Optimal Training Signals for MIMO OFDM Channel Estimation in the Presence of Frequency Offset and Phase Noise

Hlaing Minn, *Member, IEEE*, Naofal Al-Dhahir, *Senior Member, IEEE*, and Yinghui Li, *Student Member, IEEE*

***Abstract*—We develop robust mean-square error (MSE)-optimal training signal designs for multiple-input multiple-output orthogonal frequency-division multiplexing channel estimation with frequency offset and phase noise (PN), and present analytical and simulation results for the frequency-offset and PN effects on channel estimation. The proposed designs are more advantageous for moderate-to-high values of signal-to-noise ratio (SNR), residual frequency offset, and PN level. At SNR = 10 dB, the normalized MSE reductions of our proposed training signals at normalized frequency offset $|v| = 0.1, 0.5$ are about 9 and 19 dB, respectively, for one transmit antenna, and 6 and 11 dB for two transmit antennas.**

***Index Terms*—Channel estimation, frequency offset, multiple-input multiple-output (MIMO), orthogonal frequency-division multiplexing (OFDM), phase noise (PN), pilot design, training-signal design.**

I. Introduction

TRAINING-signal design for channel estimation is a well-studied problem for single-input single-output (SISO) single-carrier systems, but a relatively new one for multiple-input multiple-output (MIMO) systems [1]–[9]. To the best of our knowledge, all existing training-signal designs for channel estimation assume no frequency offset and phase noise (PN). In practice, frequency offset and PN are unavoidable due to nonideal oscillators. They cause a loss of orthogonality among the subcarriers which, in turn, seriously degrades the performance of orthogonal frequency-division multiplexing (OFDM) systems [10], [11]. Hence, frequency offset and PN estimation and compensation techniques are typically applied at the receiver. However, in practice, there will still be a nonzero residual frequency offset. In addition, PN compensation techniques (e.g., [12]) require channel estimates, and hence, obtaining robust channel estimates in the presence of PN is important. It is unclear how the existing optimal training signals behave in the presence of (residual) frequency offset and PN.

In this letter, we derive the optimal (in the mean-square error (MSE) sense) training signals for MIMO OFDM channel estimation which are the most robust to frequency offset and PN among the existing training signals. Our proposed training signal is a particular code-division multiplexing (CDM)-type pilot allocation given in (27). When the number of subcarriers K is L_0 times the number of transmit antennas, where L_0 is the smallest integer greater than or equal to the number of channel taps and K/L_0 is an integer, additional optimal training signals are given by the optimal (without frequency offset) pilot signals in [9] which satisfy the following additional condition: "...for each transmit antenna k, the optimal pilot-tone symbols $c_k^\star[n-lN_{\mathrm{Tx}}-m]$ for different l are the same." Our results show that: 1) MIMO OFDM systems are more sensitive to frequency offsets and PN than SISO OFDM systems; 2) robustness requires a certain correlation of the pilot tones; 3) under the same total training-signal energy constraint, using one training symbol is more robust than multiple training symbols; and 4) the performance improvement of the proposed optimal training signals becomes more significant for moderate-to-high values of signal-to-noise ratio (SNR), residual frequency offset, and PN level.

Note that we consider a quasi-static channel within the training symbol. For a time-varying channel within a symbol, intercarrier-interference cancellation techniques (e.g., [13], [14]) can be applied. We consider least-squares estimation of sample-spaced CIRs. Other channel-estimation methods (e.g., [15], which estimates delays, gains, and the number of channel paths) will also be affected by the frequency offset and PN, and optimal training signals for them need further investigation.

The rest of this letter is organized as follows. Section II describes the signal model. Section III presents the training signal designs in the presence of both frequency offset and PN. Simulation results and discussions are presented in Section IV, and the letter is concluded in Section V.

II. Signal Model

Consider a MIMO OFDM system where training signals from N_{Tx} 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 $\boldsymbol{C}_{n,q} = [c_{n,q}[0], \ldots, c_{n,q}[K-1]]^T$ be the pilot tones vector of the nth transmit antenna at the qth symbol interval, where K is the number of OFDM subcarriers and the superscript T denotes the transpose. Furthermore, let $\{s_{n,q}[k] : k = -N_g, \ldots, K-1\}$ be the corresponding time-domain complex baseband training samples, including $N_g(\geq L-1)$ cyclic prefix (CP) samples. Define $\boldsymbol{S}_n[q]$ as the training-signal matrix of size $K \times L$ for the nth transmit antenna at the qth symbol interval, whose elements

Paper approved by S. K. Wilson, the Editor for Multicarrier Modulation of the IEEE Communications Society. Manuscript received July 9, 2004; revised May 12, 2005 and November 11, 2005. The work of H. Minn and Y. Li was supported in part by the Erik Jonsson School Research Excellence Initiative, University of Texas at Dallas. The work of N. Al-Dhahir was supported in part by the Texas Advanced Technology Program under Contract 009741-0023-2003, and in part by the National Science Foundation under Contracts 04-30654 and 05-28010. This paper was presented in part at the IEEE WCNC, New Orleans, LA, March 2005, and in part at the ICICS, Bangkok, Thailand, December 2005.

