Training Signal Design for MIMO OFDM Channel Estimation in the Presence of Frequency Offsets

Hlaing Minn, *Member, IEEE* and Naofal Al-Dhahir, *Senior Member, IEEE*
Department of Electrical Engineering, University of Texas at Dallas
Email: { hlaing.minn, aldhahir }@utdallas.edu

***Abstract*— All existing training signal designs for channel estimation in OFDM systems assume no frequency offsets. In practice, frequency offset is unavoidable and seriously degrades the performance of OFDM systems. In this paper, we address the problem of designing optimal training signals for the estimation of frequency-selective channels in MIMO OFDM systems with frequency offsets. The mean square error advantage of our proposed optimal training signals can be quite significant for moderate-to-high values of SNR and frequency offsets.**

I. INTRODUCTION

Training signal design for channel estimation is a well-studied problem for SISO systems (e.g., see [1]-[6] and references therein) but a relatively new one for MIMO systems. For MIMO single-carrier systems, optimal design and placement of pilot symbols for channel estimation were presented in [4] while space-time-coded training signal designs were studied in [7]. Recently, optimal training signal designs for MIMO OFDM systems were addressed in [8]-[10].

To the best of our knowledge, all existing training signal designs for OFDM channel estimation assume no frequency offset. In practice, frequency offset is unavoidable due to local oscillator mismatches and Doppler shifts of the mobile wireless channels. Frequency offset causes a loss of orthogonality among the sub-carriers and seriously degrades the performance of OFDM systems. To mitigate OFDM's high sensitivity to frequency offsets, several highly-accurate frequency offset estimators, were proposed. However, there will still be some residual frequency offset after applying these frequency offset compensation techniques. Hence, in this paper, we derive the optimal training signals for MIMO OFDM channel estimation which are the most robust to frequency offsets among the existing training signals. Our results show that MIMO OFDM systems with a larger number of transmit antennas are more sensitive to frequency offsets and the performance improvement of the proposed optimal training signals can be quite significant for moderate-to-high values of SNR and frequency offsets.

II. SIGNAL MODEL AND OPTIMAL TRAINING SIGNALS IN THE ABSENCE OF FREQUENCY OFFSET

Consider a MIMO OFDM system with K sub-carriers where training signals from N_{Tx} transmit antennas are transmitted over Q OFDM symbols. Since the same channel estimation procedure is performed at each receive antenna, we only need to consider one receive antenna in designing optimal training signals. The channel impulse response (CIR) for each transmit-receive antenna pair (including all transmit/receive filtering effects) is assumed to have L taps and is quasi-static over Q OFDM symbols. Let $[c_{n,q}[0], \ldots, c_{n,q}[K-1]]^T$ be the pilot tones vector of the n-th transmit-antenna at the q-th symbol interval and $\{s_{n,q}[k] : k = -N_g, \ldots, K-1\}$ be the corresponding time-domain complex baseband training samples, including N_g $(\geq L-1)$ cyclic prefix samples where the superscript T denotes the transpose. Define $\boldsymbol{S}_n[q]$ as the training signal matrix of size $K \times L$ for the n-th transmit antenna at the q-th symbol interval whose elements are given by $[\boldsymbol{S}_n[q]]_{m,l} = s_{n,q}[l-m]$ for $m \in \{0, \ldots, K-1\}$ and $l \in \{0, \ldots, L-1\}$. Assume that $K = ML_0$ where $M \in \{1, 2, \ldots\}$, and $L_0 \geq L$ and let $\boldsymbol{h}_n$ denote the length-L CIR vector corresponding to the n-th transmit antenna.

After cyclic prefix removal at the receiver, denote the received vector of length K at the q-th symbol interval by $\boldsymbol{r}_q$. Then, the received vector over the Q symbol intervals is

$$\boldsymbol{r} = \boldsymbol{S}\boldsymbol{h} + \boldsymbol{n} \tag{1}$$

where

$$\boldsymbol{r} = [\boldsymbol{r}_0^T \; \boldsymbol{r}_1^T \; \ldots, \boldsymbol{r}_{Q-1}^T]^T \tag{2}$$

$$\boldsymbol{h} = [\boldsymbol{h}_0^T \boldsymbol{h}_1^T \ldots \boldsymbol{h}_{N_{\mathrm{Tx}}-1}^T]^T \tag{3}$$

$$\boldsymbol{S} = \begin{bmatrix} \boldsymbol{S}_0[0] & \boldsymbol{S}_1[0] & \ldots & \boldsymbol{S}_{N_{\mathrm{Tx}}-1}[0] \\ \boldsymbol{S}_0[1] & \boldsymbol{S}_1[1] & \ldots & \boldsymbol{S}_{N_{\mathrm{Tx}}-1}[1] \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{S}_0[Q-1] & \boldsymbol{S}_1[Q-1] & \ldots & \boldsymbol{S}_{N_{\mathrm{Tx}}-1}[Q-1] \end{bmatrix} \tag{4}$$

and $\boldsymbol{n}$ is a length-KQ vector of zero-mean, circularly-symmetric, uncorrelated complex Gaussian noise samples with equal variance of σ_n^2.

The least-squares channel estimate (also maximum likelihood in this case), assuming $\boldsymbol{S}^H\boldsymbol{S}$ has full rank, is given by

$$\hat{\boldsymbol{h}} = (\boldsymbol{S}^H\boldsymbol{S})^{-1}\boldsymbol{S}^H\boldsymbol{r} \tag{5}$$

and the corresponding mean square error (MSE) is given by $\sigma_n^2 \, tr\{(\boldsymbol{S}^H\boldsymbol{S})^{-1}\}$. The minimum MSE is achieved if and only if

$$\boldsymbol{S}^H\boldsymbol{S} = E_{\mathrm{av}}\boldsymbol{I} \tag{6}$$

$$\text{where} \qquad E_{\mathrm{av}} = \frac{1}{N_{\mathrm{Tx}}} \sum_{n=0}^{N_{\mathrm{Tx}}-1} \sum_{q=0}^{Q-1} \sum_{k=0}^{K-1} |s_{n,q}[k]|^2. \tag{7}$$

The corresponding minimum MSE is $LN_{\mathrm{Tx}}\sigma_n^2/E_{\mathrm{av}}$. Condition (6) can be equivalently stated as

$$\begin{aligned} \mathit{Condition}-A: \quad \sum_{q=0}^{Q-1} \boldsymbol{S}_i^H[q]\boldsymbol{S}_i[q] &= E_{\mathrm{av}}\boldsymbol{I},\ \forall i \qquad (8) \\ \mathit{Condition}-B: \quad \sum_{q=0}^{Q-1} \boldsymbol{S}_i^H[q]\boldsymbol{S}_j[q] &= \boldsymbol{0},\ \forall i \neq j. \qquad (9) \end{aligned}$$

Based on Conditions A and B, we presented in [10] general classes of optimal training signals in the absence of frequency offset which achieve this minimum MSE. Optimal pilot tone allocations across the transmit antennas were classified as frequency-division multiplexing (FDM), time-division multiplexing (TDM), code-division multiplexing in frequency-domain (CDM(F)), code-division multiplexing in time-domain (CDM(T)), and combinations thereof. In practical systems where frequency offsets are unavoidable, the optimal training signals presented in [10] may not result in the same MSE performance and the best training signals among them need to be investigated.

