EFFECT OF CHANNEL ESTIMATION ON PAIR-WISE ERROR PROBABILITY IN OFDM

Khawza I. Ahmed, Cihan Tepedelenlioglu and Andreas Spanias

Department of Electrical Engineering
Arizona State University, Tempe, AZ 85287-5706
{khawza.ahmed, cihan, spanias}@asu.edu

ABSTRACT

Pair-wise error probability (PEP) is analyzed in the presence of channel estimation error (CEE) for orthogonal frequency division multiplexing (OFDM) in a quasi-static Rayleigh fading channel. Subcarriers in OFDM are grouped in an equi-spaced manner with the number of subcarriers in a group equal to the number of channel taps. One group is dedicated for training and the rest are for data transmission. A linear minimum mean square error (LMMSE) based channel estimation for coherent detection is considered in this paper. This enables the signal dependent CEE to be uncorrelated to the data. Since ML decoding employed with perfect CSI becomes suboptimal in the presence of CEE, an optimal ML decoding is derived. It is observed from our training-based PEP expression that the CEE does not reduce the diversity order but contributes to a loss of coding gain. Moreover, a code design criterion is established. To reduce the loss of coding gain, an optimal training scheme is developed based on the PEP expression. Loss of performance due to an imperfect channel estimate is quantified in terms of bounds on bit error probability (BEP) for high SNR. Our analytical findings are corroborated by simulation examples.

1. INTRODUCTION

With the increasing demand for high data rates, orthogonal frequency division multiplexing (OFDM) has become an integral component of various standards used in applications such as digital TV, wireless LAN and asymmetric digital subscriber lines (ADSL). The popularity of OFDM as an effective multicarrier technique for wireless transmission is due to its ability to transform a frequency-selective fading channel into parallel flat fading subchannels.

For coherent detection, the receiver needs to acquire the channel state information (CSI). The acquisition of CSI is conventionally performed by using the transmitted training sequences known to the receiver. The accuracy of the estimated CSI greatly affects the performance of the receiver. Therefore, it is imperative to assess the effect of channel estimation error (CEE) and to optimally design the training scheme to enhance the overall performance of a communication system.

A unified approach for analyzing the BEP performance of various uncoded OFDM systems using an arbitrary linear pilot assisted estimate of multiplicative channel response (CR) was presented in [1]. The authors obtained a closed-form BEP expression that enabled the derivation of the LMMSE type optimal linear CR estimate. In [2], for uncoded square QAM constellations, the effect of CEE on BEP was captured in terms of normalized MSE. The average BEP was approximated in [3] in the presence of intercarrier interference (ICI) under the assumption that the estimated CR and the CEE were almost uncorrelated. For uncoded BPSK transmission and a receiver equipped with one or several antennas in an uncorrelated scattering Gaussian multipath environment, an analytical expression for the BER performance was derived in [4] using the imperfect channel estimates in the decoder. Here, the OFDM symbol period was assumed to be of a much shorter duration than the coherence time. One OFDM block was used for channel estimation while the rest of the OFDM blocks were decoded based on the estimated channel. In [5], the BER was estimated using an analytical method in the presence of CEE that involved truncation of the union bound for a coded system. Although this method was demonstrated as being effective for a channel of more than two paths, there was no expression capturing the effect of CEE on the pairwise error probability.

Unlike the approaches in the literature, we investigate the effect of CEE on the coding and diversity gains of (pre)coded OFDM through PEP analysis. The PEP through the union-Chernoff bound provides a good approximation for BER at high SNR and is extensively used in performance analyses [6] to provide code design guidelines such as the order of achievable diversity. In this paper, the derived training-based PEP provides more insight about the diversity order, the loss of coding gain and the code design criteria in the presence of CEE. Furthermore, an optimal power distribution between training and data symbols is found, and bounds on the performance loss in dB due to CEE are derived for equal and optimal power allocation on the basis of our novel training-based PEP expression. While similar optimal training designs based on the maximization of a lower bound on average capacity were reported in [7] and [8] they do not capture the loss in BEP peformance due to CEE.