The authors are with the Department of Electrical Engineering, University of Texas at Dallas, Richardson, TX 75083 USA (e-mail: hlaing.minn@utdallas.edu; aldhahir@utdallas.edu; yxl044000@utdallas.edu).

Digital Object Identifier 10.1109/TCOMM.2006.881358

are given by $[\boldsymbol{S}_n[q]]_{m,l} = s_{n,q}[m-l]$ for $m \in \{0,\ldots,K-1\}$ and $l \in \{0,\ldots,L-1\}$.

Let $\boldsymbol{h}_n$ denote the length-L CIR vector corresponding to the nth transmit antenna. After CP removal, denote the received vector of length K at the qth symbol interval by $\boldsymbol{r}_q$. In the presence of a frequency offset and PN, the received vector over the Q symbol intervals is given by

$$\boldsymbol{r} = \bar{\boldsymbol{W}}(v)\boldsymbol{S}h + \boldsymbol{n} \tag{1}$$

where $\boldsymbol{r} = \left[\boldsymbol{r}_0^T\, \boldsymbol{r}_1^T \ldots \boldsymbol{r}_{Q-1}^T\right]^T$

$$\boldsymbol{h} = \left[\boldsymbol{h}_0^T\, \boldsymbol{h}_1^T \ldots \boldsymbol{h}_{N_{\mathrm{Tx}}-1}^T\right]^T, \; \bar{\boldsymbol{W}}(v) = \boldsymbol{W}(v)\boldsymbol{\Phi} \tag{2}$$

$$[\boldsymbol{S}]_{k,m} = \boldsymbol{S}_m[k], \; m = 0,1,\ldots,N_{\mathrm{Tx}}-1, \quad k = 0,1,\ldots,Q-1 \tag{3}$$

$$\boldsymbol{W}(v) = \mathrm{diag}\Big\{\boldsymbol{W_0}(v),\; e^{\frac{j2\pi v(K+N_g)}{K}}\boldsymbol{W_0}(v),\ldots, e^{\frac{j2\pi v(Q-1)(K+N_g)}{K}}\boldsymbol{W_0}(v)\Big\} \tag{4}$$

$$\boldsymbol{W_0}(v) = \mathrm{diag}\left\{1, e^{\frac{j2\pi v}{K}}, e^{\frac{j2\pi 2v}{K}},\ldots, e^{\frac{j2\pi(K-1)v}{K}}\right\} \tag{5}$$

$$\boldsymbol{\Phi} = \mathrm{diag}\Big\{e^{j\phi_0},\ldots,e^{j\phi_{K-1}}, e^{j\phi_{K+N_g}},\ldots, e^{j\phi_{2K+N_g-1}},\ldots,e^{j\phi_{(Q-1)(K+N_g)}},\ldots, e^{j\phi_{(Q-1)(K+N_g)+K-1}}\Big\} \tag{6}$$

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, v is the frequency offset normalized by the subcarrier spacing, and $\{\phi_k\}$ are the PN samples. A continuous-time PN $\phi(t)$ is commonly modeled by a Wiener process with $E[\phi(t)] = 0$ and $E[(\phi(t+\tau)-\phi(t))^2] = 4\pi\beta|\tau|$, where β denotes the one-sided 3-dB linewidth of the Lorentzian power density spectrum of the oscillator [10]. In the discrete-time domain, the PN samples can be modeled by $\phi_n = \phi_{n-1} + \Delta\phi$, where $\Delta\phi$ is a sample of the white Gaussian process with variance σ_p^2, and ϕ_0 is uniformly distributed over $[-\pi,\pi)$. Note that $\sigma_p^2 = 4\pi\beta/(K\Delta_f)$, where Δ_f is the subcarrier spacing. Since ϕ_0 can be embedded in the CIR vector, it can simply be set to zero. We assume that channels of different transmit and receive antenna pairs are independent, and have the same power delay profile $\boldsymbol{C_{h_0}}$, i.e., $\boldsymbol{C_h} = \boldsymbol{I}_{N_{\mathrm{Tx}}} \otimes \boldsymbol{C_{h_0}}$, where $\otimes$ is the Kronecker product. Each channel is assumed to have a diagonal correlation matrix $\boldsymbol{C_{h_0}} = \mathrm{diag}\{\sigma_0^2, \sigma_1^2,\ldots,\sigma_{L-1}^2\}$.

III. Proposed Training-Signal Designs

In the absence of a frequency offset and PN, the optimality of the training signal in terms of minimizing the MSE of the least-squares channel estimate is achieved if and only if (iff)

$$\boldsymbol{S}^H\boldsymbol{S} = E_{\mathrm{av}}\boldsymbol{I} \tag{7}$$

$$\text{where} \quad E_{\mathrm{av}} = \frac{1}{N_{\mathrm{Tx}}}\sum_{n=0}^{N_{\mathrm{Tx}}-1}\sum_{q=0}^{Q-1}\sum_{k=0}^{K-1}|s_{n,q}[k]|^2. \tag{8}$$

