{I}_K)\mathbf{D}^* + \sigma_w^2 \mathbf{I}_{JM} \\ &= \mathbf{D}(\mathbf{P}\mathbf{P}^* \otimes \mathbf{I}_K)\mathbf{D}^* + \sigma_w^2 \mathbf{I}_{JM} \\ &= \mathbf{D}(\mathbf{G} \otimes \mathbf{I}_K)\mathbf{D}^* + \sigma_w^2 \mathbf{I}_{JM}. \end{aligned} \quad (3.1)$$

Since \mathbf{P} is a circulant matrix, $\mathbf{G} = \mathbf{P}\mathbf{P}^* \in \mathbb{R}^{M \times M}$ is also a circulant matrix [16] with $\mathbf{g} = [g_1 \ g_2 \ \dots \ g_M]^T$ being its first column. Let $\mathbf{J} \in \mathbb{R}^{M \times M}$ be a circulant matrix with the first column equal to $[0 \ 1 \ 0 \ \dots \ 0]^T \in \mathbb{R}^M$. Thus, \mathbf{G} can be expressed as

$$\mathbf{G} = [\mathbf{g} \ \mathbf{J}\mathbf{g} \ \mathbf{J}^2\mathbf{g} \ \dots \ \mathbf{J}^{M-1}\mathbf{g}]. \quad (3.2)$$

Using (3.2), (3.1) can be expressed as

$$\begin{aligned} \mathbf{R} &= \mathbf{D}([\mathbf{g} \ \mathbf{J}\mathbf{g} \ \mathbf{J}^2\mathbf{g} \ \dots \ \mathbf{J}^{M-1}\mathbf{g}] \otimes \mathbf{I}_K)\mathbf{D}^* + \sigma_w^2 \mathbf{I}_{JM} \\ &= [\mathbf{D}(\mathbf{g} \otimes \mathbf{I}_K) \ \mathbf{D}(\mathbf{J}\mathbf{g} \otimes \mathbf{I}_K) \ \mathbf{D}(\mathbf{J}^2\mathbf{g} \otimes \mathbf{I}_K) \ \dots \ \mathbf{D}(\mathbf{J}^{M-1}\mathbf{g} \\ &\quad \otimes \mathbf{I}_K)]\mathbf{D}^* + \sigma_w^2 \mathbf{I}_{JM} \\ &= \begin{bmatrix} g_1 \mathbf{D}(1) & g_M \mathbf{D}(1) & \dots & g_2 \mathbf{D}(1) \\ g_2 \mathbf{D}(2) & g_1 \mathbf{D}(2) & \dots & g_3 \mathbf{D}(2) \\ \vdots & \vdots & \ddots & \vdots \\ g_M \mathbf{D}(M) & g_{M-1} \mathbf{D}(M) & \dots & g_1 \mathbf{D}(M) \end{bmatrix} \mathbf{D}^* + \sigma_w^2 \mathbf{I}_{JM} \\ &= \begin{bmatrix} g_1 \mathbf{D}(1)\mathbf{D}(1)^* & g_M \mathbf{D}(1)\mathbf{D}(2)^* & \dots & g_2 \mathbf{D}(1)\mathbf{D}(M)^* \\ g_2 \mathbf{D}(2)\mathbf{D}(1)^* & g_1 \mathbf{D}(2)\mathbf{D}(2)^* & \dots & g_3 \mathbf{D}(2)\mathbf{D}(M)^* \\ \vdots & \vdots & \ddots & \vdots \\ g_M \mathbf{D}(M)\mathbf{D}(1)^* & g_{M-1} \mathbf{D}(M)\mathbf{D}(2)^* & \dots & g_1 \mathbf{D}(M)\mathbf{D}(M)^* \end{bmatrix} \\ &\quad + \sigma_w^2 \mathbf{I}_{JM}. \end{aligned} \quad (3.3)$$

Note that the autocorrelation matrix \mathbf{R} in (3.3) is a noise-perturbated matrix involving the coefficients g_1, g_2, \dots, g_M , and the outer-product matrices $\mathbf{D}(m)\mathbf{D}(n)^*$, $\forall m, n = 1, 2, \dots, M$. Dividing each submatrix in \mathbf{R} by the corresponding coefficient yields the matrices $\mathbf{D}(m)\mathbf{D}(m)^* + \frac{\sigma_w^2}{g_1} \mathbf{I}_J$ and $\mathbf{D}(m)\mathbf{D}(n)^*$ for $m, n = 1, 2, \dots, M$, $m \neq n$. Those matrices are then used to form the following matrix

$$\mathbf{Q}_F = \mathbf{D}_F \mathbf{D}_F^* + \frac{\sigma_w^2}{g_1} \mathbf{I}_{JM}, \quad (3.4)$$

where $\mathbf{D}_F = [\mathbf{D}(1)^T \ \mathbf{D}(2)^T \ \dots \ \mathbf{D}(M)^T]^T \in \mathbb{C}^{JM \times K}$ is the channel frequency response matrix. Note that \mathbf{Q}_F is the outer product of \mathbf{D}_F plus a diagonal matrix $\frac{\sigma_w^2}{g_1} \mathbf{I}_{JM}$ due to noise. If the noise components imposed on \mathbf{Q}_F can be eliminated, then we can obtain the outer-product matrix $\mathbf{D}_F \mathbf{D}_F^*$. Next, we can take eigen-decomposition of this outer-product matrix to obtain an estimate $\hat{\mathbf{D}}_F$ of \mathbf{D}_F . However, taking eigen-decomposition of such a large

size ($JM \times JM$) of matrix $\mathbf{D}_F \mathbf{D}_F^*$ involves more computations and usually renders a less accurate result, especially when M , the number of subcarriers, is large. To avoid this drawback, we want to use (3.4) to obtain another matrix $\mathbf{H}\mathbf{H}^*$, which is the outer product of the channel impulse response matrix $\mathbf{H} = [\mathbf{H}(0)^T \ \mathbf{H}(1)^T \ \dots \ \mathbf{H}(L)^T]^T \in \mathbb{C}^{J(L+1) \times K}$. The size of $\mathbf{H}\mathbf{H}^*$ is $J(L+1) \times J(L+1)$, which is smaller than the size of $\mathbf{D}_F \mathbf{D}_F^*$.^a Hence, taking eigen-decomposition of $\mathbf{H}\mathbf{H}^*$ to obtain an estimate $\hat{\mathbf{H}}$ of \mathbf{H} requires less computational load.