III. THE OPTIMAL TRAINING SIGNALS IN THE PRESENCE OF FREQUENCY OFFSETS

For completeness, in the following we summarize the main DFT properties used in this paper. Let $X[n] = \sum_{k=0}^{K-1} x[k]e^{-j2\pi kn/K}$ and $x[k] = \frac{1}{K}\sum_{n=0}^{K-1} X[n]e^{j2\pi kn/K}$, i.e. $X[n] \xleftrightarrow{\mathcal{F}} x[k]$.

Property-1: For any K, if $X[n] = a$, $\forall n$, where $a \in \mathbb{C}$ and $\mathbb{C}$ is the field of complex numbers, then $x[k] = a\delta[k]$ where $\delta[k]$ is a discrete unit impulse function, and vice versa.

Property-2: Assume that $K = ML_0$ for M=1, 2, ….

$$\text{If} \quad X[n] = \begin{cases} a, & n = mM; m = 0, \ldots, L_0 - 1; a \in \mathbb{C} \\ 0, & \text{elsewhere,} \end{cases}$$

$$\text{then} \quad x[k] = \begin{cases} aL_0/K, & k = nL_0; n = 0, \ldots, M - 1 \\ 0, & \text{elsewhere,} \end{cases}$$

and vice versa.

Property-3:

$$X[(n-l)_K] \xleftrightarrow{\mathcal{F}} e^{j2\pi lk/K}x[k]$$

where $(\cdot)_K$ denotes modulo-K operation, hence representing a cyclically-shifted version. Its dual form is given by

$$x[(k-m)_K] \xleftrightarrow{\mathcal{F}} e^{-j2\pi mn/K}X[n].$$

This section derives training signals that are optimal in terms of their robustness to frequency offsets among the optimal training signals presented in [10]. In the presence of a normalized (by the sub-carrier spacing) carrier frequency offset v, the received vector in (2) becomes

$$\boldsymbol{r} = \boldsymbol{W}(v)\ \boldsymbol{S}\boldsymbol{h} + \boldsymbol{n} \qquad (12)$$

where $\boldsymbol{W}(v) = diag\{\boldsymbol{W}_0(v)\ ,\ e^{j2\pi v(K+N_g)/K}\boldsymbol{W}_0(v), \ldots, e^{j2\pi v(Q-1)(K+N_g)/K}\boldsymbol{W}_0(v)\}$, $\boldsymbol{W}_0(v) = diag\{1,\ e^{j2\pi v/K}, e^{j2\pi 2v/K}, \ldots, e^{j2\pi(K-1)v/K}\}$. We simply consider the least-squares channel estimator for the sake of complexity and robustness [1]. The corresponding channel estimate obtained using an optimal training signal from [10] is

$$\begin{aligned} \hat{\boldsymbol{h}} &= \frac{1}{E_{\mathrm{av}}}\boldsymbol{S}^H\boldsymbol{r} \qquad (13) \\ &= \boldsymbol{h} - \frac{\{\boldsymbol{S}^H(\boldsymbol{I}-\boldsymbol{W}(v))\boldsymbol{S}\boldsymbol{h} - \boldsymbol{S}^H\boldsymbol{n}\}}{E_{\mathrm{av}}} \equiv \boldsymbol{h} - \Delta_{\boldsymbol{h}}. \qquad (14) \end{aligned}$$

Define the normalized MSE as

$$\begin{aligned} \mathrm{NMSE} &= \frac{\mathrm{MSE}}{LN_{\mathrm{Tx}}} = \frac{\mathrm{E}[\Delta_{\boldsymbol{h}}{}^{\mathrm{H}}\Delta_{\boldsymbol{h}}]}{LN_{\mathrm{Tx}}} \\ &= \frac{\sigma_n^2}{E_{\mathrm{av}}} + \frac{\mathrm{Tr}[\boldsymbol{S}^H(\boldsymbol{I}-\boldsymbol{W}(v))\boldsymbol{S}\boldsymbol{C_h}\boldsymbol{S}^H(\boldsymbol{I}-\boldsymbol{W}(v))^H\boldsymbol{S}]}{LN_{\mathrm{Tx}}E_{\mathrm{av}}^2} \qquad (15) \\ &\equiv \mathrm{NMSE}_0 + \Delta_{\mathrm{NMSE}}. \qquad (16) \end{aligned}$$

In (15), the first term is the NMSE obtained without any frequency offset which is the same for all optimal training signals from [10] and the second term is the extra NMSE caused by the frequency offset v.

We will find the best training signal matrices $\boldsymbol{S}^\star$ which give the minimum extra NMSE, i.e.,

$$\boldsymbol{S}^\star = \arg\min_{\boldsymbol{S}} \mathrm{Tr}[\boldsymbol{S}^H\boldsymbol{V}(v)\boldsymbol{S}\boldsymbol{C_h}\boldsymbol{S}^H\boldsymbol{V}(v)^H\boldsymbol{S}] \qquad (17)$$

where $\boldsymbol{V}(v) \equiv (\boldsymbol{I} - \boldsymbol{W}(v))$ and $\boldsymbol{S}$ is constrained to be circulant as described in Section II.