This condition gives several optimal training-signal designs, as presented in [9]. However, in the presence of a frequency offset and PN, the normalized MSE (NMSE) of those training signals becomes

$$\begin{aligned}\mathrm{NMSE} &= \frac{\mathrm{MSE}}{LN_{\mathrm{Tx}}} = \frac{\sigma_n^2}{E_{\mathrm{av}}} + \frac{1}{LN_{\mathrm{Tx}}E_{\mathrm{av}}^2}E\left[\mathrm{Tr}[\boldsymbol{X}]\right] \\ &= \frac{\sigma_n^2}{E_{\mathrm{av}}} + \frac{1}{LN_{\mathrm{Tx}}E_{\mathrm{av}}^2}\mathrm{Tr}[\bar{\boldsymbol{X}}] \equiv \mathrm{NMSE}_0 + \Delta_{\mathrm{NMSE}}\end{aligned} \tag{9}$$

where the first (second) term is the NMSE (extra NMSE) obtained in the absence (presence) of frequency offset and PN, $\bar{\boldsymbol{X}} = E[\boldsymbol{X}]$, and

$$\boldsymbol{X} = \left[\boldsymbol{S}^H\left(\boldsymbol{I}-\bar{\boldsymbol{W}}(v)\right)\boldsymbol{S}\boldsymbol{C_h}\boldsymbol{S}^H\left(\boldsymbol{I}-\bar{\boldsymbol{W}}(v)\right)^H\boldsymbol{S}\right]. \tag{10}$$

The expectation in (9) is with respect to the PN, and is unnecessary if only the frequency offset is considered. We will investigate which training signals are the best (most robust to frequency offset and PN) among the optimal training signals presented in [9]. Equivalently, we will find the best training-signal matrices $\boldsymbol{S}^\star$ as

$$\boldsymbol{S}^\star = \arg\min_{\boldsymbol{S}}\Delta_{\mathrm{NMSE}} = \arg\min_{\boldsymbol{S}} E\left[\mathrm{Tr}[\boldsymbol{X}]\right] \tag{11}$$

where $\boldsymbol{S}$ is constrained to be circulant due to the CP.

First, we consider the training signal structures that minimize $\mathrm{Tr}[\boldsymbol{X}]$, regardless of the PN samples. The Hermitian positive semidefinite matrix $\boldsymbol{C}_h$ can be decomposed as $\boldsymbol{C}_h = \boldsymbol{G}_1\boldsymbol{G}_1^H$. Then, $\boldsymbol{X}$ can be expressed in the form $\boldsymbol{G}G^H$, and hence, is a Hermitian positive semidefinite matrix (its eigenvalues are nonnegative). Define $\boldsymbol{Y} = \boldsymbol{X} + \boldsymbol{\Lambda}_{\boldsymbol{X_d}} + b\boldsymbol{I}$, where $\boldsymbol{X_d}$ is a diagonal matrix from the subset of $\{\boldsymbol{X}\}$ which consists of diagonal matrices only, $\boldsymbol{\Lambda}_{\boldsymbol{X_d}}$ is a diagonal matrix such that $\boldsymbol{X_d} + \boldsymbol{\Lambda}_{\boldsymbol{X_d}} = a\boldsymbol{I}$, and $a, b > 0$. The determinant (product of eigenvalues) of $\boldsymbol{X_d} + \boldsymbol{\Lambda}_{\boldsymbol{X_d}} + b\boldsymbol{I} = (a+b)\boldsymbol{I}$ is $(a+b)^{N_{\mathrm{Tx}}L}$. Now, we form groups of $\boldsymbol{Y}$ (and hence, groups of $\boldsymbol{X}$), where within each group all $\boldsymbol{Y}$'s have the same determinant of $(a+b)^{N_{\mathrm{Tx}}L}$ for any possible $a, b > 0$. Using the arithmetic-geometric mean inequality,[1] we conclude that the trace (sum of the eigenvalues) of $\boldsymbol{Y}$ will be minimum when $\boldsymbol{Y} = (a+b)\boldsymbol{I}$, which corresponds to $\boldsymbol{X} = \boldsymbol{X_d}$ within each group. Hence, we just need to consider diagonal matrices $\boldsymbol{X_d}$. Since $\boldsymbol{C_h}$ is diagonal, $\boldsymbol{X}$ will also be diagonal when $\boldsymbol{S}^H\bar{\boldsymbol{V}}(v)\boldsymbol{S}$ is diagonal, where $\bar{\boldsymbol{V}}(v) \equiv (\boldsymbol{I}-\bar{\boldsymbol{W}}(v)) = \mathrm{diag}\{\bar{V}_0,\ldots,\bar{V}_{QK-1}\}$.

For a MIMO system, $\boldsymbol{S}$ is composed of $\{\boldsymbol{S}_k[q] : q = 0,\ldots,Q-1;\; k = 0,\ldots,N_{\mathrm{Tx}}-1\}$ [see (3)]. Since all $\boldsymbol{S}_k[q]$'s are circulant, $\boldsymbol{S}^H\bar{\boldsymbol{V}}(v)\boldsymbol{S}$ will be diagonal for *any* v iff

$$s_{k,q}[n] = \sum_{i=0}^{d_{k,q}-1}A_{k,q,i}\delta[n-l_{k,q,i}], \quad k = 0,\ldots,N_{\mathrm{Tx}}-1 \tag{12}$$

$$d_q \equiv \sum_{k=0}^{N_{\mathrm{Tx}}-1}d_{k,q} \leq \frac{K}{L} \tag{13}$$

where $d_{k,q}$ is the number of nonzero samples of the qth OFDM training symbol (excluding CP samples) for the kth transmit