To obtain $\mathbf{H}\mathbf{H}^*$ from (3.4), we first define an $M \times (L+1)$ matrix $\mathbf{F}_1 = \mathbf{F}(:, 1:L+1)$, which is the matrix containing the first $(L+1)$ columns of \mathbf{F} . In addition, the relationship between the channel frequency response matrix \mathbf{D}_F and channel impulse response matrix \mathbf{H} can be described as follows:

$$\mathbf{D}_F = \sqrt{M}(\mathbf{F}_1 \otimes \mathbf{I}_J)\mathbf{H}. \quad (3.5)$$

With the aid of (3.5) and (3.4), we obtain the following matrix \mathbf{Q}_H :

$$\begin{aligned} \mathbf{Q}_H &= \frac{1}{M}(\mathbf{F}_1^* \otimes \mathbf{I}_J)\mathbf{Q}_F(\mathbf{F}_1 \otimes \mathbf{I}_J) \\ &= \frac{1}{M}(\mathbf{F}_1^* \otimes \mathbf{I}_J)(\mathbf{D}_F \mathbf{D}_F^* + \frac{\sigma_w^2}{g_1} \mathbf{I}_{JM})(\mathbf{F}_1 \otimes \mathbf{I}_J) \\ &= \frac{1}{M}(\mathbf{F}_1^* \otimes \mathbf{I}_J)\mathbf{D}_F \mathbf{D}_F^*(\mathbf{F}_1 \otimes \mathbf{I}_J) \\ &\quad + \frac{1}{M}(\mathbf{F}_1^* \otimes \mathbf{I}_J)\frac{\sigma_w^2}{g_1} \mathbf{I}_{JM}(\mathbf{F}_1 \otimes \mathbf{I}_J) \\ &= \mathbf{H}\mathbf{H}^* + \frac{\sigma_w^2}{Mg_1} \mathbf{I}_{J(L+1)}. \end{aligned} \quad (3.6)$$

Since the matrix \mathbf{H} is of full column rank by assumption (ii), the rank of $\mathbf{H}\mathbf{H}^*$ is K . This implies that the associated smallest $J(L+1) - K$ eigenvalues of \mathbf{Q}_H in (3.6) are equal to the scaled-noise variance $\frac{\sigma_w^2}{Mg_1}$. Hence in practice, we can estimate the scaled-noise variance as the average of the smallest $J(L+1) - K$ eigenvalues of \mathbf{Q}_H . Then the outer-product matrix $\mathbf{H}\mathbf{H}^*$ can be obtained by subtracting $\frac{\sigma_w^2}{Mg_1} \mathbf{I}_{J(L+1)}$ from \mathbf{Q}_H in (3.6). Finally, taking eigen-decomposition of the Hermitian and positive semi-definite matrix $\mathbf{H}\mathbf{H}^*$ with rank K yields K positive eigenvalues and the associated unit-norm eigenvectors, say, $\lambda_1, \dots, \lambda_K$ and $\mathbf{d}_1, \dots, \mathbf{d}_K$, respectively. We can thus choose the channel impulse response matrix to be

$$\hat{\mathbf{H}} = [\sqrt{\lambda_1} \mathbf{d}_1 \ \sqrt{\lambda_2} \mathbf{d}_2 \ \dots \ \sqrt{\lambda_K} \mathbf{d}_K] \in \mathbb{C}^{J(L+1) \times K}, \quad (3.7)$$

up to a unitary matrix ambiguity $\mathbf{U} \in \mathbb{C}^{K \times K}$, i.e., $\mathbf{H} = \hat{\mathbf{H}}\mathbf{U}$, since $\hat{\mathbf{H}}\hat{\mathbf{H}}^* = \mathbf{H}\mathbf{H}^* = \mathbf{Q}$. The ambiguity matrix \mathbf{U} is intrinsic to semiblind estimation of multiple input systems using only second-order statistics technique [17]. This

ambiguity can be resolved using a short pilot sequence [18].

3.2 Precoding design

In Section 3.1, we obtain \mathbf{Q}_F from the autocorrelation matrix \mathbf{R} . However, in practice, we have $\widehat{\mathbf{R}} = \mathbf{R} + \Delta\mathbf{R}$ instead of \mathbf{R} , where $\Delta\mathbf{R}$ is the error matrix due to the presence of finite-sample estimation error. As a result, dividing each submatrix in the autocorrelation matrix $\widehat{\mathbf{R}}$ by the corresponding coefficient g_m to obtain $\widehat{\mathbf{Q}}_F$ involves an error term, i.e., $\widehat{\mathbf{Q}}_F = \mathbf{Q}_F + \Delta\mathbf{Q}_F$, where $\Delta\mathbf{Q}_F$ is the error term, and $\Delta\mathbf{Q}_F((k-1)J+1 : kJ, (l-1)J+1 : lJ) = \frac{1}{g_m} \Delta\mathbf{R}((k-1)J+1 : kJ, (l-1)J+1 : lJ)$ with $m = (k-l)_M + 1$, $\forall 1 \leq k, l \leq M$. Here $(\cdot)_M$ is the modulo- M operation. It is obvious that a large value of the corresponding g_m attenuates the error term $\Delta\mathbf{Q}_F$, which in turn increases the accuracy of estimation for $\widehat{\mathbf{Q}}_F$.

As a result, we need to design the precoding coefficients p_1, p_2, \dots, p_M to maximize g_1, g_2, \dots, g_M to reduce the error term. However, this results in a multi-objective optimization problem which does not seem to easily yield a tractable way to design. Hence, we present another feasible approach to design the precoding in the following.

