BANDWIDTH AND POWER EFFICIENCY CONSIDERATIONS FOR OPTIMAL TRAINING IN OFDM

Ning Chen[1] *G. Tong Zhou*[1]* *Ananthram Swami*[2]

[1]School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0250, USA
[2]US Army Research Laboratory, Adelphi, MD 20783, USA

ABSTRACT

Orthogonal Frequency Division Multiplexing (OFDM) has been adopted by many high speed transmission standards. Pilot tone assisted modulation (PTAM) is a well known technique to estimate the channel state information (CSI) in OFDM but is not bandwidth efficient. To achieve high bandwidth efficiency, we propose to carry out superimposed training with techniques described in this paper to improve the quality of the CSI estimate. It is well known that OFDM signals have high peak-to-average power ratios (PARs) and are thus not power efficient. To improve the power efficiency, we employ a blind Selected Mapping (SLM) PAR reduction method in conjunction with superimposed training. Simulation results demonstrate that under medium to low SNR conditions, combining superimposed training with PAR reduction yields BER performance that is comparable with PTAM; the benefit of the former is that bandwidth efficiency is preserved.

1. INTRODUCTION

Orthogonal Frequency Division Multiplexing (OFDM) has attracted a great deal of attention in recent years mainly because of its simple equalizer structure. OFDM has been adopted by many standards, such as digital audio broadcast (DAB) and digital video broadcast (DVB) in Europe, as well as in asymmetric digital subscriber line (ADSL), Wireless Local Area Network (WLAN) and Broadband Wireless Access (BWA) applications.

To mitigate the inter-symbol interference (ISI), pilot tone assisted modulation (PTAM) has been proposed to acquire the channel state information (CSI) for frequency-selective channels [1–3]. By maximizing the training-based channel capacity, it was shown in [3] that the optimal way to place the pilot tones is to use equi-spaced subcarriers. On the other hand, by minimizing the mean-squared error (MSE) of the symbol estimate, it was shown in [2] that equi-spaced and equi-powered pilot tones are optimal. PTAM has a superior performance in acquiring the CSI due to the decoupling of the pilot tones and the information subsymbols. However, two important issues arise in PTAM-OFDM systems: reduced bandwidth efficiency due to the dedicated slots for training and generally low power efficiency due to the high peak-to-average power ratio (PAR) of the OFDM signal. In addition, the high PAR values of the OFDM signals make them sensitive to nonlinear effects in the power amplifier (PA). To minimize nonlinear distortions, either very linear PAs have to be used, or the input signal has to be backed-off significantly. Both options are power inefficient.

Based on the optimal pilot tone design in [2], we propose here a novel joint channel estimation and PAR reduction technique for OFDM, the aim being to improve the bandwidth efficiency as well as to reduce the PAR (thus improving the power efficiency). We adopt superimposed training for the purpose of channel estimation; i.e., we add pilot tones onto the information subsymbols. We can achieve a high bandwidth efficiency, but the CSI estimate will not be as accurate as in the dedicated training case. We will propose two methods to improve the CSI estimate. For better power efficiency, we employ the selected mapping (SLM) [4] technique to reduce the PAR, and embed the side information about the SLM index in the position of the pilot tones. We demonstrate the PAR reducing capability of the proposed method as well as the corresponding BER performance. In comparison with PTAM, our proposed method is advantageous at medium to low SNRs (< 10 dB): higher bandwidth efficiency can be achieved without sacrificing the BER.

Notations: A time-domain sequence is represented by a lower case letter; e.g., $s[n]$. Its frequency-domain representation is denoted by the corresponding capitalized letter; e.g., $S[k]$. A bold-faced lower case letter denotes a vector of the frequency-domain sequence; e.g., $\mathbf{s} = [S[0],\ S[1],\ \ldots,\ S[N-1]]^T$. Its time-domain counterpart is denoted by adding a subscript $_t$; e.g., $\mathbf{s}_t = [s[0],\ s[1],\ \ldots,\ s[N-1]]^T$. A bold-faced capital letter denotes a matrix.

2. REVIEW OF PTAM-OFDM

In OFDM, the length N frequency-domain subsymbols $\mathbf{s} = [S[0],\ S[1],\ \ldots,\ S[N-1]]^T$ are transformed into the time-domain to form the transmitted signal as $\mathbf{s}_t = \mathbf{F}^H\ \mathbf{s}$, where $\mathbf{F}$ is the $N \times N$ normalized discrete Fourier transform (DFT) matrix with $\mathbf{F}(k,n) = 1/\sqrt{N}\ e^{-j2\pi kn/N}$, and $\mathbf{F}^H$ is its conjugate transpose. In PTAM-OFDM, P pilot tones are inserted in the frequency domain in order to acquire the CSI; $P \geq L$ is assumed where L is the length of the finite impulse response (FIR) channel. The frequency-domain transmitted signal is

$$\mathbf{x} = \mathbf{s} + \mathbf{b}, \tag{1}$$

where $\mathbf{s}$ consists of $N-P$ information subsymbols and P zeros, $\mathbf{b} = [B[0],\ B[1],\ \ldots,\ B[N-1]]^T$ consists of P pilot tones and $N-P$ zeros. The locations of the zeros in $\mathbf{s}$ and

*This work was supported in part by the US Army Research Laboratory Communications and Networks Collaborative Technology Alliance Program and the US National Science Foundation Grant No. 0218778.