[1]For positive numbers λ_i, $\prod_{i=1}^N\lambda_i \leq ((1/N)\sum_{i=1}^N\lambda_i)^N$, and the equality holds iff all λ_i's are equal.

antenna, and for each q, $\{l_{k,q,i} : \forall k, i\}$ are any permutation of $\{m_p\}$ with $m_{p+1} - m_p \geq L$, $K + m_0 - m_{d_q-1} \geq L$, and $0 \leq m_p \leq K-1$. To satisfy the condition in (7), each transmit antenna must have the same total transmitted training energy [9], which, in turn, implies that

$$\sum_{q=0}^{Q-1} \sum_{i=0}^{d_{k,q}-1} |A_{k,q,i}|^2 = E_{\mathrm{av}} \quad \forall k. \tag{14}$$

Using (12), we obtain

$$\begin{aligned} \mathrm{Tr}[\bar{\boldsymbol{X}}_{\boldsymbol{d}}] = & \sum_{k=0}^{N_{\mathrm{Tx}}-1} \sum_{m=0}^{L-1} \sigma_m^2 E \\ & \times \left[\left| \sum_{q=0}^{Q-1} \sum_{i=0}^{d_{k,q}-1} |A_{k,q,i}|^2 \bar{V}_{m+l_{k,q,i}+qK} \right|^2 \right] \quad (15) \\ = & \sum_{k=0}^{N_{\mathrm{Tx}}-1} \sum_{m=0}^{L-1} \sigma_m^2 \\ & \times \left\{ \sum_{q=0}^{Q-1} \sum_{i=0}^{d_{k,q}-1} |A_{k,q,i}|^4 E\left[\left| \bar{V}_{m+l_{k,q,i}+qK} \right|^2 \right] \right. \\ & + 2 \sum_{q_1=0}^{Q-1} \sum_{i_1=0}^{d_k-1} \sum_{\substack{q_2=0 \\ q_2Q+i_2>q_1Q+i_1}}^{Q-1} \sum_{i_2=0}^{d_k-1} \left| A_{k,i_1} \right|^2 \left| A_{k,i_2} \right|^2 \\ & \left. \times \Re\left[E\left[\bar{V}_{m+l_{k,i_1}+q_1K} \bar{V}^*_{m+l_{k,i_2}+q_2K} \right] \right] \right\} \quad (16) \end{aligned}$$

where $\bar{V}_{l_{k,q,i}} \in \{\bar{V}_l : l = 0, \ldots, QK-1\}$ is associated with the ith nonzero training sample of the qth symbol for the kth antenna, and

$$\begin{aligned} E\left[\bar{V}_l \bar{V}_m^*\right] = 1 & - e^{j2\pi k_l v/K} e^{-k_l\sigma_p^2/2} - e^{-j2\pi k_m v/K} e^{-k_m\sigma_p^2/2} \\ & + e^{j2\pi(k_l-k_m)v/K} e^{-|k_l-k_m|\sigma_p^2/2}. \end{aligned} \tag{17}$$

The optimal training signal is then defined by

$$\left\{ \left\{ \left| A^\star_{k,q,i} \right| \right\}, \left\{ l^\star_{k,q,i} \right\}, \left\{ d^\star_{k,q} \right\} \right\} = \arg\min_{\{\{|A_{k,q,i}|\},\{l_{k,q,i}\},\{d_{k,q}\}\}} \mathrm{Tr}[\bar{\boldsymbol{X}}_{\boldsymbol{d}}]. \tag{18}$$

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 typical small values of v and σ_p^2, we obtain

$$\Re\left[E\left[\bar{V}_l \bar{V}_m^*\right]\right] > 0 \tag{19}$$

$$\Re\left[E\left[\bar{V}_l \bar{V}_m^*\right]\right] < \Re\left[E\left[\bar{V}_{l+qK} \bar{V}^*_{m+qK}\right]\right], \quad q = 1, 2, \ldots \tag{20}$$

$$\begin{aligned} \left[E\left[|\bar{V}_l|^2\right]\right] & = 2\left[1 - e^{\frac{-k_l\sigma_p^2}{2}} \cos\left(\frac{2\pi k_l v}{K}\right)\right] \\ & < \left[E\left[|\bar{V}_{l+qK}|^2\right]\right], \quad q = 1, 2, \ldots \end{aligned} \tag{21}$$

which indicates that $Q = 1$ minimizes (16). Hence, we can conclude that under the same total training-signal energy constraint, using one OFDM training symbol is more robust to frequency offset and PN than using multiple training symbols. The intuitive explanation is that the phase offset caused by the frequency offset is smaller for a smaller time interval, hence giving a smaller MSE for one OFDM training symbol.