Since $\boldsymbol{S}^H\boldsymbol{V}(v)\boldsymbol{S}\boldsymbol{C_h}\boldsymbol{S}^H\boldsymbol{V}(v)^H\boldsymbol{S}$ is in the form of $\boldsymbol{G}\boldsymbol{G}^H$, it is a Hermitian positive semi-definite matrix. Let $\boldsymbol{X} = \boldsymbol{G}\boldsymbol{G}^H = \boldsymbol{S}^H\boldsymbol{V}(v)\boldsymbol{S}\boldsymbol{C_h}\boldsymbol{S}^H\boldsymbol{V}(v)^H\boldsymbol{S}$. Define $\boldsymbol{Y}$=$\boldsymbol{X} + \boldsymbol{\Lambda}_{\boldsymbol{X_d}} + b\boldsymbol{I}$ where $\boldsymbol{X_d}$ is from the subset of $\boldsymbol{X}$ which are diagonal, $\boldsymbol{X_d} + \boldsymbol{\Lambda}_{\boldsymbol{X_d}}$= $a\boldsymbol{I}$, and $a, b > 0$. Consider groups of $\boldsymbol{Y}$ with the same determinant $(a+b)^{N_{\mathrm{Tx}}L}$. Using the arithmetic-geometric mean inequality [2], we conclude that Tr[$\boldsymbol{Y}$] and hence Tr[$\boldsymbol{X}$] will be minimum for $\boldsymbol{X} = \boldsymbol{X_d}$ within each group. Hence, we just need to consider diagonal matrices $\boldsymbol{X_d}$. After obtaining a diagonal matrix with minimum trace, denoted by $\boldsymbol{X_d^\star}$, we can find other non-diagonal matrices $\boldsymbol{X}^\star$ with the same minimum trace by using the relation Tr[$\boldsymbol{P}\boldsymbol{X_d^\star}\boldsymbol{P}^H$] = Tr[$\boldsymbol{X_d^\star}$] where $\boldsymbol{P}$ is a unitary matrix.

We assume that channels of different transmit and receive antenna pairs are independent and have the same power delay profile. Each channel is assumed to have a diagonal correlation matrix $\boldsymbol{C_{h_0}}$ = diag$\{\sigma_0^2, \sigma_1^2, \ldots, \sigma_{L-1}^2\}$. Since $\boldsymbol{C_h}$ is diagonal, $\boldsymbol{X}$ will also be diagonal when $\boldsymbol{S}^H\boldsymbol{V}(v)\boldsymbol{S}$ is diagonal. Since $\boldsymbol{S}$ is circulant, $\boldsymbol{S}^H\boldsymbol{V}(v)\boldsymbol{S}$ will be diagonal if and only if

$$s_{k,q}[n] = \sum_{i=0}^{d_{k,q}} A_{k,q,i}\delta[n - l_{k,q,i}],\ k = 0,\ \ldots,\ N_{\mathrm{Tx}} - 1 \qquad (18)$$

$$\sum_{k=0}^{N_{\mathrm{Tx}}-1} (1 + d_{k,q}) \leq \frac{N}{L} \qquad (19)$$

where for each q, $\{l_{k,q,i}\ :\ \forall k, i\}$ are any permutation of $\{m_p\}$ with $m_{p+1} - m_p \geq L$ and $0 \leq m_p \leq K - 1$. To satisfy Condition (6), each transmit antenna must have the same total transmitted training energy [10] which, in turn, implies that

$$\sum_{q=0}^{Q-1}\sum_{i=0}^{d_{k,q}} |A_{k,q,i}|^2 = E_{\mathrm{av}},\ \forall k. \qquad (20)$$

---

[1] For other estimators such as MMSE or Bayesian estimators which utilize some statistical properties, sensitivity to model mismatches could be an issue.

[2] For positive numbers λ_i, $\prod_{i=1}^{N}\lambda_i \leq (\frac{1}{N}\sum_{i=1}^{N}\lambda_i)^N$ and the equality holds if and only if all λ_i are equal.

Then, using (18), we obtain a diagonal matrix $(\boldsymbol{S}^H\boldsymbol{V}(v)\boldsymbol{S})$ with elements given by

$$[\boldsymbol{S}^H\boldsymbol{V}(v)\boldsymbol{S}]_{l,l} = \sum_{q=0}^{Q-1}\sum_{i=0}^{d_{\lfloor l/L\rfloor,q}} |A_{\lfloor l/L\rfloor,q,i}|^2 V_{l_{\lfloor l/L\rfloor,q,i}+(K+N_g)q+(l)_L} \tag{21}$$

where $\lfloor x\rfloor$ denotes the maximum integer less than or equal to x and

$$\mathrm{Tr}[\boldsymbol{X_d}] = \sum_{k=0}^{N_{\mathrm{Tx}}-1}\sum_{m=0}^{L-1}\sigma_m^2\left|\sum_{q=0}^{Q-1}\sum_{i=0}^{d_{k,q}}|A_{k,q,i}|^2 V_{m+l_{k,q,i}+(K+N_g)q}\right|^2 \tag{22}$$

where $V_{l_{k,q,i}} \in \{V_l :\ l = 0,\ \ldots,\ QK-1\}$ is associated with the i-th non-zero training sample of the q-th symbol for the k-th antenna. The best $\boldsymbol{X_d}$ is defined by

$$\boldsymbol{X}_{\boldsymbol{d}}^{\star} = \arg\min_{\boldsymbol{X_d}}\mathrm{Tr}[\boldsymbol{X_d}]. \tag{23}$$

However, since $\{V_l\}$ depend on v, there is no single $\boldsymbol{S}^\star$ which remains optimum for all v and $\{\sigma_i^2\}$.

In practical systems, frequency offset estimation and compensation are typically performed before channel estimation. Hence, during channel estimation, the residual frequency offset is usually very small. For very small values of v, we have

$$\begin{aligned} V_l &= 1-e^{j2\pi k_l v/K} \simeq -j2\pi k_l v/K && (24)\\ \text{where}\quad k_l &= \lfloor l/K\rfloor N_g + l. && (25)\end{aligned}$$

It can be concluded from (24) and (22) that under the same total training signal energy constraint, using one training symbol (i.e., $Q=1$) is more robust to frequency offsets than using multiple training symbols. Hence, in the following, we just consider $Q=1$ case and the index q will be omitted. Note that $\boldsymbol{X}_{\boldsymbol{d}}^{\star}$ is completely determined by the best $\{|A_{k,i}|\}$ and $\{l_{k,i}\}$, denoted by $\{|A_{k,i}^\star|\}$ and $\{l_{k,i}^\star\}$, respectively. In this case, we obtain from (23)(22) using (24) that

$$\begin{aligned} |A_{k,i}^\star| &= \begin{cases}\sqrt{E_{\mathrm{av}}} & ,\ \text{if } i=0\\ 0 & ,\ \text{if } i\neq 0\end{cases} && (26)\\ l_{k,0}^\star &\in \{0,\ L,\ 2L,\ \ldots,(N_{\mathrm{Tx}}-1)L\} && (27)\\ \text{and}\quad & l_{k,0}^\star \neq l_{m,0}^\star \text{ if } k\neq m. \end{aligned}$$