$\mathbf{b}$ are carefully selected so that when $S[k] = 0,\ B[k] \neq 0$, and when $B[k] = 0,\ S[k] \neq 0$. The placement of the pilot tones and the allocation of the power between $\mathbf{b}$ and $\mathbf{s}$ have been studied in [1–3]. The same problem has also been investigated for single carrier systems [5]. All these studies point to the same optimality conditions for the pilot tones for estimating frequency selective block fading channels.

According to [2], the optimal placement strategy is to modulate $P = L$ pilot tones with equal power onto equi-spaced subcarriers. For simplicity, let us assume that the number of sub-carriers N is an integer multiple of P; i.e., $R = N/P$ is an integer. Next define a set of P equi-spaced pilot tone indices as

$$\mathbf{\Omega}_0 \triangleq \left\{ k_i \,\middle|\, k_i = iR + \theta_0, 0 \le i \le P-1, 0 \le \theta_0 \le R-1 \right\}, \tag{2}$$

which can be characterized by the shift parameter θ_0 alone.

We consider a frequency selective block fading channel, which is modeled by a time-invariant (over each OFDM block) FIR filter $\mathbf{h} = [h[0],\ h[1],\ \ldots,\ h[L-1]]^T$. With the insertion/removal of the cyclic prefix (CP), the frequency selectivity of the channel appears flat on each subcarrier; i.e.,

$$\mathbf{y} = \mathbf{D}_H\ \mathbf{x} + \mathbf{v}, \tag{3}$$

where $\mathbf{y} = [Y[0],\ Y[1],\ \ldots,\ Y[N-1]]^T$ is the received signal after the CP removal and the application of the DFT, $\mathbf{D}_H$ is a diagonal matrix with the channel DFT coefficients (i.e., $H[k] = \sum_{n=0}^{L-1} h[n]e^{-j2\pi kn/N},\ 0 \le k \le N-1$) on the diagonal, and $\mathbf{v} = [V[0],\ V[1],\ \ldots,\ V[N-1]]^T$ is the normalized DFT of the additive noise $v[n]$; i.e., $V[k] = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} v[n]e^{-j2\pi kn/N}$.

Substituting (1) into (3), we obtain

$$\begin{aligned} \mathbf{y} &= \mathbf{D}_H\ (\mathbf{s} + \mathbf{b}) + \mathbf{v} \\ &= \mathbf{D}_H\ \mathbf{s} + \mathbf{D}_B\ \mathbf{F}_{1:L}\ \mathbf{h} + \mathbf{v}, \end{aligned} \tag{4}$$

where $\mathbf{D}_B$ is a diagonal matrix with $\mathbf{b}$ on the diagonal, and $\mathbf{F}_{1:L} = \sqrt{N}\mathbf{F}(:, 1:L)$ is an $N \times L$ DFT matrix consisting of the first L columns of $\sqrt{N}\mathbf{F}$.

Let $\mathbf{B} = \mathbf{D}_B\ \mathbf{F}_{1:L}$ and

$$\mathbf{B}^\dagger = (\mathbf{F}_{1:L}^H \mathbf{D}_B^H \mathbf{D}_B \mathbf{F}_{1:L})^{-1}\ \mathbf{F}_{1:L}^H \mathbf{D}_B^H \tag{5}$$

be the pseudoinverse of $\mathbf{B}$. Since $\mathbf{B}$ has full column rank, we can pre-multiply $\mathbf{B}^\dagger$ on both sides of (4) to obtain

$$\mathbf{B}^\dagger\ \mathbf{y} = \mathbf{B}^\dagger\ \mathbf{D}_H\ \mathbf{s} + \mathbf{h} + \mathbf{B}^\dagger\ \mathbf{v}. \tag{6}$$

Since the pilot tones and the information subsymbols are decoupled, we have $\mathbf{D}_B^H\ \mathbf{s} = \mathbf{0}$, which allows us to write $\mathbf{D}_B^H \mathbf{D}_H\ \mathbf{s} = \mathbf{0}$ as well. Thus, the first term on the RHS of (6) is zero and

$$\mathbf{B}^\dagger\ \mathbf{y} = \mathbf{h} + \mathbf{B}^\dagger\ \mathbf{v}. \tag{7}$$

We infer from (7) that the least squares estimate of the channel $\mathbf{h}$ is

$$\hat{\mathbf{h}} = \mathbf{B}^\dagger\ \mathbf{y}. \tag{8}$$

In other words, $\hat{\mathbf{h}} = \mathbf{h} + \mathbf{B}^\dagger \mathbf{v}$; and hence the unknown $\mathbf{s}$ does not affect the accuracy in $\hat{\mathbf{h}}$.