In the following, we will use $Q = 1$ and the corresponding index q will be omitted for clarity. Then, our objective function to minimize becomes

$$\begin{aligned} \mathrm{Tr}[\bar{\boldsymbol{X}}_{\boldsymbol{d}}] = & \sum_{k=0}^{N_{\mathrm{Tx}}-1} \sum_{m=0}^{L-1} \sigma_m^2 \left\{ \sum_{i=0}^{d_k-1} |A_{k,i}|^4 E\left[\left| \bar{V}_{m+l_{k,i}} \right|^2 \right] \right. \\ & \left. + 2 \sum_{i=0}^{d_k-1} \sum_{n=1;n>i}^{d_k-1} |A_{k,i}|^2 |A_{k,n}|^2 \Re\left[E\left[\bar{V}_{m+l_{k,i}} \bar{V}^*_{m+l_{k,n}} \right] \right] \right\}. \end{aligned} \tag{22}$$

Since v and σ_p^2 are typically very small, by using the Taylor series approximation, we obtain

$$\Re\left[E\left[\bar{V}_l \bar{V}_m^*\right]\right] \simeq \frac{4\pi^2 v^2 lm}{K^2}. \tag{23}$$

Substituting (23) into (22) gives

$$\mathrm{Tr}[\bar{\boldsymbol{X}}_{\boldsymbol{d}}] \simeq \frac{4\pi^2 v^2}{K^2} \sum_{k=0}^{N_{\mathrm{Tx}}-1} \sum_{m=0}^{L-1} \sigma_m^2 \left[\sum_{i=0}^{d_k-1} |A_{k,i}|^2 (m + l_{k,i}) \right]^2. \tag{24}$$

By using the fact that $l_{k,i} \neq l_{m,n}$ if $(k,i) \neq (m,n)$, $l_{k,i} - l_{m,n} \geq L$, $K + l_{k,i} - l_{m,n} \geq L$ for any $l_{m,n} > l_{k,i}$, together with $l_{k,i} < l_{k,n}$ for $i < n$, from (24), we obtain the following optimal values of the parameters:

$$\begin{aligned} & d_k^\star = 1, \quad |A_k^\star|^2 = E_{\mathrm{av}} \\ & \{l_k^\star : k = 0, \ldots, N_{\mathrm{Tx}}-1\} = \{mL : m = 0, \ldots, N_{\mathrm{Tx}}-1\} \end{aligned} \tag{25}$$

where we have dropped the index i. The corresponding optimal pilot tones are

$$\begin{aligned} & \{c_k^\star[n] : k = 0, 1, \ldots, N_{\mathrm{Tx}}-1\} \\ & = \{\sqrt{E_{\mathrm{av}}} e^{j\phi_m} e^{-j2\pi mnL/K} : m = 0, 1, \ldots, N_{\mathrm{Tx}}-1\} \end{aligned} \tag{26}$$

which are of CDM(F) (CDM in the frequency domain [9]) pilot allocation over all subcarriers.

Under different system parameters and conditions, such as spectral constraints (some subcarriers are nulled) in emerging cognitive radio systems and peak-to-average energy-ratio constraint of the power amplifier, some training signals may give better performance/flexibility than others among those training signals with the same minimum NMSE. Hence, we investigate other training signal matrices which give the same minimum trace of $\bar{\boldsymbol{X}}$ by using the relation $\mathrm{Tr}[\boldsymbol{P}\bar{X}\boldsymbol{P}^H] = \mathrm{Tr}[\bar{\boldsymbol{X}}]$, where $\boldsymbol{P}$ is a unitary matrix (see [9] for the detailed steps).[2] For $K > L_0 N_{\mathrm{Tx}}$, the obtained pilot tones are the same as those in (26). For $K = L_0 N_{\mathrm{Tx}}$, we obtain the following additional pilot designs:

$$c_k^\star[n] = \sum_{m=0}^{N_{\mathrm{Tx}}-1} \sum_{l=0}^{L_0-1} \beta_{k,m} \delta[n - lN_{\mathrm{Tx}} - m] \tag{27}$$

[2]Although [9] did not include PN, these steps and results are applicable to the system with PN.

where $\beta_{k,m} \in \mathbb{C}$. For constant-modulus pilot tones, $|\beta_{k,m}|$ is either zero or a constant. We conclude from (27) that for $K = L_0 N_{\text{Tx}}$, an additional condition for the optimal training signals in the presence of frequency offset and PN is that *for each transmit antenna k, the optimal pilot-tone symbols $c_k^\star[n - lN_{\text{Tx}} - m]$ for different l are the same*. This can be explained intuitively as follows. This condition allocates training-signal energy only to the first sample of the mth segment $\{s_k[n] : n = mL_0, \ldots, (m+1)L_0 - 1\}$ for any $m \in \{0, 1, \ldots, N_{\text{Tx}} - 1\}$. Within each segment, the phase offset due to the frequency offset is the smallest at the first sample, which, in turn, yields the smallest MSE.

The NMSE for the proposed training signal in the presence of frequency offset and PN is given by

$$\text{NMSE} = \frac{\sigma_n^2}{E_{\text{av}}} + \frac{2}{LN_{\text{Tx}}} \sum_{k=0}^{N_{\text{Tx}}-1} \sum_{m=0}^{L-1} \sigma_m^2 \times \left[1 - e^{\frac{-(m+kL)\sigma_p^2}{2}} \cos\left(2\pi(m+kL)v/K\right)\right]. \quad (28)$$

If $\sigma_p^2 = 0 (v = 0)$, (28) gives the NMSE in the presence of frequency offset (PN) only. For typical small (residual) frequency offsets, we can approximate the NMSE as

$$\text{NMSE} \simeq \frac{\sigma_n^2}{E_{\text{av}}} + \frac{1}{LN_{\text{Tx}}} \sum_{k=0}^{N_{\text{Tx}}-1} \sum_{m=0}^{L-1} \sigma_m^2 \left[2\left(1 - e^{\frac{-(m+kL)\sigma_p^2}{2}}\right) + e^{\frac{-(m+kL)\sigma_p^2}{2}} \left(\frac{2\pi(m+kL)v}{K}\right)^2\right]. \quad (29)$$

From (28) and (29), we observe that the PN level σ_p^2 (the frequency offset v) affects the extra NMSE exponentially (quadratically).