Since no prior information of the distortion $\Delta\mathbf{R}$ can be obtained in advance, we combine all the M objective functions into a single cost with the same weight, i.e., $g = g_1 + g_2 + \dots + g_M$, and try to design the precoding to maximize g . In addition, it is easy to verify that $g = (p_1 + p_2 + \dots + p_M)^2$. Then the optimization problem can be formulated as follows:

$$\begin{aligned} & \max_{p_1, p_2, \dots, p_M} (p_1 + p_2 + \dots + p_M)^2 \\ \text{subject to } & \sum_{n=1}^M p_n^2 = 1. \end{aligned} \quad (3.8)$$

The constraint in (3.8) normalizes the power gain of each precoded symbol in the precoded vector $\mathbf{P}\mathbf{x}^{(k)}(n)$ to 1. Appendix shows that the optimal solution to (3.8) is

$$p_1 = p_2 = \dots = p_M = \frac{1}{\sqrt{M}}. \quad (3.9)$$

Although (3.9) is the optimal solution for channel estimation, it makes symbol detection impossible because (3.9) produces a singular matrix \mathbf{P} that can not decode the precoded vector $\mathbf{P}\mathbf{x}^{(k)}(n)$ at the receiver. To make symbol detection possible after channel estimation, we modify the optimal solution (3.9) as the following precoding scheme

$$\begin{cases} p_1 = \sqrt{\frac{1}{M} + \tau} \\ p_n = \sqrt{\frac{1}{M} - \frac{\tau}{M-1}}, \quad n = 2, 3, \dots, M, \end{cases} \quad (3.10)$$

to make a nonsingular matrix \mathbf{P} , where $0 < \tau < \frac{M-1}{M}$ is small. The solution in (3.10) is a small perturbation of the optimal solution in (3.9). In addition, if we increase τ from

0 to $\frac{M-1}{M}$, then p_1 is larger than p_n , $n = 2, 3, \dots, M$, which would improve the channel equalization performance. In the following subsection, we will prove this fact by evaluating the equalization performance under the precoding scheme (3.10).

3.3 Analysis of PEP

One approach to evaluating the equalization performance is BER analysis, but it is generally quite complex. Hence, we use PEP analysis, a technology which is widely used in space-time communications and OFDM systems, to examine the equalization performance [19-25]. In addition, to better understand the intrinsic impact of the precoding (3.10) on equalization, we assume that the channel state information is known at the receivers. This assumption also appears in [26,27] to evaluate the equalization performance. Now, let us consider the system model (2.4) with zero-forcing (ZF) equalization and drop the time index n for notational convenience.

The PEP analysis measures the probability that a symbol vector \mathbf{x} is sent but another $\tilde{\mathbf{x}} \neq \mathbf{x}$ is detected. Let $\|\cdot\|$ denote the two-norm of a vector. Then by definition, the PEP conditioned on the channel impulse response matrix \mathbf{H} is given by

$$\Pr[\mathbf{x} \rightarrow \tilde{\mathbf{x}} | \mathbf{H}] = \Pr[\|\widehat{\mathbf{x}} - \tilde{\mathbf{x}}\| < \|\widehat{\mathbf{x}} - \mathbf{x}\| | \mathbf{H}], \quad (3.11)$$

where $\widehat{\mathbf{x}} = \mathbf{x} + (\mathbf{P}^{-1} \otimes \mathbf{I}_K) \mathbf{D}^\dagger \mathbf{v}$ is the estimate of \mathbf{x} after ZF equalization, and \mathbf{D}^\dagger is the pseudo-inverse of \mathbf{D} . Let $d = \|\mathbf{x} - \tilde{\mathbf{x}}\|$ be the distance between \mathbf{x} and $\tilde{\mathbf{x}}$, and let $\mathbf{e} = \frac{\tilde{\mathbf{x}} - \mathbf{x}}{d}$ be the normalized error vector. Then (3.11) can be directly simplified to

$$\Pr[\mathbf{x} \rightarrow \tilde{\mathbf{x}} | \mathbf{H}] = \Pr[u > \frac{d}{2} | \mathbf{H}], \quad (3.12)$$

where $u = \text{Re}[\mathbf{v}^* (\mathbf{D}^\dagger)^* (\mathbf{P}^{-T} \otimes \mathbf{I}_K) \mathbf{e}]$ and $\text{Re}[\cdot]$ denotes the real part. Since each element in \mathbf{v} is a zero-mean circular Gaussian random variable with variance σ_w^2 , the random variable u is also a zero-mean Gaussian with variance

$$\begin{aligned} E[u^2] &= E[|\text{Re}[\mathbf{v}^* (\mathbf{D}^\dagger)^* (\mathbf{P}^{-T} \otimes \mathbf{I}_K) \mathbf{e}]|^2] \\ &= \frac{\sigma_w^2}{2} \|(\mathbf{D}^\dagger)^* (\mathbf{P}^{-T} \otimes \mathbf{I}_K) \mathbf{e}\|^2. \end{aligned} \quad (3.13)$$

Hence, the conditional PEP in (3.12) becomes

$$\Pr[\mathbf{x} \rightarrow \tilde{\mathbf{x}} | \mathbf{H}] = Q\left(\frac{d}{\sqrt{2\sigma_w^2} \|(\mathbf{D}^\dagger)^* (\mathbf{P}^{-T} \otimes \mathbf{I}_K) \mathbf{e}\|}\right), \quad (3.14)$$

where $Q(\cdot)$ is the Q -function [28]. Let $\|\cdot\|_F$ denote the Frobenius norm of a matrix. Then by the submultiplicative property of matrix norms [29], we have

$$\begin{aligned} \|(\mathbf{D}^\dagger)^*(\mathbf{P}^{-T} \otimes \mathbf{I}_K)\mathbf{e}\| &\leq \|(\mathbf{D}^\dagger)^*\|_F \cdot \|\mathbf{P}^{-T} \otimes \mathbf{I}_K\|_F \cdot \|\mathbf{e}\| \\ &= \sqrt{K}\|(\mathbf{D}^\dagger)^*\|_F \cdot \|\mathbf{P}^{-T}\|_F \quad (3.15) \\ &= \sqrt{K}\|(\mathbf{D}^\dagger)^*\|_F \cdot \|\mathbf{P}^{-1}\|_F. \end{aligned}$$

The first equality in (3.15) holds since the two-norm of the unit vector \mathbf{e} is 1, and the second equality holds since the Frobenius norm of a matrix \mathbf{A} equals the Frobenius norm of \mathbf{A}^T .