Afterwards, the information subsymbols can be estimated as

$$\hat{\mathbf{s}} = \mathbf{D}_{\hat{H}}^{-1}\ \mathbf{y} - \mathbf{b}, \tag{9}$$

where $\mathbf{D}_{\hat{H}}^{-1}$ is a diagonal matrix with $1/\hat{H}[k]$ on the diagonal[1]. $\hat{\mathbf{s}}$ is then decoded to yield the $\bar{\mathbf{s}}$ estimate belonging to the symbol constellation.

Remarks:

1. The inserted pilot tones in PTAM reduce the bandwidth efficiency. For example, when $N = 64,\ L = 4$, there is a $L/N \times 100\% = 6.25\%$ loss in bandwidth efficiency. To avoid this loss, we propose to keep the equi-spaced and equi-powered pattern of the pilot tones, and superimpose these pilot tones onto the information data. We will describe methods to improve the quality of the CSI estimate under such superimposed training framework.
2. OFDM signals suffer from high PARs. To alleviate the high PAR problem, we propose to employ SLM for PAR reduction. We will describe a method to embed the SLM side information in the transmitted data, which can be recovered blindly at the receiver.

3. SUPERIMPOSED TRAINING FOR OFDM

To improve the bandwidth efficiency of a PTAM-OFDM system, we propose to superimpose the pilot tones onto the information subsymbols.

3.1. Improved CSI Estimation

In superimposed training, the pilot tones are no longer decoupled from the information subsymbols; i.e., $\mathbf{D}_B^H \mathbf{s} \neq \mathbf{0}$, and thus (7) does not hold any more. Since $E[\mathbf{D}_H\ \mathbf{s}] = \mathbf{0}$, equation (8) can still be used to estimate the CSI, although the performance of the channel estimate will not be as accurate as in the PTAM case. To improve (8) for the superimposed training case, we propose two modifications as described below.

1. Under the superimposed training framework, it can be shown (e.g., [6,7]) that the variance of the channel estimate in (8) increases with the information signal power σ_s^2, which is undesirable. Therefore, we can reduce the variance of $\hat{h}[l]$ by lowering the average power of those information subsymbols that are co-located with the pilot tones. For example, let us set

$$E[|S[k]|^2] = \begin{cases} \alpha\sigma_s^2, & k \in \mathbf{\Omega}_0 \\ \sigma_s^2, & k \in \mathbf{\Omega}_0^\perp, \end{cases} \tag{10}$$

where $\mathbf{\Omega}_0^\perp$ denotes the complement of $\mathbf{\Omega}_0$ and $0 \le \alpha \le 1$. $\alpha = 0$ corresponds to PTAM, whereas $\alpha = 1$ corresponds to conventional superimposed training. Let us denote by β, the power allocation factor, which is the ratio between the total power allocated to the pilot

[1] To use this one-tap equalizer, it is assumed that $\hat{H}[k] \neq 0,\ \forall\ k$. To cope with this, typically the symbol stream is coded across subcarriers. The MMSE estimator can also be used.

tones and the total transmitted power. For example, in PTAM,

$$\beta = \frac{P\sigma_p^2}{P\sigma_p^2 + (N-P)\sigma_s^2}, \qquad (11)$$

whereas for superimposed training with (10),

$$\beta = \frac{P\sigma_p^2}{P\sigma_p^2 + [N-(1-\alpha)P]\sigma_s^2}, \qquad (12)$$

2. We can explore the CSI contained in $\mathbf{D}_H\mathbf{s}$ as described below. We infer from (4)

$$\mathbf{u} = \mathbf{y} - \mathbf{D}_B\mathbf{F}_{1:L}\mathbf{h} = \mathbf{D}_H\mathbf{s} + \mathbf{v}. \qquad (13)$$