IV. Simulation Results and Discussions

Due to the space limitation, we refer readers to [16] for some examples of optimal training signals in the absence or presence of frequency offsets, for the range of the extra NMSEs of the optimal training signals from [9], and the corroborating simulation results for the NMSE comparison between the proposed training signal and the other training signals from [9]. The minimum NMSEs achieved with the proposed optimal training signals are plotted in Fig. 1 for different values of v and SNR $(= E_{\text{av}} N_{\text{Tx}}/(QK\sigma_n^2))$. At moderate-to-high SNR, $v = 0.1$ introduces a significant degradation in channel estimation, while $v \leq 0.01$ causes insignificant degradation.

Simulation results for the performance comparison of several training signals in the presence of frequency offset $v = 0.01$ and Wiener PN (with $\sigma_p^2 = 0.0001, 0.0025$) are presented in Fig. 2. Due to the space limitation, readers are referred to [17] for more simulation results for the effects of PN. PN with $\sigma_p^2 \geq 0.0025$ introduces an NMSE floor (a larger σ_p^2 gives a larger floor). For SNR values of practical interest, the performance degradation due to PN with $\sigma_p^2 \leq 0.0001$ is negligible.

Based on Figs. 1 and 2, together with (28) and (29), the following remarks are in order.

1) In the absence of frequency offset and PN, NMSE depends only on $(E_{\text{av}}/\sigma_n^2)$ regardless of N_{Tx}. In the presence of frequency offset or/and PN, NMSE depends on $(E_{\text{av}}/\sigma_n^2)$, $\boldsymbol{C_h}$, N_{Tx}, v, or/and σ_p^2.
2) A larger N_{Tx} results in a larger NMSE in the presence of frequency offset or/and PN. This implies that channel estimation in a MIMO system is more sensitive to frequency offset and PN than in a SISO system.
3) At very low SNR, the NMSE is mainly dominated by the NMSE_0, and the effect of Δ_{NMSE} is insignificant. At moderate-to-high SNR, as v and (or) σ_p^2 increase(s), Δ_{NMSE} becomes the dominating factor and the advantage of the optimal training signals has a greater impact.

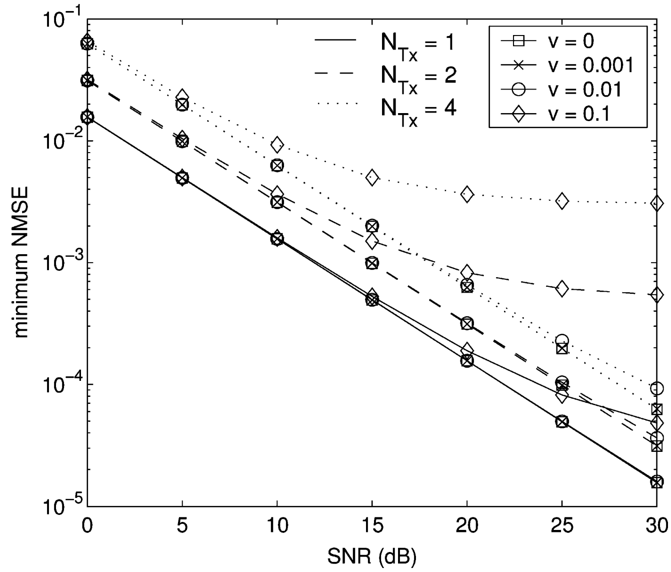

Fig. 1. Minimum NMSE for different values of N_{Tx}, v, and SNR in MIMO OFDM systems with $K = 64$, $N_g = 16$ in an eight-tap multipath Rayleigh fading channel with an exponential power delay profile.

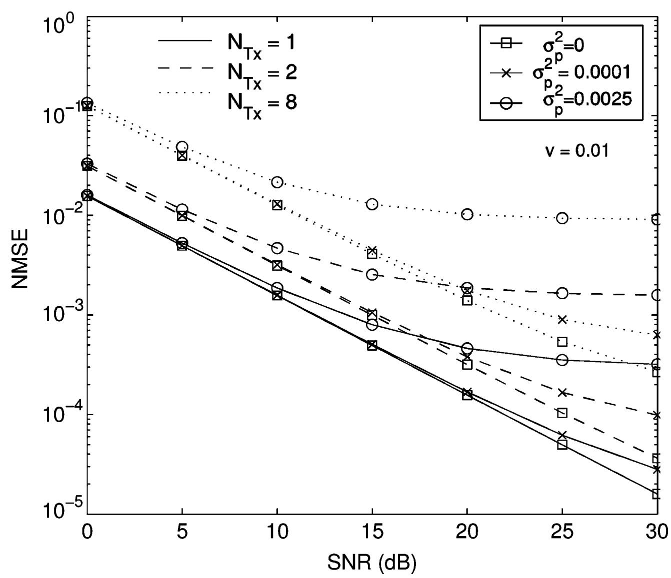

Fig. 2. Effect of PN on the NMSE of proposed training signal (training #1) for MIMO OFDM system with $K = 64$, $N_g = 16$, and $v = 0.01$ in an eight-tap multipath Rayleigh fading channel with an exponential power delay profile. ($\sigma_p^2 = 4\pi\beta/(K\Delta_f)$ reflects the PN level, where β is the one-sided 3-dB linewidth of the Lorentzian PN power density spectrum).