Let us now focus on $\|\mathbf{P}^{-1}\|_F$ in (3.15). Since \mathbf{P} is a circulant matrix, it can be decomposed as $\mathbf{P} = \mathbf{F}^{-1}\mathbf{D}_p\mathbf{F}$, and the inverse of \mathbf{P} can be expressed as $\mathbf{P}^{-1} = \mathbf{F}^{-1}\mathbf{D}_p^{-1}\mathbf{F}$, where \mathbf{D}_p is a diagonal matrix with eigenvalues of \mathbf{P} on its diagonal [16]. For \mathbf{P} with coefficients $\{p_1, p_2, \dots, p_M\}$ given in (3.10), the first row of \mathbf{P} is $[a \ b \ b \ \dots \ b]$, where $a = \sqrt{\frac{1}{M} + \tau}$ and $b = \sqrt{\frac{1}{M} - \frac{\tau}{M-1}}$. Then the eigenvalues of \mathbf{P} are given by the DFT of the first row of \mathbf{P} [16] to form the diagonal matrix $\mathbf{D}_p = \text{diag}[a + (M-1)b, a - b, \dots, a - b] \in \mathbb{R}^{M \times M}$. This leads to

$$\begin{aligned} \|\mathbf{P}^{-1}\|_F &= \|\mathbf{F}^{-1}\mathbf{D}_p^{-1}\mathbf{F}\|_F \\ &\leq \|\mathbf{F}^{-1}\|_F \cdot \|\mathbf{D}_p^{-1}\|_F \cdot \|\mathbf{F}\|_F \\ &= \sqrt{M} \cdot \sqrt{[a+(M-1)b]^{-2} + (M-1)(a-b)^{-2}} \cdot \sqrt{M} \\ &\leq M\sqrt{(a-b)^{-2} + (M-1)(a-b)^{-2}} \\ &= M\sqrt{M}(a-b)^{-1}. \end{aligned} \quad (3.16)$$

From (3.15) and (3.16), we know

$$\|(\mathbf{D}^\dagger)^*(\mathbf{P}^{-T} \otimes \mathbf{I}_K)\mathbf{e}\| \leq \sqrt{KM^3}\|(\mathbf{D}^\dagger)^*\|_F(a-b)^{-1}. \quad (3.17)$$

Using (3.17), we know the conditional PEP (3.14) is upper bounded by

$$\Pr[\mathbf{x} \rightarrow \tilde{\mathbf{x}} | \mathbf{H}] \leq Q\left(\frac{d(a-b)}{\sqrt{2\sigma_w^2 KM^3}\|(\mathbf{D}^\dagger)^*\|_F}\right). \quad (3.18)$$

From (3.18), it is obvious that we can increase a (i.e., p_1 or τ) to decrease the upper bound of PEP, which in turn reduces the symbol/bit detection error. However, it is easy to check that increasing τ from 0 to $\frac{M-1}{M}$ would decrease the value of the objective function $g = (p_1 + p_2 + \dots + p_M)^2$ in (3.8), which means the estimation performance deteriorates. Hence, there is a tradeoff in the selection of τ between channel estimation and equalization. In the work of [4,11], this tradeoff is also observed. We will give a simulation example to demonstrate this tradeoff in Section 4.

3.4 Discussion

We now give some further comments about the proposed method.

- (1) *Channel identifiability and the case of more transmitters:* The channel identifiability condition, $\text{rank}(\mathbf{H}) = K$ (assumption (ii)), for the proposed method is the same as that in methods [11,12,30], but is more relaxed than the identifiability conditions for methods [9,10]. If assumption (ii) does not hold, i.e., the matrix \mathbf{H} is rank deficient with $\text{rank}(\mathbf{H}) = W < K$, then $\text{rank}(\mathbf{H}\mathbf{H}^*) = W < K$. In this case, we could only choose W positive eigenvalues and the associated eigenvectors from $\mathbf{H}\mathbf{H}^*$, which can not form the matrix $\hat{\mathbf{H}}$ in (3.7) in theory.

In addition, since the size of the channel impulse response matrix \mathbf{H} is $J(L+1) \times K$, $\text{rank}(\mathbf{H}) = K$ implies

$$J(L+1) \geq K, \quad (3.19)$$

i.e., the product of the number of receivers (J) and the channel length ($L+1$) should be no less than the number of transmitters (K). Hence, the proposed method is capable of identifying not only the more receivers case ($J \geq K$), but also the more transmitters case ($K > J$) as long as (3.19) is fulfilled.

- (2) *Channel order overestimation:* So far we have assumed that the channel order L is known. If L is unknown, we can set P , the length of CP, as an upper bound of L since $P \geq L$ is required to avoid interblock interference. With this upper bound $\hat{L} = P$ and following the process given in Section 3.1, the corresponding matrix \mathbf{Q}_H in (3.6) can be similarly constructed as $\mathbf{Q}_H = \mathbf{H}_{\text{ov}}\mathbf{H}_{\text{ov}}^* + \frac{\sigma_w^2}{Mg_1}\mathbf{I}_{J(\hat{L}+1)}$, where $\mathbf{H}_{\text{ov}} = [\mathbf{H}^T \ \mathbf{0} \ \dots \ \mathbf{0}]^T \in \mathcal{C}^{J(\hat{L}+1) \times K}$. Then, similar

($\hat{L}-L$) blocks

to the method given in Section 3.1, we eliminate the noise contribution imposed on \mathbf{Q}_H to obtain $\mathbf{H}_{\text{ov}}\mathbf{H}_{\text{ov}}^*$. Note that the last $(\hat{L}-L)$ block columns and block rows of $\mathbf{H}_{\text{ov}}\mathbf{H}_{\text{ov}}^*$ are zero. Hence, $\text{rank}(\mathbf{H}_{\text{ov}}\mathbf{H}_{\text{ov}}^*) = K$ and $\mathbf{H}_{\text{ov}}\mathbf{H}_{\text{ov}}^*$ has K positive eigenvalues. Each of the associated eigenvectors has the form $\hat{\mathbf{d}} = [\mathbf{d}^T \ \mathbf{0} \ \dots \ \mathbf{0}]^T \in \mathcal{C}^{J(\hat{L}+1)}$ where $\mathbf{d} \in \mathcal{C}^{J(L+1)}$. From these K positive eigenvalues and the associated K eigenvectors, we can estimate $\hat{\mathbf{H}}_{\text{ov}}$, and then obtain the channel impulse response matrix $\hat{\mathbf{H}}$ from $\hat{\mathbf{H}}_{\text{ov}}$ up to a matrix ambiguity.