If $S[k]$ has a constant modulus; i.e., $|S[k]| = \sigma_s$, we can take the magnitude squared on both sides of (13) to obtain $|\mathbf{u}|^2 = |\mathbf{h}_f|^2\sigma_s^2 + |\mathbf{v}|^2 + 2\mathcal{R}\{\mathbf{D}_H\mathbf{D}_s\mathbf{v}^*\}$, where $|\cdot|^2$ is an element-wise magnitude squared operator for a scalar, a vector, or a matrix, $\mathbf{h}_f$ is the diagonal of $\mathbf{D}_H$, $\mathbf{D}_s$ is a diagonal matrix with $\mathbf{s}$ on the diagonal, $\mathcal{R}\{\cdot\}$ denotes the real part, and $(\cdot)^*$ denotes complex conjugation. Treating the last two items in $|\mathbf{u}|^2$ as noise and (8) as the initial estimate of $h[l]$, we can estimate the magnitude response of $h[l]$ as

$$|\hat{\mathbf{h}}_f| = \sqrt{\frac{|\mathbf{y} - \mathbf{D}_B\mathbf{F}_{1:L}\hat{\mathbf{h}}|^2 - \sigma_v^2}{\sigma_s^2}}, \qquad (14)$$

where again, the $\sqrt{\cdot}$ operation is carried out on each element. We propose to take the magnitude response from (14) but take the phase response from (8) to form an improved CSI estimate; i.e.,

$$\hat{H}[k] = |\hat{H}_2[k]|\; e^{j\angle \hat{H}_1[k]}, \quad 0 \le k \le N-1, \qquad (15)$$

where $\hat{H}_1[k]$ is the DFT of the $\hat{h}[l]$ in (8), and $|\hat{H}_2[k]|$ comes from the kth element of $|\hat{\mathbf{h}}_f|$ in (14).

Knowing that the channel has length $L < N$, the CSI estimate in (15) can be further improved by smoothing $\hat{H}[k]$. Specifically, we take the N-point IFFT of $\hat{H}[k]$, truncate the resulting $\hat{h}[l]$ to retain only the first L samples, and subsequently take the FFT of the truncated $\hat{h}[l]$ to obtain its DFT coefficients for use in the equalizer.

3.2. Pilot Tone Selection

The main difference between superimposed training and PTAM is that superimposed training has both data and pilot on the pilot subcarriers, but PTAM has only pilot on those subcarriers. Both schemes use the CSI estimator in (8) and they have the same optimality condition for the placement of the pilot tones, i.e., equi-spaced and equi-powered pilot tones [2,7]. Among such optimal $B[k]$'s, two classes are of interest; i.e.,

$$B_1[k] = \begin{cases} \sigma_p\; e^{j\phi_p}, & k \in \mathbf{\Omega}_0, \\ 0, & k \in \mathbf{\Omega}_0^{\perp}, \end{cases} \quad 0 \le k \le N-1, \qquad (16)$$

$$B_2[k] = \mathrm{DFT}\{b_2[n]\}, \qquad (17)$$

where the time-domain sequence $b_2[n]$ has a constant-magnitude and is periodic with period P. The general rules for constructing $B_2[k]$ sequences (equi-powered in both time and frequency domains and periodic in time/equi-spaced in frequency) are not clear, but the following periodic chirp sequence is a solution [7,8]:

$$b_2[n] = \begin{cases} \sigma_p\; e^{j\frac{\pi}{P}([n]_P^2 + [n]_P)}, & P \text{ odd},\; 0 \le n \le N-1, \\ \sigma_p\; e^{j\frac{\pi}{P}([n]_P^2 + 2[n]_P)}, & P \text{ even},\; 0 \le n \le N-1, \end{cases} \qquad (18)$$

where $[n]_P$ is the residue of n divided by P. We will elaborate next on the use of $B_1[k]$ and $B_2[k]$.

When $B[k] = B_1[k]$, the CSI estimate in (8) can be greatly simplified. By substituting (16) into (8), we infer that $\mathbf{D}_B^H\mathbf{D}_B = \sigma_p^2\,\mathbf{I}_{\mathbf{\Omega}_0}$, where $\mathbf{I}_{\mathbf{\Omega}_0}(k,k) = 1$ for $k \in \mathbf{\Omega}_0$ and zero otherwise. In this case, the first term on the RHS of (5) becomes a scaled identity matrix $(\mathbf{F}_{\mathbf{\Omega}_0,1:L}^H\,\mathbf{F}_{\mathbf{\Omega}_0,1:L})^{-1}/\sigma_p^2 = \mathbf{I}/(P\sigma_p^2)$, where $\mathbf{F}_{\mathbf{\Omega}_0,1:L}$ is a replica of $\mathbf{F}_{1:L}$ with the kth row, $k \in \mathbf{\Omega}_0^{\perp}$, being zeroed out. Therefore, (8) can be rewritten as