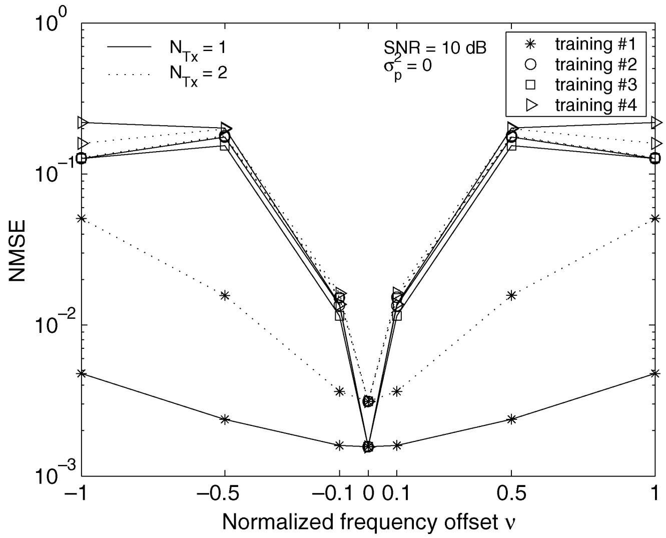

Fig. 3. Effects of different frequency offsets on the channel-estimation NMSE of different training signals without PN. ($\sigma_p^2 = 4\pi\beta/(K\Delta_f)$ reflects the PN level, where β is the one-sided 3-dB linewidth of the Lorentzian PN power density spectrum).

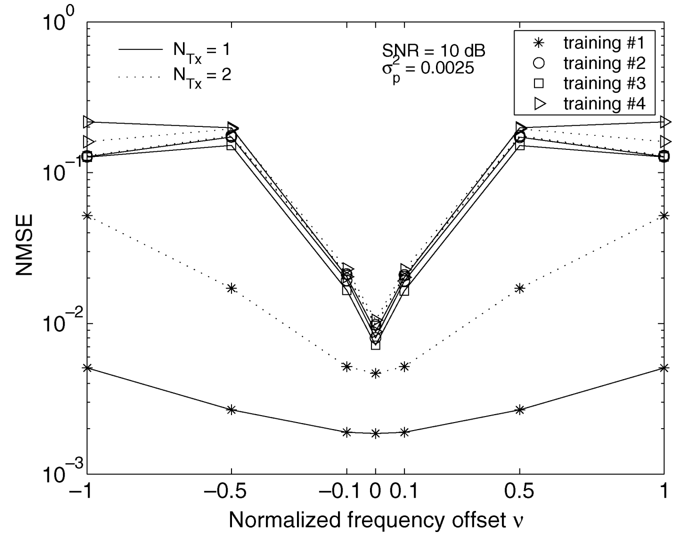

Fig. 4. Effects of different frequency offsets on the channel-estimation NMSE of different training signals with PN. ($\sigma_p^2 = 4\pi\beta/(K\Delta_f)$ reflects the PN level, where β is the one-sided 3-dB linewidth of the Lorentzian PN power density spectrum).

4) For $K = LN_{\mathrm{Tx}}$, NMSE improvement of the proposed training signals is marginal, but for $K > LN_{\mathrm{Tx}}$, which is a more practical scenario, NMSE improvements of the proposed training signals are significant. A smaller N_{Tx} gives a larger NMSE improvement of the proposed training signals.

In Figs. 3 and 4, simulation results for the effects of different frequency offsets on the channel-estimation NMSE of several training signals are presented for scenarios without PN and with PN, respectively, for SNR = 10 dB. (Results for SNR = 0 and 20 dB are not plotted due to space limitation.) These simulation results match the theoretical NMSE results in (9) and (28) (not shown in the figures for clarity). In the figures, training#1 represents an optimal training signal (proposed), training#2 employs a frequency-division multiplexing (FDM) pilot allocation with L tones for each antenna, training#3 is of a CDM(F) allocation over $2L$ subcarriers, and training#4 uses a CDM(F) allocation over all subcarriers (not the optimal one) (see [9] for the details of the FDM and CDM(F) pilot allocations). Although the proposed optimal training signals are derived based on the condition of very small v, our numerical evaluation in [16] and simulation results in Figs. 3 and 4 show that the proposed training signals become more effective for a larger residual frequency offset $|v| < 0.5$, and still remain the most robust among all training signals from [9] even at $|v| = 1$. The NMSE reductions of our proposed training signal at $|v| = 1, 0.5, 0.1, 0$ with $\sigma_p^2 = 0$ are about 9, 10, 2, and 0 dB, respectively, for $N_{\mathrm{Tx}} = 1$, and 3, 7, 1.3, and 0 dB for $N_{\mathrm{Tx}} = 2$ at SNR = 0 dB. The corresponding values at SNR = 10 dB are about 15, 19, 9, 0 dB for $N_{\mathrm{Tx}} = 1$, and 4, 11, 6, 0 dB for $N_{\mathrm{Tx}} = 2$. At SNR = 20 dB, they are about 16, 22, 18, and 0 dB for $N_{\mathrm{Tx}} = 1$, and 4, 11, 11, and 0 dB for $N_{\mathrm{Tx}} = 2$. With $\sigma_p^2 = 0.0025$, the corresponding reductions for $N_{\mathrm{Tx}} = 1, 2$ are about (9, 10, 3, 1.4) dB and (3, 6.5, 1.7, 0.6) dB at SNR = 0 dB, (14, 18, 10, 6) dB and (4, 10, 6, 3) dB at SNR = 10 dB, and (16, 21, 15, 11) dB and (4, 11, 9, 6) dB at SNR = 20 dB. Our proposed training signals are even more advantageous in the presence of both residual frequency offset and PN than either one alone. As the PN variance and (or) residual frequency offset ($|v| < 0.5$) increase(s), the NMSE difference between the proposed training signals and the reference training signals increases.