2. SIGNAL MODEL

Suppose we have an OFDM system with optimum equi-spaced subcarrier grouping [9]. The input output relationship of the m-th subgroup can be given by

$$\mathbf{x}_m = \mathbf{D}_{H_m}\mathbf{s}_m + \mathbf{w}_m, \qquad m = 1, \cdots, M \tag{1}$$

where M is the total number of groups each having K subcarriers making for a total of $P = MK$ subcarriers. $\mathbf{x}_m$, $\mathbf{s}_m$ and $\mathbf{w}_m$ are the received, the transmitted and the additive noise subblocks respectively with each of the vectors having K elements. $\mathbf{D}_{H_m}$ is a diagonal matrix with the i-th diagonal element $[\mathbf{D}_{H_m}]_{ii} = H(p_{m_i})$, where $H(p_{m_i}) := \sum_{l=0}^{L} h(l)\exp(-2\pi l p_{m_i}/P)$. $h(l)$ represents the l-th channel tap of the underlying L-th order FIR channel $\mathbf{h} := [h(0), h(1), \cdots, h(L)]^T$ and $\{p_{m_i} = (i-1)P/K + (m-1)\}_{i=1}^{K}$ denotes equi-spaced subcarrier indices. In general

The work in this paper was supported in part by NSF Career Grant No. CCR-0133841 and by Arizona State University.

$\mathbf{s}_m \in \mathcal{S}$, where $\mathcal{S}$ denotes a set of complex numbers of finite size constituting the alphabet of the constellation points for the uncoded case. K symbols $\{s(p_{m_i})\}_{i=1}^{K}$ are grouped together in $\mathbf{s}_m$ to be transmitted through the equi-spaced subcarriers indexed by $\{p_{m_i}\}_{i=1}^{K}$. We assume that $\mathbf{w}_m \sim \mathcal{CN}(\mathbf{0}, \sigma_w^2 \mathbf{I}_K)$ and $\mathbf{h} \sim \mathcal{CN}(\mathbf{0}, \sigma_h^2 \mathbf{I}_{L+1})$ where $\mathbf{I}_K$ is the identity matrix of size K. K is assumed to be equal to $L+1$ throughout this paper. Note that, for simplicity, we omit the OFDM time symbol index, as all processing is carried within a particular block. We will consider a scheme where the t-th group is used for channel estimation:

$$\mathbf{x}_t = \sigma_\tau \mathbf{D}_{H_t} \mathbf{s}_t + \mathbf{w}_t \tag{2}$$

where $t \in [1, M]$, σ_τ^2 denotes the average power of each of the training subcarriers and thus $\frac{1}{K}\text{tr}(\mathbf{s}_t \mathbf{s}_t^{\mathcal{H}}) = 1$. Also note that $\mathbf{s}_t$ are selected from an equi-powered constellation where each member in the constellation has normalized unit energy, e.g., PSK. This is because equi-spaced and equi-powered pilot symbols ensure optimal performance of the channel estimator [8],[10]. Also, choosing the size of the training subblock as $K = L+1$ ensures maximum throughput over the data subcarriers when optimal power allocation over the training and data symbols is performed [7],[8]. Here, we chose equi-spaced and equi-powered $L+1$ training symbols whose power will be optimized based on the training-based PEP expression. The remaining $(M-1)$ groups are used for the data transmission and the d-th group can be considered for the analysis of the performance:

$$\mathbf{x}_d = \sigma_\delta \mathbf{D}_{H_d} \mathbf{s}_d + \mathbf{w}_d \tag{3}$$

where $d \neq t$, $d \in [1, M]$, σ_δ^2 denotes the average power at each of data subcarriers and thus $\frac{1}{K}E\{\text{tr}(\mathbf{s}_d \mathbf{s}_d^{\mathcal{H}})\} = 1$. The training and data noise vectors $\mathbf{w}_t$ and $\mathbf{w}_d$ are independent and have the same variance σ_w^2. The transmit power is normalized to satisfy

$$\begin{aligned} \frac{1}{P}\left[\sigma_\tau^2 \text{tr}\left(\mathbf{s}_t \mathbf{s}_t^{\mathcal{H}}\right) + \sigma_\delta^2 (M-1) E\left\{\text{tr}\left(\mathbf{s}_d \mathbf{s}_d^{\mathcal{H}}\right)\right\}\right] &= 1 \\ \Rightarrow \frac{1}{M}\left[\sigma_\tau^2 + \sigma_\delta^2 (M-1)\right] &= 1 \end{aligned} \tag{4}$$