- (3) *Algorithm:* We now summarize the proposed approach as the following algorithm:

- (1) Collect the received data $\mathbf{y}(n)$, and then estimate the autocorrelation matrix \mathbf{R} via the

following time average

$$\hat{\mathbf{R}} = \frac{1}{S} \sum_{n=1}^S \mathbf{y}(n)\mathbf{y}(n)^*, \quad (3.20)$$

where S is the number of data blocks.

- (2) Use the precoding coefficients in (3.10) to compute \mathbf{Q}_F from the autocorrelation matrix of $\hat{\mathbf{R}}$.
- (3) Form the matrix \mathbf{Q}_H using (3.6) and \mathbf{Q}_F obtained from the previous step.
- (4) Use the method given in Section 3.1 to remove the noise components imposed on \mathbf{Q}_H to obtain $\mathbf{H}\mathbf{H}^*$.
- (5) Finally, obtain the channel impulse response matrix \mathbf{H} by computing the K largest eigenvalues and the associated eigenvectors of $\mathbf{H}\mathbf{H}^*$.

4 Simulation

In this section, we generate 100 2-input 2-output random channels with $L = 6$ for each simulation (except simulation 3) to demonstrate the performance of the proposed method. The number of subcarriers for one OFDM block is $M = 36$, and the length of CP is $P = 6$. Each channel coefficient in the channel impulse response matrix is generated according to the independent complex-valued Gaussian distribution with zero-mean and unit variance. The normalized mean-square-error (NMSE) of the channel impulse response matrix is defined as $\text{NMSE} = (1/I) \sum_{i=1}^I \|\hat{\mathbf{H}}^{(i)} - \mathbf{H}\|_F^2 \cdot \|\mathbf{H}\|_F^{-2}$, where $I = 100$ is the number of independent trials. $\hat{\mathbf{H}}^{(i)} = [\hat{\mathbf{H}}^{(i)}(0)^T \ \hat{\mathbf{H}}^{(i)}(1)^T \ \dots \ \hat{\mathbf{H}}^{(i)}(6)^T]^T$ is the i th estimate of the channel impulse response matrix \mathbf{H} after removing the unitary matrix ambiguity by the least squares method [17]. The number of symbol blocks is $S = 100$. The input source symbols are quadrature-phase-shift-keying (QPSK) signals. The channel noise is zero-mean, temporally and spatially white Gaussian. The signal-to-noise ratio (SNR) at the output is defined as $\text{SNR} = \frac{E[\|\mathbf{y}(n) - \mathbf{v}(n)\|_2^2]}{E[\|\mathbf{v}(n)\|_2^2]}$.

4.1 Simulation 1: the effect of the precoding on channel estimation and equalization

In this simulation, we use 4 different precoders based on (3.10) with $\tau = 0.2, 0.4, 0.6,$ and 0.8 , to illustrate the effect of the precoding on channel estimation and ZF equalization. Figure 2 shows that NMSE decreases as SNR increases for each precoder. In this figure, we also see that the estimation performs better for smaller τ , which is consistent with the analysis at the end in Section 3.3. Figure 3 shows that the BER improves as τ increases, since the analysis of PEP shows that a larger τ can lower the upper

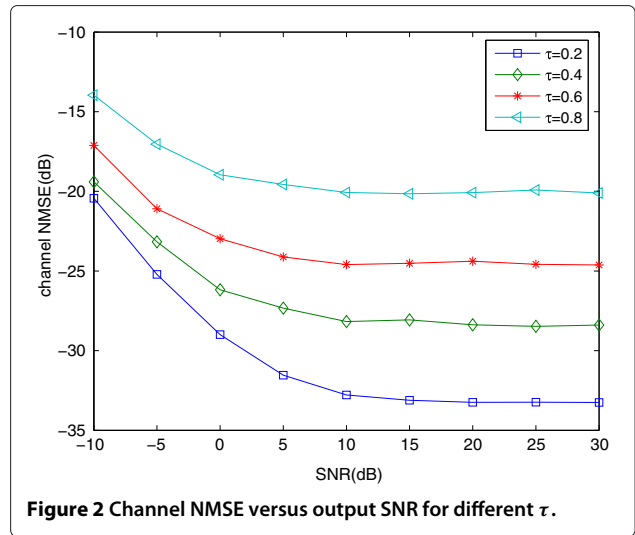


Figure 2 Channel NMSE versus output SNR for different τ .

bound of PEP, which in turn improves symbol/bit detection at the receiver. From Figures 2 and 3, we know there is a tradeoff between channel estimation and equalization, and the selection of τ should depend on the scenarios we meet.

4.2 Simulation 2: robustness to channel order overestimation

In this simulation, we use the precoding scheme that satisfies (3.10) with different τ and fix SNR= 10 dB. For each upper bound \hat{L} with $0 \leq (\hat{L} - L) \leq 5$, we choose $P = \hat{L}$ and $M = 9P$ for simulation such that the transmission efficiency is maintained at 90%. Figure 4 shows that the proposed method is reasonably robust to channel order overestimation since the NMSE increases slowly for each τ as $(\hat{L} - L)$ increases from 0 to 5.

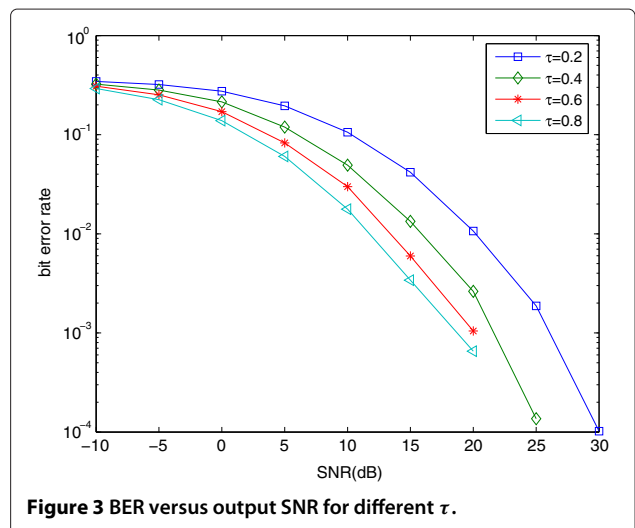
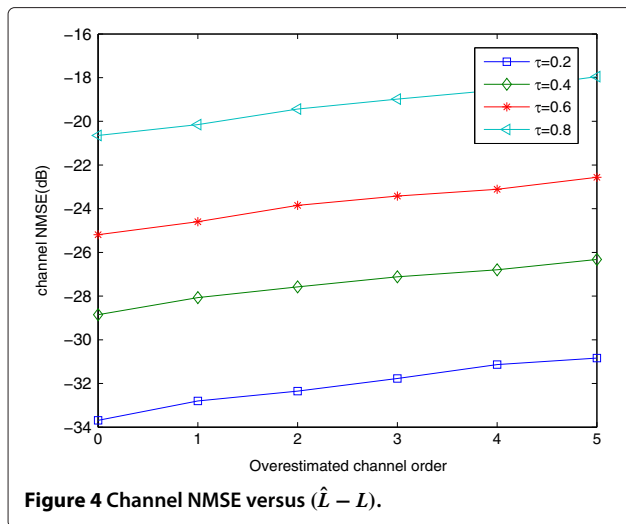


Figure 3 BER versus output SNR for different τ .