The corresponding minimum extra NMSE is

$$\begin{aligned} (\Delta_{\mathrm{NMSE}})_{\min} &= \frac{1}{LN_{\mathrm{Tx}}}\sum_{m=0}^{N_{\mathrm{Tx}}-1}\sum_{l=0}^{L-1}\sigma_l^2|V_{mL+l}|^2 && (28)\\ &\simeq \frac{1}{LN_{\mathrm{Tx}}}(2\pi v/K)^2\sum_{m=0}^{N_{\mathrm{Tx}}-1}\sum_{l=0}^{L-1}\sigma_l^2 k_{mL+l}^2. && (29)\end{aligned}$$

Hence, the optimal training signals are given by

$$\begin{aligned} &\{|s_k^\star[n]| :\ k=0,\ 1,\ \ldots,N_{\mathrm{Tx}}-1\} =\\ &\{\sqrt{E_{\mathrm{av}}}\delta[n-mL] : m=0,\ 1,\ \ldots,N_{\mathrm{Tx}}-1\} \qquad (30)\end{aligned}$$

and the corresponding optimal pilot tones are

$$\begin{aligned} &\{c_k^\star[n] :\ k=0,\ 1,\ \ldots,N_{\mathrm{Tx}}-1\} =\\ &\{\sqrt{E_{\mathrm{av}}}e^{j\phi_m}e^{-j2\pi mnL/K} :\ m=0,\ 1,\ \ldots,\ N_{\mathrm{Tx}}-1\} \qquad (31)\end{aligned}$$

where $\{\phi_m\}$ are arbitrary phases. The above pilot allocation is of CDM(F) type over all sub-carriers.

Now, consider other training signal matrices, denoted by $\bar{\boldsymbol{S}}$, which give the same minimum trace for $\boldsymbol{X}$ as $\boldsymbol{S}^\star$ does. Using the relation $\mathrm{Tr}[\boldsymbol{P}\boldsymbol{X}_{\boldsymbol{d}}^{\star}\boldsymbol{P}^H] = \mathrm{Tr}[\boldsymbol{X}_{\boldsymbol{d}}^{\star}]$ where $\boldsymbol{P}$ is a unitary matrix, we obtain

$$\begin{aligned} \bar{\boldsymbol{S}} &= \boldsymbol{S}^\star\boldsymbol{P} && (32)\\ \boldsymbol{P}^H\boldsymbol{C_h}\boldsymbol{P} &= \boldsymbol{C_h}. && (33)\end{aligned}$$

Denote the $L\times L$ sub-matrices of the $N_{\mathrm{Tx}}L\times N_{\mathrm{Tx}}L$ matrix $\boldsymbol{P}$ by $\boldsymbol{P}_{k,l}$ for $k,\ l = 0,\ 1,\ \ldots,\ N_{\mathrm{Tx}}-1$. Then (33) can be expressed as

$$\begin{aligned} \sum_{k=0}^{N_{\mathrm{Tx}}-1}\boldsymbol{P}_{k,l}^H\boldsymbol{C_{h_0}}\boldsymbol{P}_{k,l} &= \boldsymbol{C_{h_0}}, \forall l && (34)\\ \sum_{k=0}^{N_{\mathrm{Tx}}-1}\boldsymbol{P}_{k,l}^H\boldsymbol{C_{h_0}}\boldsymbol{P}_{k,m} &= \boldsymbol{0},\ \forall l\neq m. && (35)\end{aligned}$$

Since $\boldsymbol{C_{h_0}}$ is diagonal, $\boldsymbol{P}_{k,l}$ must be diagonal for (35) to hold for arbitrary diagonal $\boldsymbol{C_{h_0}}$. Since $\boldsymbol{S}^\star$ and $\bar{\boldsymbol{S}}$ are circulant, (32) implies that $\boldsymbol{P}$ and hence $\boldsymbol{P}_{k,l}$ are circulant. Hence, we obtain

$$\boldsymbol{P}_{k,l} = a_{k,l}\boldsymbol{I},\ a_{k,l}\in\mathbb{C}. \tag{36}$$

Since $\boldsymbol{P}^H\boldsymbol{P}=\boldsymbol{I}$, we have

$$\sum_{k=0}^{N_{\mathrm{Tx}}-1}|a_{k,l}|^2 = 1,\ \sum_{k=0}^{N_{\mathrm{Tx}}-1}a_{k,l}\ a_{k,m}^* = 0,\ \forall l\neq m. \tag{37}$$

In finding $\bar{\boldsymbol{S}}$, we will simply use $\boldsymbol{S}^\star$=$\sqrt{E_{\mathrm{av}}}$ $[\boldsymbol{I}_{N_{\mathrm{Tx}}L}, \boldsymbol{0}_{(K-N_{\mathrm{Tx}}L)\times N_{\mathrm{Tx}}L}]^T$ because any permutation of training signals over the antennas does not affect the performance. Then (32) becomes

$$\bar{\boldsymbol{S}} = \sqrt{E_{\mathrm{av}}}\ [\boldsymbol{P}, \boldsymbol{0}_{K-N_{\mathrm{Tx}}L\times N_{\mathrm{Tx}}L}]^T. \tag{38}$$

Note that the optimal training signal matrices $\boldsymbol{S}$ presented in [10] satisfy $\boldsymbol{S}^H\boldsymbol{S} = E_{\mathrm{av}}\boldsymbol{I}$ and we also have $\bar{\boldsymbol{S}}^H\bar{\boldsymbol{S}} = E_{\mathrm{av}}\boldsymbol{I}$. Hence, from (38) together with (36) and (37), we obtain the following optimal training signals:

$$\bar{s}_k[n] = \sqrt{E_{\mathrm{av}}}\sum_{l=0}^{N_{\mathrm{Tx}}-1}a_{k,l}\ \delta[n-lL],\ a_{k,l}\in\mathbb{C}. \tag{39}$$

Recall that $\bar{s}_k[n]$ must also satisfy the relationship $\bar{s}_k[n]$ = $\sum_{m=0}^{K-1}c_k[m]e^{j2\pi mn/K}$. For $K > N_{\mathrm{Tx}}L$, if we restrict ourselves to the constant-modulus optimal pilot tones of [10], the obtained training signals $\{\bar{s}_k[n] :\ \forall k\}$ will be any permutation of $\{\sqrt{E_{\mathrm{av}}}a_{m,m}\ \delta[n-mL] :\ m = 0,\ 1,\ \ldots,\ N_{\mathrm{Tx}}-1\}$ which are already included in the training signals corresponding to $\boldsymbol{S}^\star$. There are no non-diagonal matrices $\boldsymbol{X}$ with the same minimum trace as $\boldsymbol{X_d}^\star$ for $K > N_{\mathrm{Tx}}L$ and constant-modulus pilot tones.