$$\begin{aligned} \hat{\mathbf{h}} &= \frac{1}{P\sigma_p \exp(j\phi_p)}\; \mathbf{F}_{\mathbf{\Omega}_0,1:L}^H\; \mathbf{F}\; \mathbf{y}_t \\ &= \frac{1}{P\sigma_p \exp(j\phi_p)/\sqrt{N}}\frac{1}{R}\; \Big[e^{-j2\pi\theta_0\frac{0}{R}}\mathbf{I},\; e^{-j2\pi\theta_0\frac{1}{R}}\mathbf{I}, \\ &\qquad \ldots,\; e^{-j2\pi\theta_0\frac{R-1}{R}}\mathbf{I}\Big]\,\mathbf{y}_t, \end{aligned} \qquad (19)$$

or in scalar form,

$$\hat{h}[l] = \frac{1}{P\sigma_p e^{j\phi_p}}\frac{1}{R}\sum_{r=0}^{R-1} y[rP+l]\; e^{-j2\pi\theta_0\frac{l}{R}}, \quad 0 \le l \le L-1, \qquad (20)$$

which is the phase-shifted synchronized average of the received time-domain signal $y[n]$. The receiver is very simple since the channel estimate (20) only uses a first-order statistic, and the equalizer is a simple one-tap equalizer which is an appealing feature of OFDM.

The normalized IDFT of $B_1[k]$ is

$$b_1[n] = \frac{P}{\sqrt{N}}\sigma_p\; e^{j\phi_p}\sum_{l=0}^{R-1}\delta[n-lP]\; e^{j2\pi\theta_0\frac{l}{R}}, \qquad (21)$$

which is a periodic impulse pilot sequence. As we show in [8], among all periodic sequences with the same period, superimposed training with $b_1[n]$ gives rise to the worst case PAR whereas superimposed training with $b_2[n]$ (c.f. (18)) leads to the best case PAR. The gap between the complementary cumulative distribution function (CCDF) of the best case PAR and that of the worst case PAR is a function of the number of subcarriers N, the number of pilot tones P and the power allocation factor β. For example, when $N = 128$, $P = 8$ and $\beta = 0.5$, the PAR of the transmitted OFDM signal with $b_1[n]$ is 2 dB higher than that of the OFDM signal with $b_2[n]$ at the clipping probability of 10^{-4}.

We will show in Section 5.1. that when SLM is employed to reduce the PAR of the transmitted signal, $B_1[k]$ and $B_2[k]$ do not differ much in their impact on the final PAR value after SLM.

4. BLIND SELECTED MAPPING

One major disadvantage of OFDM is the significant amplitude fluctuations; i.e., the high PAR values, requiring large backoff of the average operating power of a RF power amplifier in order to linearly amplify the signal, thus resulting in low power efficiency.

The PAR of a time-domain OFDM signal $x(t)$ is defined as [10]

$$\mathrm{PAR}(x(t)) = \frac{\mathcal{P}_{\max}}{\mathcal{P}_{av}}, \qquad (22)$$

where $\mathcal{P}_{\max} = \max_{0\le t\le T_s} |x(t)|^2$ is the peak power over one symbol duration T_s, $\mathcal{P}_{av} = \bar{E}[|x(t)|^2]$ is the average power of the OFDM symbol, and $\bar{E}[\cdot]$ denotes expectation, or time-averaged expectation if $x(t)$ is nonstationary.

We are interested in the selected mapping (SLM) approach, first proposed by Bäuml, Fischer and Huber in [4]. SLM is distortionless, and is an effective PAR reduction method. Denote by $\phi^{(m)}[k]$, $0 \le k \le N-1$, $0 \le m \le M-1$, a set of M (random) phase sequences, which is a known table that is available to both the transmitter and the receiver. Each row (indexed by m) represents a different phase rotation sequence; the column index k corresponds to the frequency-domain subcarriers.

In SLM, we first rotate the phases of $X[k]$ as described by

$$Z^{(m)}[k] = X[k]e^{j\phi^{(m)}[k]}. \qquad (23)$$

It is clear that $Z^{(m)}[k]$ and $X[k]$ contain the same information, but their time-domain counterparts $z^{(m)}(t)$ and $x(t)$ can have very different PAR values. From the M candidate $z^{(m)}(t)$ signals, $z^{(\bar{m})}(t)$, which has the lowest PAR, is transmitted. The index $\bar{m}$ ($\log_2 M$ bits) may be transmitted as side information, which is of critical importance to the receiver for decoding and is generally protected by channel coding [4].

To avoid the information rate loss caused by the transmission of the optimum index $\bar{m}$, we describe a blind SLM method that embeds the information about $\bar{m}$ in the pilot tones.