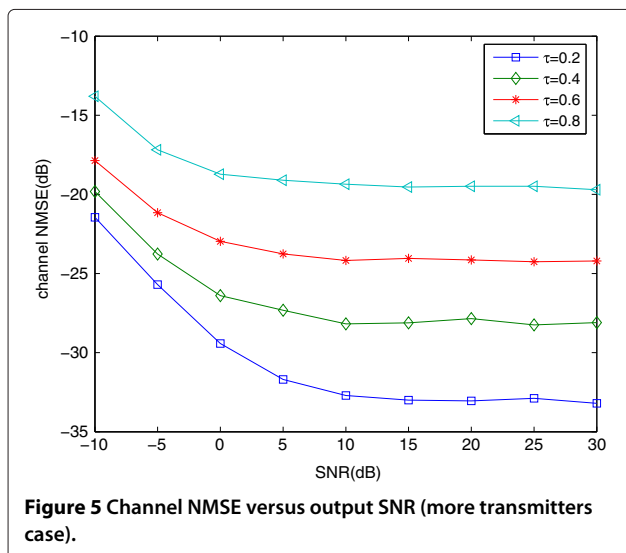


4.3 Simulation 3: channels with more transmitters than receivers

In this simulation, we generate 100 3-input 2-output random channels with $L = 6$ to illustrate the performance of the proposed method for channels with more transmitters than receivers. The precoding scheme is chosen based on (3.10) with different τ . Figure 5 shows that the NMSE decreases as SNR increases. This figure also shows that the proposed method can apply to the channels with more transmitters than receivers.

4.4 Simulation 4: comparison with existing methods

In this simulation, we compare the ZF equalization performances achieved by the proposed method (with $\tau = 0.8$), one subspace method [9], and three precoding methods [10-12]. For the precoding matrices in [10-12], the precoding coefficients are $\{\beta = 0.432, \alpha = 0.9, \gamma = 1.1\}$,

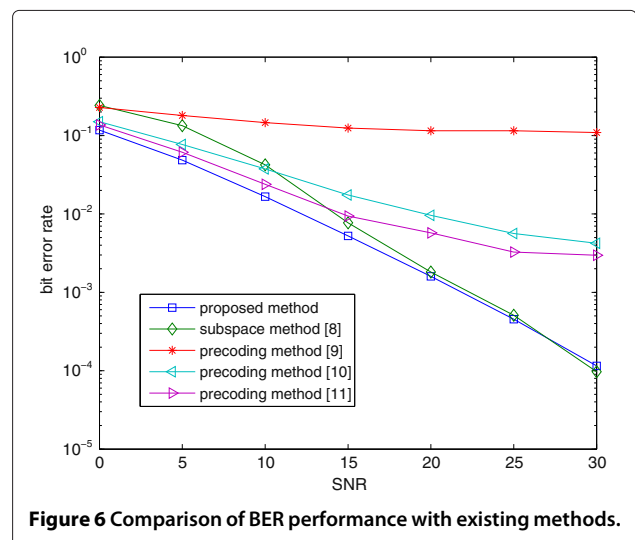


($p = 0.2$), and ($\alpha = 0.8, \delta = 0.05$), respectively. Figure 6 shows that the proposed method outperforms the three precoding methods. The reason may be due to no systematic procedures for the precoding designs are given in [10-12] to combat against the noise effects and numerical errors; while the proposed method not only works out a way to remove the noise components, but also appropriately develops a precoding to combat against the numerical errors.

Figure 6 also shows that the proposed method performs better than the subspace method in the low-to-medium SNR region ($\text{SNR} < 25$ dB), and for high SNR, the subspace method performs better than the proposed method. Since the subspace method enjoys the so-called “finite sample convergence” property [22-24], that is, in the noiseless case (or sufficiently high SNR), the channels can be almost exactly identified by using a finite number of samples for autocorrelation estimation, it is expected that the subspace-based solution can yield improved channel estimation accuracy and the resultant BER in the high SNR region, as compared with the proposed method.

5 Conclusions

In this article, we propose a semiblind channel estimation method for MIMO-OFDM systems based on circular precoding. By taking advantage of circular precoding, we obtain the outer product of the channel impulse response matrix \mathbf{H} from the autocorrelation matrix of the received data. Then the channel impulse response matrix can be obtained by computing the positive eigenvalues and the corresponding eigenvectors of the outer-product matrix $\mathbf{H}\mathbf{H}^*$. We also study the precoding design to combat the numerical error of estimation for the autocorrelation matrix, and discuss the effects of precoding on channel estimation and equalization. With the proposed



framework, the method is reasonably robust to channel order overestimation and the identifiability condition is simply that the channel impulse response matrix has full column rank. Thanks to the identifiability condition, the proposed method is applicable to MIMO channels with more transmitters or more receivers. The simulations in this study also demonstrate the performance of the proposed method.

Endnote

^a Since the CP is actually a copy of the last portion of $\mathbf{s}^{(k)}(n) \in \mathbb{C}^M$, the length of CP, P , is less than M (i.e., $P < M$). In general, for transmission efficiency, P is usually less than or equal to $0.25M$. In addition, in Section 2, we know the length of CP is longer than or equal to L (i.e., $L \leq P$) to combat against the channel delay spread. Hence we have $L + 1 < M$, which implies the size of $\mathbf{H}\mathbf{H}^*$ is smaller than the size of $\mathbf{D}_F\mathbf{D}_F^*$.