For $K = LN_{\mathrm{Tx}}$, by using DFT Properties 2 and 3, the optimal pilot tones corresponding to (39) can be expressed as

$$c_k^\star[n] = \sum_{m=0}^{N_{\mathrm{Tx}}-1}\sum_{l=0}^{L-1}\beta_{k,m}\delta[n-lN_{\mathrm{Tx}}-m] \tag{40}$$

where $\beta_{k,m}\in\mathbb{C}$. For constant-modulus pilot tones, $|\beta_{k,m}|$ is either zero or a constant. We conclude from (40) that for $K = N_{\mathrm{Tx}}L_0$, an additional condition for the optimal training signals in the presence of frequency offset is:

Condition-C : For each transmit antenna k, the optimal pilot tone symbols $c_k^\star[n-lN_{\mathrm{Tx}}-m]$ for different l are the same.

Now we will relate (40) to the pilot allocations presented in [10]. If $|\beta_{k,m}|$ is non-zero for all m and k, the corresponding allocation is CDM(F). In addition to Condition-C, $\beta_{k,m}$

must satisfy the requirements for CDM allocation as well as the equal-energy requirement from [10]. In this case, these requirements are given by

$$\mathrm{CDM(F)}: \beta_{k,m} = \beta_{0,m} e^{-j2\pi mk/N_{\mathrm{Tx}}} e^{j\phi_k}, \quad |\beta_{k,m}| = \sqrt{E_{\mathrm{av}}} \tag{41}$$

where ϕ_k is an arbitrary phase. If we consider a pure FDM pilot tone allocation, then (40) reduces to the following optimal pilot tones:

$$\mathrm{FDM}: \; c_k^\star[n] = \sum_{l=0}^{L-1} \beta_k \delta[n - lN_{\mathrm{Tx}} - k], \quad |\beta_k| = \sqrt{N_{\mathrm{Tx}} E_{\mathrm{av}}}. \tag{42}$$

Suppose we use a combination of FDM and CDM(F) allocation, say 2-FDM + $\frac{N_{\mathrm{Tx}}}{2}$-CDM(F). Then, the optimal pilot tones for the $N_{\mathrm{Tx}}/2$ antennas in the first FDM group and those in the second FDM group, respectively, are given by

$$\text{2-FDM} + \frac{N_{\mathrm{Tx}}}{2}\text{-CDM(F)}:$$

$$c_k^\star[n] = \sum_{m=0}^{\frac{N_{\mathrm{Tx}}}{2}-1} \sum_{l=0}^{L-1} \beta_{k,m} \, \delta[n - lN_{\mathrm{Tx}} - m], \tag{43}$$

$$k = 0, \ldots, \frac{N_{\mathrm{Tx}}}{2} - 1,$$

$$c_k^\star[n] = \sum_{m=\frac{N_{\mathrm{Tx}}}{2}}^{N_{\mathrm{Tx}}-1} \sum_{l=0}^{L-1} \beta_{k,m} \, \delta[n - lN_{\mathrm{Tx}} - m], \tag{44}$$

$$k = \frac{N_{\mathrm{Tx}}}{2}, \ldots, N_{\mathrm{Tx}} - 1.$$

Within each FDM group, $\beta_{k,m}$ must satisfy the requirements for CDM allocation as well as the equal-energy requirement from [10]. In this case, these requirements are given by

$$\beta_{k,m} = \beta_{0,m} e^{\frac{-j4\pi mk}{N_{\mathrm{Tx}}}} e^{j\phi_k}, \quad k, m = 0, \ldots, \frac{N_{\mathrm{Tx}}}{2} - 1 \tag{45}$$

$$\beta_{k,m} = \beta_{\frac{N_{\mathrm{Tx}}}{2},m} e^{\frac{-j4\pi mk}{N_{\mathrm{Tx}}}} e^{j\phi_k}, \tag{46}$$

$$k, m = \frac{N_{\mathrm{Tx}}}{2}, \ldots, N_{\mathrm{Tx}} - 1$$

$$|\beta_{k,m}| = \sqrt{2E_{\mathrm{av}}} \tag{47}$$

where ϕ_k is an arbitrary phase and $\phi_0 = \phi_{N_{\mathrm{Tx}}/2} = 0$. To summarize, for $K = N_{\mathrm{Tx}}L$, the mapping of $\boldsymbol{S}^\star$ to $\bar{\boldsymbol{S}}$ gives us the most robust optimal training signals with any pilot allocation presented in [10] but with the additional constraint of Condition-C while for $K > N_{\mathrm{Tx}}L$, the optimal pilot tones are given by (31).

IV. EXAMPLES, DISCUSSIONS, AND SIMULATION RESULTS

We numerically evaluated the extra NMSEs for all optimal training signals and confirmed the minimum extra NMSE of the optimal training signals presented in the previous section. For the convenience of numerical evaluation and presentation, $K = 8$ and $L = 2$ will be assumed unless stated otherwise.

Some representative examples of the optimal pilot tone vectors in the absence/presence of frequency offset are presented in Table I/II for $N_{\mathrm{Tx}} = 2$, ($K > LN_{\mathrm{Tx}}$) and Tables III/IV for $N_{\mathrm{Tx}} = 4$, ($K = LN_{\mathrm{Tx}}$). $\{\alpha_i\}$ are constant modulus symbols and $\{\phi_i\}$ are arbitrary phases. One can observe that Condition-C for the optimal training signals is satisfied.

Table V lists the maximum and minimum values of the extra NMSEs of the optimal training signals from [10] obtained from numerical evaluation. The minimum values are achieved by the proposed optimal training signals. The following remarks are in order:

1) The minimum extra NMSE increases with increasing N_{Tx} while the maximum extra NMSE decreases with increasing N_{Tx}.
2) The ratio of the maximum to minimum extra NMSE reflects the importance of using the optimal training signals. In Table V, this ratio for the SISO OFDM system is around 100 while for MIMO OFDM systems with $K = LN_{\mathrm{Tx}}$, it is quite small but for a more practical situation where $K > LN_{\mathrm{Tx}}$, it is around 10. Hence, for moderate and high values of SNR and residual frequency offsets where the extra NMSE dominates the performance degradation, the advantage of using the optimal training signals derived in this paper is significant.
3) The extra NMSE is approximately proportional to v^2, as evident from (29).

The minimum NMSEs achieved with the proposed optimal training signals are plotted in Figure 1 for different values ofN_{Tx}, v and SNR ($= E_{\mathrm{av}}N_{\mathrm{Tx}}/(QK\sigma_n^2)$). At moderate to high SNR, $v = 0.1$ introduces a significant degradation in channel estimation while $v = 0.01$ or less causes insignificant degradation. Based on Figure 1 together with (15) and (29), the following remarks are in order:

1) In the absence of frequency offset, the NMSE depends only on $\frac{E_{\mathrm{av}}}{\sigma_n^2}$ regardless of N_{Tx}.
2) In the presence of frequency offset, the NMSE depends on $\frac{E_{\mathrm{av}}}{\sigma_n^2}$, $\boldsymbol{C_h}$, v, and N_{Tx}.
3) A larger N_{Tx} results in a larger NMSE in the presence of frequency offset.