where the last equality is obtained from the relation $P = MK$. We can rewrite (2) and (3) as

$$\mathbf{x}_t = \sigma_\tau \mathbf{D}_{s_t} \mathbf{V}_t \mathbf{h} + \mathbf{w}_t \tag{5}$$

$$\mathbf{x}_d = \sigma_\delta \mathbf{D}_{s_d} \mathbf{V}_d \mathbf{h} + \mathbf{w}_d \tag{6}$$

where $\mathbf{D}_{s_t}$ and $\mathbf{D}_{s_d}$ are diagonal matrices with the (i,i)-th element as $[\mathbf{D}_{s_t}]_{ii} = s(p_{t_i})$, $[\mathbf{D}_{s_d}]_{ii} = s(p_{d_i})$, $\mathbf{v}^T(p_{t_i})$ and $\mathbf{v}^T(p_{d_i})$ are the i-th row of $\mathbf{V}_t$ and $\mathbf{V}_d$ respectively with $\mathbf{v}(p_{m_i}) := [\ 1, \exp(-2\pi p_{m_i}/P), \cdots, \exp(-2\pi L p_{mi}/P)]^T$. Also note that our choice of p_{m_i} corresponds to a grouping that periodically interleaves the subcarriers and hence yield $\mathbf{V}_t$ and $\mathbf{V}_d$ that are unitary: $\mathbf{V}_t \mathbf{V}_t^{\mathcal{H}} = \mathbf{V}_t^{\mathcal{H}} \mathbf{V}_t = K\mathbf{I}_K$ and also $\mathbf{V}_d \mathbf{V}_d^{\mathcal{H}} = \mathbf{V}_d^{\mathcal{H}} \mathbf{V}_d = K\mathbf{I}_K$.

3. MMSE CHANNEL ESTIMATION AND ML DECODING

Our goal is to derive the PEP when MMSE channel estimation is performed. Therefore, we will assume that $\hat{\mathbf{h}}_{\text{MMSE}}$ is computed from (5) using the knowledge of $\mathbf{s}_t$ and use $\hat{\mathbf{h}}_{\text{MMSE}}$ in (6) to decode the unknown data $\mathbf{s}_d$. Using the orthogonality relationship, it is straight forward to show that

$$\begin{aligned} \hat{\mathbf{h}}_{\text{MMSE}} &= \sigma_\tau \mathbf{V}_t^{\mathcal{H}} \mathbf{D}_{s_t}^{\mathcal{H}} \left(\frac{\sigma_h^2}{\sigma_\tau^2 \sigma_h^2 K + \sigma_w^2}\right) \mathbf{I}_K \mathbf{x}_t \\ &= K_1 \mathbf{h} + K_2 \mathbf{w}_t' \end{aligned} \tag{7}$$

where $K_1 = \frac{K\sigma_\tau^2 \sigma_h^2}{\sigma_\tau^2 \sigma_h^2 K + \sigma_w^2}$, $K_2 = \frac{\sqrt{K}\sigma_h^2 \sigma_\tau \sigma_w}{\sigma_\tau^2 \sigma_h^2 K + \sigma_w^2}$ and $\mathbf{w}_t' := \frac{1}{\sqrt{K}\sigma_w} \mathbf{V}_t^{\mathcal{H}} \mathbf{D}_{s_t}^{\mathcal{H}} \mathbf{w}_t$ with the assumption that $\mathbf{D}_{s_t}^{\mathcal{H}} \mathbf{D}_{s_t} = \mathbf{I}_K$. Thus $\mathbf{w}_t' \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_K)$. Note that $\hat{\mathbf{h}}_{\text{MMSE}} \sim \mathcal{CN}(\mathbf{0}, \sigma_{\hat{h}}^2 \mathbf{I}_K)$, with $\sigma_{\hat{h}}^2 = K_1^2 \sigma_h^2 + K_2^2$. Now let us rewrite (3) as

$$\mathbf{x}_d = \sigma_\delta \hat{\mathbf{D}}_{H_d} \mathbf{s}_d + \sigma_\delta \tilde{\mathbf{D}}_{H_d} \mathbf{s}_d + \mathbf{w}_d, \tag{8}$$