Appendix

The optimal solution to (3.8)

To find the optimal solution to (3.8), we first formulate (3.8) as

$$\begin{aligned} \min_{p_1, p_2, \dots, p_M} & f(p_1, p_2, \dots, p_M) \\ \text{subject to} & h(p_1, p_2, \dots, p_M) = 0, \end{aligned} \quad (\text{A.1})$$

where $f(p_1, p_2, \dots, p_M) = -(p_1 + p_2 + \dots + p_M)^2$ and $h(p_1, p_2, \dots, p_M) = \sum_{n=1}^M p_n^2 - 1$. Then we form a Lagrangian function γ as (A.2)

$$\gamma(p_1, p_2, \dots, p_M, \lambda) = f(p_1, p_2, \dots, p_M) + \lambda h(p_1, p_2, \dots, p_M), \quad (\text{A.2})$$

where λ is the Lagrange multiplier [31]. The first-order necessary condition provided from Lagrange's theorem says that for a point (p_1, p_2, \dots, p_M) to be a minimizer, the partial derivatives of the Lagrangian function must satisfy

$$\begin{cases} \frac{\partial \gamma}{\partial p_1} = -2 \sum_{n=1}^M p_n + 2\lambda p_1 = 0 \\ \frac{\partial \gamma}{\partial p_2} = -2 \sum_{n=1}^M p_n + 2\lambda p_2 = 0 \\ \vdots \\ \frac{\partial \gamma}{\partial p_M} = -2 \sum_{n=1}^M p_n + 2\lambda p_M = 0 \\ \frac{\partial \gamma}{\partial \lambda} = \sum_{n=1}^M p_n^2 - 1 = 0. \end{cases} \quad (\text{A.3})$$

From the first M equations in (A.3), we know

$$\sum_{n=1}^M p_n = \lambda p_1 = \lambda p_2 = \dots = \lambda p_M. \quad (\text{A.4})$$

There are two cases for (A.4), $\sum_{n=1}^M p_n = 0$ and $\sum_{n=1}^M p_n \neq 0$. For the first case, it is obvious that the objective function $f = 0$. However, it is easy to find a solution that conforms to the constraint $h = 0$, for example, $(p_1, p_2, p_3, \dots, p_M) = (1, 0, 0, \dots, 0)$, such that the objective function $f = -1$ which is smaller than $f = 0$. Hence, the first case can not minimize the objective function. Thus, it is necessary to focus the second case, $\sum_{n=1}^M p_n \neq 0$.

For $\sum_{n=1}^M p_n \neq 0$, (A.4) implies $\lambda \neq 0$ and $p_1 = p_2 = \dots = p_M$. Combining (A.4) with the last equation in (A.3) shows that $(p_1^*, p_2^*, \dots, p_M^*, \lambda^*) = (\frac{1}{\sqrt{M}}, \frac{1}{\sqrt{M}}, \dots, \frac{1}{\sqrt{M}}, M)$ is the unique solution to (A.3) for $\sum_{n=1}^M p_n \neq 0$, which implies $(p_1^*, p_2^*, \dots, p_M^*) = (\frac{1}{\sqrt{M}}, \frac{1}{\sqrt{M}}, \dots, \frac{1}{\sqrt{M}})$ is the unique candidate for being a minimizer. We now want to verify if $(p_1^*, p_2^*, \dots, p_M^*, \lambda^*)$ satisfies the second-order sufficient condition.

According to [31], we form the Hessian matrix $\Gamma(p_1, p_2, \dots, p_M, \lambda)$ of the Lagrangian function $\gamma(p_1, p_2, \dots, p_M, \lambda)$ at $(p_1^*, p_2^*, \dots, p_M^*, \lambda^*)$ as

$$\Gamma(p_1^*, p_2^*, \dots, p_M^*, \lambda^*) = \begin{bmatrix} 4M-2 & -2 & -2 & \dots & -2 \\ -2 & 4M-2 & -2 & \dots & -2 \\ -2 & -2 & 4M-2 & \dots & -2 \\ \vdots & \vdots & \dots & \ddots & \vdots \\ -2 & -2 & -2 & \dots & 4M-2 \end{bmatrix}. \quad (\text{A.5})$$

Let $\mathbf{r} = [r_1 \quad r_2 \dots r_M]^T$ be a vector and let $\nabla h(p_1^*, p_2^*, \dots, p_M^*) = [\frac{2}{\sqrt{M}} \quad \frac{2}{\sqrt{M}} \dots \frac{2}{\sqrt{M}}]$ be the gradient of h at $(p_1^*, p_2^*, \dots, p_M^*)$. If we can verify that the Hessian matrix Γ in (A.5) is positive definite for all $\mathbf{r} \neq \mathbf{0}$ with $\nabla h(p_1^*, p_2^*, \dots, p_M^*)\mathbf{r} = 0$, then $(p_1^*, p_2^*, \dots, p_M^*)$ is the minimizer. To see this, we know

$$\begin{aligned} \nabla h(p_1^*, p_2^*, \dots, p_M^*)\mathbf{r} &= 0 \\ \Leftrightarrow r_1 + r_2 + \dots + r_M &= 0. \end{aligned} \quad (\text{A.6})$$

Now

$$\begin{aligned} & \mathbf{r}^T \Gamma(p_1^*, p_2^*, \dots, p_M^*, \lambda^*)\mathbf{r} \\ &= (4Mr_1^2 - 2r_1 \sum_{i=1}^M r_i) + (4Mr_2^2 - 2r_2 \sum_{i=1}^M r_i) \\ & \quad + \dots + (4Mr_M^2 - 2r_M \sum_{i=1}^M r_i) \\ &= 4M(r_1^2 + r_2^2 + \dots + r_M^2) > 0, \quad \forall \mathbf{r} \neq \mathbf{0}. \end{aligned} \quad (\text{A.7})$$

Note that the second equality in (A.7) holds due to (A.6). From (A.7), we know the Hessian matrix $\Gamma(p_1^*, p_2^*, \dots, p_M^*, \lambda^*)$ is positive definite for all $\mathbf{r} \neq \mathbf{0}$ with $\nabla h(p_1^*, p_2^*, \dots, p_M^*)\mathbf{r} = 0$. Hence, the solution $(p_1^*, p_2^*, \dots, p_M^*) = (\frac{1}{\sqrt{M}}, \frac{1}{\sqrt{M}}, \dots, \frac{1}{\sqrt{M}})$ is the optimal solution to (A.1), and the objective function $f(p_1^*, p_2^*, \dots, p_M^*) = -M$.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

This research was sponsored by the National Science Council under grant NSC-98-2221-E035-043.