According to [2], in PTAM, as long as the pilot tones are equi-spaced and equi-powered and the additive noise is white, channel estimation performance is not affected by the choice of θ_0 (c.f. (2)). In superimposed PTAM, instead of using a pre-selected θ_0, we can try different delays (positions) $\theta_0^{(m)}$ for the pilot tones; i.e.,

$$\boldsymbol{\Theta} \triangleq \{\theta_0^{(0)},\ \theta_0^{(1)},\ \ldots,\ \theta_0^{(m)}, \ldots\ \ldots,\ \theta_0^{(M-1)}\}, \qquad (24)$$

where the same index m is used for both the pilot position and the row index in the SLM phase table. There are totally $R = N/P$ distinct pilot tone positions for $\theta_0^{(m)}$ and typically $R > M$. Assume that R is an integer multiple of M. We suggest for the M delays to be equi-spaced in order to minimize the detection error in $\bar{m}$; i.e., the possible pilot shifts are $\{\theta_0^{(0)} = 0,\ \theta_0^{(1)} = R/M,\ \ldots,\ \theta_0^{(M-1)} = \frac{M-1}{M}R\}$.

Among M equivalent representations, the mth superimposed training OFDM signal is given by

$$X^{(m)}[k] = \begin{cases} \sqrt{\alpha}S[k] + B[k], & k \in \boldsymbol{\Omega}_m, \\ S[k], & k \in \boldsymbol{\Omega}_m^{\perp}, \end{cases},$$
$$0 \le k \le N-1,\ 0 \le m \le M-1, \qquad (25)$$

where $\boldsymbol{\Omega}_m$ is represented by $\theta_0^{(m)}$ similar to the way that $\boldsymbol{\Omega}_0$ is represented by θ_0.

Next perform the phase rotations,

$$Z^{(m)}[k] = X^{(m)}[k]e^{j\phi^{(m)}[k]}. \qquad (26)$$

Similar to SLM, $z^{(m)}(t)$ and $\mathrm{PAR}\left(z^{(m)}(t)\right)$ are evaluated and $z^{(\bar{m})}(t)$, which has the lowest PAR among $\{z^{(m)}(t)\}$, is transmitted. In other words, the optimum pilot tone location - phase sequence index is

$$\bar{m} = \operatorname*{argmin}_{0\le m\le M-1} \left\{\mathrm{PAR}\left(z^{(m)}(t)\right)\right\}. \qquad (27)$$

The expected power of $X^{(m)}[k]$ is

$$E[|X^{(m)}[k]|^2] = \begin{cases} \alpha\sigma_s^2 + \sigma_p^2, & k \in \boldsymbol{\Omega}_m \\ \sigma_s^2, & k \in \boldsymbol{\Omega}_m^{\perp}. \end{cases} \qquad (28)$$

Since $\sigma_p^2 \gg \sigma_s^2$ [2, 3, 9], we see that there is a disparity in the average power at the pilot tone locations vs. at the non-pilot tone locations. Since the phase rotations in (26) do not affect the power, the same average power profile holds for $Z^{(m)}[k]$.

The received frequency-domain signal is

$$Y[k] = Z^{(\bar{m})}[k]H[k] + V[k]. \qquad (29)$$

As in [9], we first form the synchronized average of the instantaneous power of $Y[k]$ as

$$\rho_r = \frac{1}{P}\sum_{l=0}^{P-1} |Y[lR + r]|^2. \qquad (30)$$

At the receiver, we determine the optimum index $\bar{m}$, i.e., the actual pilot shift $\theta_0^{(\bar{m})}$ through peak detection in ρ_r; i.e.,

$$\hat{\theta}_0^{(\bar{m})} = \operatorname*{argmax}_{r\in\boldsymbol{\theta}} \{\rho_r\}. \qquad (31)$$

Since the receiver knows the set of possible values for $\boldsymbol{\Theta}$, from $\hat{\theta}_0^{(\bar{m})}$, a simple lookup table search yields $\hat{\bar{m}}$.

The accuracy in $\hat{\bar{m}}$ is critical for the decoding of $\hat{S}[k]$ at the receiver. If $\hat{\bar{m}}$ is erroneous for one particular OFDM block, the BER will be high for that block. The success of the algorithm largely hinges upon ensuring $\sigma_p^2 \gg \sigma_s^2$. This as well as the finite alphabet nature of $\boldsymbol{\Theta}$ help to make (31) robust over frequency selective channels. Similar findings for the PTAM-SLM case were presented in [9].

5. SIMULATIONS

In the examples in this section, we assume that the number of sub-carriers $N = 128$, the length of the FIR channel $L = 4$, and the number of pilot tones $P = L = 4$. According to [2], the optimum power allocation factor for PTAM for the above parameters is $\beta = 0.15$. For superimposed training, we chose $\beta = 0.3$ because it yields relatively good performance for both low and high SNR scenarios. The weighting factor in (25) was chosen to be $\alpha = 0.5$.