For validation purposes, we simulated the NMSE performance in the presence of frequency offsets for the proposed optimal training signal and other training signals which were optimal in the absence of frequency offsets. The results are presented in Figures 2 and 3 for $N_{\mathrm{Tx}} = 2$ and 8, respectively. For $K > N_{\mathrm{Tx}}L$ in Figure 2, training#1 represents an optimal training signal, training#2 employs an FDM allocation with L tones (a pseudo-noise (PN) sequence), training#3 is of a CDM(F) allocation over $2L$ sub-carriers (a PN sequence), and training#4 uses a CDM(F) allocation over all sub-carriers (a PN sequence). For $K = N_{\mathrm{Tx}}L$ in Figure 3, training#1 represents an optimal training signal with FDM allocation, training#2 employs an FDM pilot allocation with L tones (a PN sequence), training#3 represents the optimal training signals with CDM(F) allocation, and training#4 uses a CDM(F) allocation over all sub-carriers (a PN sequence). The MSE advantages of the proposed optimal training signals are more significant for a larger v, a higher SNR, and a smaller N_{Tx}.

V. CONCLUSIONS

We presented optimal training signals for MIMO OFDM channel estimation in the presence of frequency offsets. The effect of frequency offset on channel estimation is more severe for a larger number of transmit antennas. Under the same total training energy constraint, using only one OFDM training

TABLE II

OPTIMAL PILOT TONE VECTORS IN THE PRESENCE OF FREQUENCY OFFSET FOR $K > N_{\mathrm{Tx}}L$, ($K = 8, N_{\mathrm{Tx}} = 2, L = 2$)

Sub-carrier Index	CDM(F)	
	Ant. 0	Ant. 1
0	α_1	$\alpha_1 e^{j\phi_1}$
1	α_1	$\alpha_1 e^{-j\pi/2} e^{j\phi_1}$
2	α_1	$\alpha_1 e^{-j2\pi/2} e^{j\phi_1}$
3	α_1	$\alpha_1 e^{-j3\pi/2} e^{j\phi_1}$
4	α_1	$\alpha_1 e^{-j4\pi/2} e^{j\phi_1}$
5	α_1	$\alpha_1 e^{-j5\pi/2} e^{j\phi_1}$
6	α_1	$\alpha_1 e^{-j6\pi/2} e^{j\phi_1}$
7	α_1	$\alpha_1 e^{-j7\pi/2} e^{j\phi_1}$

TABLE V

THE RANGE OF FREQUENCY-OFFSET-INDUCED EXTRA NMSE FOR THE TRAINING SIGNALS WHICH ARE OPTIMAL FOR $v = 0$ ($K = 8$, $L = 2$)

N_{Tx}	Δ_{NMSE}	
	$v = 0.01$	$v = 0.1$
1	$[1.03 \times 10^{-5},\ 1.24 \times 10^{-3}]$	$[1.03 \times 10^{-3},\ 1.22 \times 10^{-1}]$
2	$[9.26 \times 10^{-5},\ 9.15 \times 10^{-4}]$	$[9.23 \times 10^{-3},\ 8.99 \times 10^{-2}]$
4	$[5.04 \times 10^{-4},\ 5.76 \times 10^{-4}]$	$[4.96 \times 10^{-2},\ 5.65 \times 10^{-2}]$

symbol for channel estimation is more robust to frequency offset than using multiple training symbols. The reduction in the frequency-offset-induced extra MSE achieved by the optimal training signals can be quite significant (an order of magnitude) at moderate and high values of SNR and frequency offsets and especially for a small number of transmit antennas.

ACKNOWLEDGMENT

The work of H. Minn was supported in part by the Erik Jonsson School Research Excellence Initiative, the University of Texas at Dallas. The work of N. Al-Dhahir was supported in part by the Texas Advanced Technology Program (ATP) project no. 009741-0023-2003.

REFERENCES

[1] C. Tellambura, M.G. Parker, Y.J. Guo, S.J. Shepherd and S.K. Barton, "Optimal sequences for channel estimation using discrete Fourier transform techniques," *IEEE Trans. Commun.*, Feb. 1999, pp. 230-238.

[2] J.C.L. Ng, K.B. Letaief, and R.D. Murch, "Complex optimal sequences with constant magnitude for fast channel estimation initialization," *IEEE Trans. Commun.,* Vol. 46, No. 3, Mar. 1998, pp. 305-308.

[3] S. Adireddy, L. Tong, and H. Viswanathan, "Optimal placement of training for frequency-selective block-fading channels," *IEEE Trans. Information Theory,* Vol. 48, No. 8, Aug. 2002, pp. 2338-2353.

[4] M. Dong and L. Tong, "Optimal design and placement of pilot symbols for channel estimation," *IEEE Trans. Signal Processing*, Vol. 50, No. 12, Dec. 2002, pp. 3055-3069.

[5] S. Ohno and G. B. Giannakis, "Optimal training and redundant precoding for block transmissions with application to wireless OFDM," *IEEE Trans. Comun.,* Vol. 50, No. 12, Dec. 2002, pp. 2113-2123.

[6] X. Ma, G.B. Giannakis and S. Ohno, "Optimal training for block transmissions over doubly selective wireless fading channels," *IEEE Trans. Signal Processing*, Vol. 51, No. 5, May 2003, pp. 1351-1365.

[7] C. Fragouli, N. Al-Dhahir, and W. Turin, "Training-based channel estimation for multiple antenna broadband transmissions," *IEEE Trans. Wireless Commun.,* Vol. 2, No. 2, Mar. 2003, pp. 384-391.

[8] I. Barhumi, G. Leus, and M. Moonen, "Optimal training design for MIMO OFDM systems in mobile wireless channels," *IEEE Trans. Signal Processing,* Vol. 51, No. 6, June 2003, pp. 1615-1624.

[9] Y. Li, "Simplified channel estimation for OFDM systems with multiple transmit antennas," *IEEE Trans. Wireless Commun.,* Jan. 2002, pp. 67-75.

[10] H. Minn and N. Al-Dhahir, "Optimal training signals for MIMO OFDM channel estimation," *IEEE Globecom*, Nov. 29 - Dec. 3, 2004.