V. Conclusions

We presented MSE-optimal training signals for MIMO OFDM channel estimation in the presence of frequency offset and PN. Frequency offset and PN introduce a channel-estimation NMSE floor which is higher for a larger residual frequency offset, a larger PN level, and a larger number of transmit antennas. Individually, the PN effect on NMSE is similar to the frequency-offset effect on NMSE. The perfomance advantage of the proposed training signals over other training signals is greater for a smaller number of transmit antennas, a larger residual frequency offset, and a larger PN level, and more significant in the presence of both residual frequency offset and PN than either one alone.

References

[1] S. Adireddy, L. Tong, and H. Viswanathan, “Optimal placement of training for frequency-selective block-fading channels,” *IEEE Trans. Inf. Theory*, vol. 48, no. 8, pp. 2338–2353, Aug. 2002.

[2] M. Dong and L. Tong, “Optimal design and placement of pilot symbols for channel estimation,” *IEEE Trans. Signal Process.*, vol. 50, no. 12, pp. 3055–3069, Dec. 2002.

[3] I. Barhumi, G. Leus, and M. Moonen, “Optimal training design for MIMO OFDM systems in mobile wireless channels,” *IEEE Trans. Signal Process.*, vol. 51, no. 6, pp. 1615–1624, Jun. 2003.

[4] R. Negi and J. Cioffi, “Pilot tone selection for channel estimation in a mobile OFDM system,” *IEEE Trans. Consum. Electron.*, vol. 44, no. 3, pp. 1122–1128, Aug. 1998.

[5] S. Ohno and G. B. Giannakis, “Optimal training and redundant precoding for block transmissions with application to wireless OFDM,” *IEEE Trans. Commun.*, vol. 50, no. 12, pp. 2113–2123, Dec. 2002.

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

[7] Y. Li, "Simplified channel estimation for OFDM systems with multiple transmit antennas," *IEEE Trans. Wireless Commun.*, vol. 1, no. 1, pp. 67–75, Jan. 2002.

[8] X. Ma, L. Yang, and G. B. Giannakis, "Optimal training for MIMO frequency-selective fading channels," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 453–466, Mar. 2005.

[9] H. Minn and N. Al-Dhahir, "Optimal training signals for MIMO OFDM channel estimation," *IEEE Trans. Wireless Commun.*, vol. 5, no. 5, pp. 1158–1168, May 2006.

[10] T. Pollet, M. Van Bladel, and M. Moeneclaey, "BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise," *IEEE Trans. Commun.*, vol. 43, no. 2–4, pp. 191–193, Feb.–Apr. 1995.

[11] M. Gudmundson and P. O. Anderson, "Adjacent channel interference in an OFDM system," in *Proc. Veh. Technol. Conf.*, Atlanta, GA, May 1996, pp. 918–922.

[12] J. Tubbax, B. Come, L. Van der Perre, S. Donnay, M. Engels, H. D. Man, and M. Moonen, "Compensation of IQ imbalance and phase noise in OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 872–877, May 2005.

[13] A. Stamoulis, S. N. Diggavi, and N. Al-Dhahir, "Intercarrier interference in MIMO OFDM," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2451–2463, Oct. 2002.

[14] Y. Mostofi and D. C. Cox, "ICI mitigation for pilot-aided OFDM mobile systems," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 765–774, Mar. 2005.

[15] Z. J. Wang, Z. Han, and K. J. R. Liu, "A MIMO-OFDM channel estimation approach using time of arrivals," *IEEE Trans. Wireless Commun.*, vol. 4, no. 3, pp. 1207–1213, May 2005.

[16] H. Minn and N. Al-Dhahir, "Training signal design for MIMO OFDM channel estimation in the presence of frequency offsets," in *Proc. IEEE WCNC*, New Orleans, LA, Mar. 2005, vol. 1, pp. 1–6.

[17] ——, "MIMO-OFDM training signal design in the presence of frequency offset and phase noise," in *Proc. ICICS*, Bangkok, Thailand, Dec. 2005, pp. 263–267.