The information subsymbols were independently drawn from a QPSK constellation with Gray coding. The signal-to-noise ratio (SNR) is defined as

$$\text{SNR} = \frac{\mathcal{P}_{dc}}{N_0}, \tag{32}$$

where $\mathcal{P}_{dc}$ is the amount of DC power drawn by the PA and $N_0 = 2\sigma_v^2$ is the power spectral density of the additive noise. The effective SNR, which directly affects the BER, can be expressed by $\text{SNRe} = \mathcal{P}_{av}/N_0$, where $\mathcal{P}_{av}$ is the average output power of the PA. If an ideal linear PA is used and the signal is to be amplified undistorted, $\mathcal{P}_{av}$ is proportional to $\mathcal{P}_{dc}/\text{PAR}$ if linear scaling is employed [10]. Therefore, $\text{SNRe} \propto \mathcal{P}_{dc}/\text{PAR}/N_0$, and the benefit of PAR reduction is realized as an increase in SNRe.

It is shown in [11] that the optimal way to construct the phase table $\{\phi^{(m)}[k]\}_{0\leq k\leq N-1,\ 0\leq m\leq M-1}$ for SLM is to make $\phi^{(m)}[k]$ i.i.d. with $E[\exp(j\phi^{(m)}[k])] = 0$. Therefore, for simplicity, we can have $\phi^{(m)}[k]$ i.i.d. drawn from $\{0, \pi\}$ with equal probability so that $\exp(j\phi^{(m)}[k]) = \pm 1$ with equal probability. As such, (26) can be implemented without multiplications but with selected sign changes only.

5.1. PAR Reduction Performance

In this example, we approximate the continuous-time PAR by evaluating the discrete-time PAR of the 4-time oversampled OFDM signal [12]. 10^6 independent Monte Carlo trials were conducted.

Fig. 1 shows the empirical CCDF curves (solid lines) of the PAR of the transmitted signal $z^{(\bar{m})}(t)$ for different number of selections, M, along with the CCDF of the PAR of the original OFDM signal and that of the PTAM-OFDM signal. We can see from the figure that the PTAM-OFDM signal had a higher PAR than the OFDM signal when $B_1[k]$ was used, but a lower PAR when $B_2[k]$ was used. By employing SLM, we can significantly reduce the PAR. We observe that when $M = 8$, the proposed algorithm achieved approximately 3.5 dB of PAR reduction (compared with the original OFDM signal) at the CCDF level of 10^{-4} with either $B_1[k]$ or $B_2[k]$. We also see from Fig. 1 that the larger the M, the smaller the resulting PAR, and the smaller the gap between the CCDF corresponding to $B_1[k]$ and that corresponding to $B_2[k]$. On the other hand, the computational complexity of SLM increases as M increases. There is also a diminishing return in the PAR reduction capability as M further increases. As a rule of thumb, we recommend to use $\min\{R, 4\} \leq M \leq \min\{R,\ 8\}$.

5.2. BER Performance

Fig. 2 compares the BER performance of our proposed scheme with that of PTAM-OFDM for $N = 128$, $P = 4$, $\beta = 0.15$, and $M = 8$. When SLM was employed, $B_1[k]$ was used for its low complexity in CSI estimation. Otherwise, $B_2[k]$ was used to achieve high power efficiency of the transmitter. When perfect CSI is available at the receiver, the BER result provides a basis for comparison. The receiver used a zero-forcing equalizer and a suboptimal but simple hard-decision decoder [2]. The frequency selective channel was modeled as Rayleigh fading channels with i.i.d. complex Gaussian taps. The BER was evaluated by averaging over 10^3 Monte Carlo trials.

From Fig. 2, we can see that with superimposed training and in the absence of PAR reduction, the simple channel estimate (8) had the worst BER performance (solid line with circles). Superimposed training with PAR reduction (SLM with $M = 8$, solid line with dots) had an SNR gain of 2 dB, but the improvement diminished at high SNRs due to the error floor. Superimposed training with the improved channel estimate (15) and SLM with $M = 8$ (solid line with squares) lowered the error floor but did not eliminate it. Next, we compare our proposed algorithm with PTAM (dash-dotted line) and the known channel case (dashed line). We observe from Fig. 2 that PTAM had a very good BER performance as it approaches the known

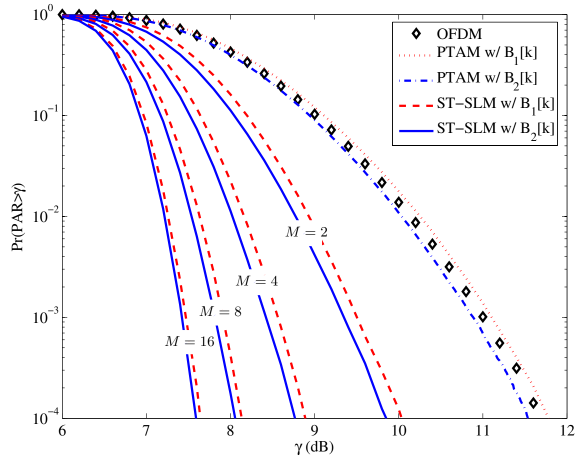

Figure 1. CCDF of the PAR in the proposed algorithm. ST stands for superimposed training here.