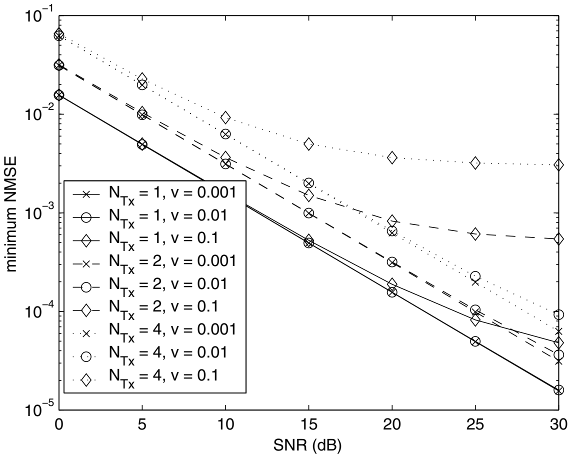

Fig. 1. The minimum NMSE for a MIMO OFDM system with $K = 64$, $N_g = 16$, $Q = 1$ in an 8-tap multipath Rayleigh fading channel with an exponential power delay profile

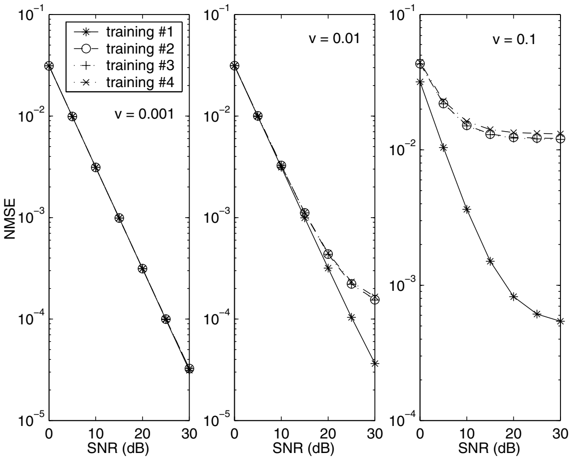

Fig. 2. The NMSE comparison of several training signals for a MIMO OFDM system with $N_{\mathrm{Tx}} = 2$, $K = 64$, $N_g = 16$, $Q = 1$ in an 8-tap multipath Rayleigh fading channel with an exponential power delay profile

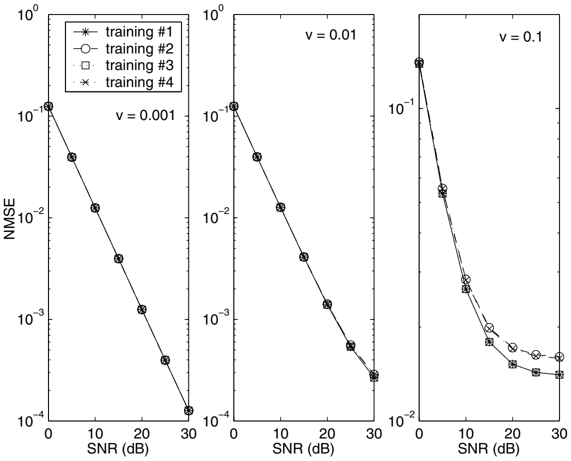

Fig. 3. The NMSE comparison of several training signals for a MIMO OFDM system with $N_{\mathrm{Tx}} = 8$, $K = 64$, $N_g = 16$, $Q = 1$ in an 8-tap multipath Rayleigh fading channel with an exponential power delay profile

TABLE I

OPTIMAL PILOT TONE VECTORS IN THE ABSENCE OF FREQUENCY OFFSET FOR A MIMO OFDM SYSTEM WITH $K > N_{\mathrm{Tx}}L$, $(K = 8, N_{\mathrm{Tx}} = 2, L = 2)$

Sub-carrier Index	Pilot Allocation							
	FDM		FDM		CDM(F)		CDM(F), $1 \le m \le 3$	
	Ant. 0	Ant. 1	Ant. 0	Ant. 1	Ant. 0	Ant. 1	Ant. 0	Ant. 1
0	$2\alpha_1$	0	$\sqrt{2}\alpha_1$	0	$\sqrt{2}\alpha_1$	$\sqrt{2}\alpha_1 e^{j\phi_1}$	α_1	$\alpha_1 e^{j\phi_1}$
1	0	0	0	$\sqrt{2}\alpha_5$	0	0	α_2	$\alpha_2 e^{-jm\pi/2} e^{j\phi_1}$
2	0	$2\alpha_3$	$\sqrt{2}\alpha_2$	0	0	0	α_3	$\alpha_3 e^{-j2m\pi/2} e^{j\phi_1}$
3	0	0	0	$\sqrt{2}\alpha_6$	$\sqrt{2}\alpha_2$	$-\sqrt{2}\alpha_2 e^{j\phi_1}$	α_4	$\alpha_4 e^{-j3m\pi/2} e^{j\phi_1}$
4	$2\alpha_2$	0	$\sqrt{2}\alpha_3$	0	$\sqrt{2}\alpha_3$	$\sqrt{2}\alpha_3 e^{j\phi_1}$	α_5	$\alpha_5 e^{-j4m\pi/2} e^{j\phi_1}$
5	0	0	0	$\sqrt{2}\alpha_7$	0	0	α_6	$\alpha_6 e^{-j5m\pi/2} e^{j\phi_1}$
6	0	$2\alpha_4$	$\sqrt{2}\alpha_4$	0	0	0	α_7	$\alpha_7 e^{-j6m\pi/2} e^{j\phi_1}$
7	0	0	0	$\sqrt{2}\alpha_8$	$\sqrt{2}\alpha_4$	$-\sqrt{2}\alpha_4 e^{j\phi_1}$	α_8	$\alpha_8 e^{-j7m\pi/2} e^{j\phi_1}$

TABLE III

OPTIMAL PILOT TONE VECTORS IN THE ABSENCE OF FREQUENCY OFFSET FOR A MIMO OFDM SYSTEM WITH $K = N_{\mathrm{Tx}}L$,$(K = 8, N_{\mathrm{Tx}} = 4, L = 2)$