Received: 8 March 2012 Accepted: 2 June 2012

Published: 3 October 2012

References

1. LJ Cimini Jr., Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing. *IEEE Trans. Commun.* **33**(7), 665–675 (1985)
2. JA Bingham, Multicarrier modulation for data transmission: an idea whose time has come. *IEEE Commun. Mag.* **28**(5), 5–14 (1990)
3. L Hanzo, M Munster, BJ Choi, T Keller, *OFDM and MC-CDMA for Broadband Multi-User Communications, WLANs and broadcasting* (Wiley, West Sussex, England, 2003)
4. R Lin, AP Petropulu, Linear precoding assisted blind channel estimation for OFDM systems. *IEEE Trans. Veh. Technol.* **54**(3), 983–995 (2005)
5. GL Stuber, JR Barry, SW McLaughlin, Y Li, MA Ingram, TG Pratt, Broadband MIMO-OFDM wireless communications. *Proc. IEEE*. **92**(2), 271–294 (2004)
6. K Abed-Meraim, W Qiu, Y Hua, Blind system identification. *Proc. IEEE*. **85**(8), 1310–1332 (1997)
7. Z Ding, Y Li, *Blind Equalization and Identification* (Marcel Dekker, Inc., New York, 2001)
8. GB Giannakis, Y Hua, P Stoica, L Tong, *Signal Processing, Advances in Wireless and Mobile Communications Volume I: Trends in Channel Identification and Equalization* (Prentice Hall, PTR Upper Saddle River, 2001)
9. F Gao, Y Zeng, A Nallanathan, TS Ng, Robust subspace blind channel estimation for cyclic prefixed MIMO OFDM systems: algorithm, identifiability and performance analysis. *IEEE J. Sel. Areas Commun.* **26**(2), 378–388 (2008)
10. S Yatawatta, AP Petropulu, blind channel estimation in MIMO OFDM systems with multiuser interference. *IEEE Trans. Signal Process.* **54**(3), 1054–1068 (2006)
11. F Gao, A Nallanathan, Blind channel estimation for MIMO OFDM systems via nonredundant linear precoding. *IEEE Trans. Signal Process.* **55**(2), 784–789 (2007)
12. C Shin, RW Heath Jr., EJ Powers, Non-redundant precoding-based blind and semi-blind channel estimation for MIMO block transmission with a cyclic prefix. *IEEE Trans. Signal Process.* **56**(6), 2509–2523 (2008)
13. AP Petropulu, R Lin, Blind OFDM channel estimation through simple linear precoding. *IEEE Trans. Wirel. Commun.* **3**(2), 647–655 (2004)
14. R Zhang, Blind channel estimation for precoded OFDM system. in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process*, vol. 3. (Philadelphia, Pennsylvania, USA, 2005), pp. 469–472
15. T Kim, I Eo, Blind channel estimation and equalization in OFDM system with circular precoding. in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process*, vol. 4. (Toulouse, France, 2006), pp. 341–344
16. TK Moon, WC Stirling, *Mathematical Methods and Algorithms for Signal Processing* (Prentice Hall, Upper Saddle River, 2000)
17. Z Ding, Matrix outer-product decomposition method for blind multiple channel identification. *IEEE Trans. Signal Process.* **45**(12), 3053–3061 (1997)
18. YS Chen, Semiblind channel estimation for MIMO single carrier with frequency-domain equalization systems. *IEEE Trans. Veh. Technol.* **59**(1), 53–62 (2010)
19. Z Liu, Y Xin, GB Giannakis, Linear constellation precoding for OFDM with maximal multipath diversity and coding gain. *IEEE Trans. Commun.* **51**(3), 416–427 (2003)
20. C Tepedelenlioğlu, Maximum multipath diversity with linear equalization in precoded OFDM systems. *IEEE Trans. Inf. Theory*. **50**(1), 232–235 (2004)
21. S Zhou, GB Giannakis, Single-carrier space-time block-coded transmissions over frequency-selective fading channels. *IEEE Trans. Inf. Theory*. **49**(1), 164–179 (2003)
22. L Tong, Q Zhao, Joint order detection and blind channel estimation by least squares smoothing. *IEEE Trans. Signal Process.* **47**(9), 2345–2355 (1999)
23. Q Zhao, L Tong, Adaptive blind channel estimation by least squares smoothing. *IEEE Trans. Signal Process.* **47**(12), 3000–3012 (1999)
24. X Yu, L Tong, Joint channel and symbol estimation by oblique projections. *IEEE Trans. Signal Process.* **49**(12), 3074–3083 (2001)
25. EG Larsson, P Stoica, *Space-Time Block Coding for Wireless Communications* (Cambridge University Press, Cambridge, 2003)
26. JY Wu, TS Lee, Periodic-modulation-based blind channel identification for the single-carrier block transmission with frequency-domain equalization. *IEEE Trans. Signal Process.* **54**(3), 1114–1130 (2006)
27. A Chevreuil, E Serpedin, P Loubaton, GB Giannakis, Blind channel identification and equalization using periodic modulation precoders: performance analysis. *IEEE Trans. Signal Process.* **48**(6), 1570–1586 (2000)
28. JG Proakis, M Salehi, *Essentials of Communication Systems Engineering* (Prentice Hall, Upper Saddle River, 2005)
29. GH Golub, CF Van Loan, *Matrix Computations*, 3rd edn. (The Johns Hopkins University Press, Baltimore, 1996)
30. CA Lin, YS Chen, Blind identification of MIMO channels using optimal periodic precoding. *IEEE Trans. Circ. Syst. Reg. Papers.* **54**(4), 901–911 (2007)
31. DG Luenberger, *Linear and Nonlinear Programming*, 2nd edn. (Kluwer Academic Publishers, Boston, 2003)

doi:10.1186/1687-6180-2012-212

Cite this article as: Chen and Song: Semiblind channel estimation for MIMO-OFDM systems. *EURASIP Journal on Advances in Signal Processing* 2012 **2012**:212.

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com