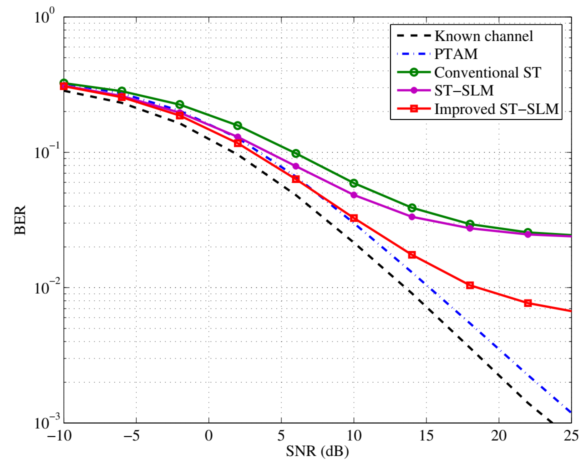

Figure 2. BER of the proposed algorithm. ST stands for superimposed training here.

channel case to within $1-2$ dBs. However, superimposed training results in an error floor at high SNRs because the information signal acts as the dominant source of "noise" during channel estimation. On the other hand, at medium to low SNRs (i.e., SNR$<$ 10 dB), our proposed algorithm performs similarly to or better than PTAM, the net benefit being the enhanced bandwidth efficiency.

6. CONCLUSIONS

It is well known in the literature that pilot tone assisted modulation (PTAM) with equi-spaced and equi-powered pilot tones is optimal for OFDM. However, the lack of bandwidth efficiency in dedicated training methods and the lack of power efficiency in OFDM are of concern. For the former, we propose to estimate the CSI using a superimposed training approach and offer two techniques to improve the CSI estimate. For the latter, we propose to combine SLM with superimposed training, and describe a simple blind detector to recover the SLM index. From the simulation results, we see that our proposed algorithm is very attractive at medium to low SNRs (e.g., SNR $<$ 10 dB): the BER is comparable to the dedicated training case, with the net gain being a higher bandwidth efficiency.

REFERENCES

[1] R. Negi and J. Cioffi, "Pilot tone selection for channel estimation in a mobile OFDM system," *IEEE Trans. Consumer Electron.,* vol. 44, pp. 1122-1128, Aug. 1998.

[2] 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.

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

[4] R. W. Bäuml, R. F. H. Fischer and J. B. Huber, "Reducing the peak-to-average power ratio of multicarrier modulation by selected mapping," *Elec. Lett.,* vol. 32, no. 22, pp. 2056-2057, Oct. 1996.

[5] M. Ghogho, D. C. McLernon, E. Alameda-Hernandez and A. Swami, "Channel estimation and symbol detection for block transmission using data-dependent superimposed training," *IEEE Signal Processing Letters*, vol. 12, no. 3, pp. 226-229, March 2005.

[6] G. T. Zhou, M. Viberg, and T. McKelvey, "A first-order statistical method for channel estimation," *IEEE Signal Processing Lett.*, vol. 10, no. 3, pp. 57-60, 2003.

[7] A. G. Orozco-Lugo, M. M. Lara, and D. C. McLernon, "Channel estimation using implicit training," *IEEE Trans. Signal Processing*, vol. 52, no. 1, pp. 240-254, January 2004.

[8] N. Chen and G. T. Zhou, "Superimposed training for OFDM: a peak-to-average power ratio analysis," *IEEE Trans. Signal Processing,* revised in April 2005.

[9] —, "Peak-to-average power ratio reduction in OFDM with blind selected pilot tone modulation," in *Proc. IEEE Intl. Conference on Acoustics, Speech, and Signal Processing*, vol. 3, pp. 845-848, Philadelphia, PA, March 2005.

[10] H. Ochiai, "Performance analysis of peak power and band-limited OFDM system with linear scaling," *IEEE Trans. Wireless Commun.*, vol. 2, no. 5, pp. 1055-1065, Sept. 2003.

[11] G. T. Zhou and L. Peng, "Optimality condition for selected mapping in OFDM," *IEEE Trans. Signal Processing*, revised in May 2005.

[12] J. Tellado, *Multicarrier Modulation with Low PAR - Applications to DSL and Wireless.* Kluwer Academic, 2000.