Pilot Allocation	Sub-carrier Index	Ant. 0	Ant. 1	Ant. 2	Ant. 3
CDM(F)	0	α_1	$\alpha_1 e^{j\phi_1}$	$\alpha_1 e^{j\phi_2}$	$\alpha_1 e^{j\phi_3}$
	1	α_2	$\alpha_2 e^{-j\pi/2} e^{j\phi_1}$	$\alpha_2 e^{-j2\pi/2} e^{j\phi_2}$	$\alpha_2 e^{-j3\pi/2} e^{j\phi_3}$
	2	α_3	$\alpha_3 e^{-j2\pi/2} e^{j\phi_1}$	$\alpha_3 e^{-j4\pi/2} e^{j\phi_2}$	$\alpha_3 e^{-j6\pi/2} e^{j\phi_3}$
	3	α_4	$\alpha_4 e^{-j3\pi/2} e^{j\phi_1}$	$\alpha_4 e^{-j6\pi/2} e^{j\phi_2}$	$\alpha_4 e^{-j9\pi/2} e^{j\phi_3}$
	4	α_5	$\alpha_5 e^{-j4\pi/2} e^{j\phi_1}$	$\alpha_5 e^{-j8\pi/2} e^{j\phi_2}$	$\alpha_5 e^{-j12\pi/2} e^{j\phi_3}$
	5	α_6	$\alpha_6 e^{-j5\pi/2} e^{j\phi_1}$	$\alpha_6 e^{-j10\pi/2} e^{j\phi_2}$	$\alpha_6 e^{-j15\pi/2} e^{j\phi_3}$
	6	α_7	$\alpha_7 e^{-j6\pi/2} e^{j\phi_1}$	$\alpha_7 e^{-j12\pi/2} e^{j\phi_2}$	$\alpha_7 e^{-j18\pi/2} e^{j\phi_3}$
	7	α_8	$\alpha_8 e^{-j7\pi/2} e^{j\phi_1}$	$\alpha_8 e^{-j14\pi/2} e^{j\phi_2}$	$\alpha_8 e^{-j21\pi/2} e^{j\phi_3}$
FDM	0	$2\alpha_1$	0	0	0
	1	0	$2\alpha_3$	0	0
	2	0	0	$2\alpha_5$	0
	3	0	0	0	$2\alpha_7$
	4	$2\alpha_2$	0	0	0
	5	0	$2\alpha_4$	0	0
	6	0	0	$2\alpha_6$	0
	7	0	0	0	$2\alpha_8$
2-FDM + 2-CDM(F)	0	$\sqrt{2}\alpha_1$	$\sqrt{2}\alpha_1 e^{j\phi_1}$	0	0
	1	0	0	$\sqrt{2}\alpha_5$	$\sqrt{2}\alpha_5 e^{j\phi_3}$
	2	0	0	$\sqrt{2}\alpha_6$	$-\sqrt{2}\alpha_6 e^{j\phi_3}$
	3	$\sqrt{2}\alpha_2$	$-\sqrt{2}\alpha_2 e^{j\phi_1}$	0	0
	4	$\sqrt{2}\alpha_3$	$\sqrt{2}\alpha_3 e^{j\phi_1}$	0	0
	5	0	0	$\sqrt{2}\alpha_7$	$\sqrt{2}\alpha_7 e^{j\phi_3}$
	6	0	0	$\sqrt{2}\alpha_8$	$-\sqrt{2}\alpha_8 e^{j\phi_3}$
	7	$\sqrt{2}\alpha_4$	$-\sqrt{2}\alpha_4 e^{j\phi_1}$	0	0

TABLE IV

OPTIMAL PILOT TONE VECTORS IN THE PRESENCE OF FREQUENCY OFFSET FOR A MIMO OFDM SYSTEM WITH $K = N_{\mathrm{Tx}}L$,$(K = 8, N_{\mathrm{Tx}} = 4, L = 2)$

Pilot Allocation	Sub-carrier Index	Ant. 0	Ant. 1	Ant. 2	Ant. 3
CDM(F)	0	α_1	$\alpha_1 e^{j\phi_1}$	$\alpha_1 e^{j\phi_2}$	$\alpha_1 e^{j\phi_3}$
	1	α_2	$\alpha_2 e^{-j\pi/2} e^{j\phi_1}$	$\alpha_2 e^{-j2\pi/2} e^{j\phi_2}$	$\alpha_2 e^{-j3\pi/2} e^{j\phi_3}$
	2	α_3	$\alpha_3 e^{-j2\pi/2} e^{j\phi_1}$	$\alpha_3 e^{-j4\pi/2} e^{j\phi_2}$	$\alpha_3 e^{-j6\pi/2} e^{j\phi_3}$
	3	α_4	$\alpha_4 e^{-j3\pi/2} e^{j\phi_1}$	$\alpha_4 e^{-j6\pi/2} e^{j\phi_2}$	$\alpha_4 e^{-j9\pi/2} e^{j\phi_3}$
	4	α_1	$\alpha_1 e^{-j4\pi/2} e^{j\phi_1}$	$\alpha_1 e^{-j8\pi/2} e^{j\phi_2}$	$\alpha_1 e^{-j12\pi/2} e^{j\phi_3}$
	5	α_2	$\alpha_2 e^{-j5\pi/2} e^{j\phi_1}$	$\alpha_2 e^{-j10\pi/2} e^{j\phi_2}$	$\alpha_2 e^{-j15\pi/2} e^{j\phi_3}$
	6	α_3	$\alpha_3 e^{-j6\pi/2} e^{j\phi_1}$	$\alpha_3 e^{-j12\pi/2} e^{j\phi_2}$	$\alpha_3 e^{-j18\pi/2} e^{j\phi_3}$
	7	α_4	$\alpha_4 e^{-j7\pi/2} e^{j\phi_1}$	$\alpha_4 e^{-j14\pi/2} e^{j\phi_2}$	$\alpha_4 e^{-j21\pi/2} e^{j\phi_3}$
FDM	0	$2\alpha_1$	0	0	0
	1	0	$2\alpha_2$	0	0
	2	0	0	$2\alpha_3$	0
	3	0	0	0	$2\alpha_4$
	4	$2\alpha_1$	0	0	0
	5	0	$2\alpha_2$	0	0
	6	0	0	$2\alpha_3$	0
	7	0	0	0	$2\alpha_4$
2-FDM + 2-CDM(F)	0	$\sqrt{2}\alpha_1$	$\sqrt{2}\alpha_1 e^{j\phi_1}$	0	0
	1	0	0	$\sqrt{2}\alpha_3$	$\sqrt{2}\alpha_3 e^{j\phi_3}$
	2	0	0	$\sqrt{2}\alpha_4$	$-\sqrt{2}\alpha_4 e^{j\phi_3}$
	3	$\sqrt{2}\alpha_2$	$-\sqrt{2}\alpha_2 e^{j\phi_1}$	0	0
	4	$\sqrt{2}\alpha_1$	$\sqrt{2}\alpha_1 e^{j\phi_1}$	0	0
	5	0	0	$\sqrt{2}\alpha_3$	$\sqrt{2}\alpha_3 e^{j\phi_3}$
	6	0	0	$\sqrt{2}\alpha_4$	$-\sqrt{2}\alpha_4 e^{j\phi_3}$
	7	$\sqrt{2}\alpha_2$	$-\sqrt{2}\alpha_2 e^{j\phi_1}$	0	0