where $\hat{\mathbf{D}}_{H_d} = \text{diag}(\mathbf{V}_d \hat{\mathbf{h}}_{\text{MMSE}})$, and $\tilde{\mathbf{D}}_{H_d} = \mathbf{D}_{H_d} - \hat{\mathbf{D}}_{H_d}$. Commuting $\tilde{\mathbf{D}}_{H_d}$ with $\mathbf{s}_d$, we obtain

$$\mathbf{x}_d = \sigma_\delta \hat{\mathbf{D}}_{H_d} \mathbf{s}_d + \underbrace{\sigma_\delta \mathbf{D}_{s_d} \mathbf{V}_d ((1-K_1)\mathbf{h} - K_2 \mathbf{w}_t') + \mathbf{w}_d}_{:= \bar{\mathbf{w}}_{s_d}} \tag{9}$$

where the term associated with estimation error due to noise at the training phase and the noise at the data phase are lumped together and denoted as the signal dependent noise $\bar{\mathbf{w}}_{s_d}$. This $\bar{\mathbf{w}}_{s_d}$ can be shown to be uncorrelated with $\sigma_\delta \hat{\mathbf{D}}_{H_d} \mathbf{s}_d$ using the orthogonality principle of LMMSE. For a fixed $\mathbf{s}_d$, $\bar{\mathbf{w}}_{s_d}$ is zero mean Gaussian distributed with covariance matrix

$$\begin{aligned} \mathbf{R}_{\bar{w}_{s_d}} &:= E\left[\bar{\mathbf{w}}_{s_d} \bar{\mathbf{w}}_{s_d}^{\mathcal{H}}\right] \\ &= \underbrace{\sigma_\delta^2 ((1-K_1)^2 \sigma_h^2 M + K_2^2 M)}_{:= K_3} \mathbf{D}_{s_d} \mathbf{D}_{s_d}^{\mathcal{H}} + \sigma_w^2 \mathbf{I}_K. \end{aligned}$$

In decoding $\mathbf{s}_d$, we have at least two options. If there is no CEE, i.e., the channel is known, the ML decision rule is

$$\textbf{ML1:} \quad \arg\min_{\mathbf{s}_d} \| \mathbf{x}_d - \sigma_\delta \mathbf{D}_{s_d} \mathbf{V}_d \mathbf{h} \|^2 \tag{10}$$

since, in this case $\mathbf{h}$ is known and $\mathbf{w}_d$ is white. Hence a possibility is to use $\hat{\mathbf{h}}_{\text{MMSE}}$ as if it were the true channel, i.e, use it in place of $\mathbf{h}$ in (10). However, ML1 is not the ML estimator of $\mathbf{s}_d$ unless $\mathbf{R}_{\bar{\mathbf{w}}_{s_d}}$ is an identity matrix. In fact, the true ML estimator is **ML2**:

$$\arg\min_{\mathbf{s}_d} \left\{ \sum_{i=1}^{K} \left[\log\left[K_3 |s_{d_i}|^2 + \sigma_w^2\right] + \frac{|x_{di} - [\hat{\mathbf{D}}_{H_d}]_{ii} s_{d_i}|^2}{K_3 |s_{d_i}|^2 + \sigma_w^2} \right] \right\} \tag{11}$$

where x_{di} and s_{d_i} are the i-th element of $\mathbf{x}_d$ and $\mathbf{s}_d$ respectively. Notice that (11) will yield the same result as (10) if $\mathbf{s}_d$ has constant modulus entries that would make $\mathbf{R}_{\bar{\mathbf{w}}_{s_d}}$ independent of $\mathbf{s}_d$. As ML2 is more complex we will focus on ML1 for PEP analysis in the presence of CEE.

4. PEP IN THE PRESENCE OF CEE

In the presence of CEE, we have the signal model in (9). Therefore, using (10) for decoding gives us the following Chernoff bound for the PEP:

$$\begin{aligned} \text{P}\left(\mathbf{s}_d \to \mathbf{s}_d' | \hat{\mathbf{h}}_{\text{MMSE}}\right) &\leq \exp\left(-\frac{\sigma_\delta^2 \hat{\mathbf{h}}_{\text{MMSE}}^{\mathcal{H}} \mathbf{V}_d^{\mathcal{H}} \mathbf{D}_e^{\mathcal{H}} \mathbf{D}_e \mathbf{V}_d \hat{\mathbf{h}}_{\text{MMSE}}}{4\mathbf{e}^{\mathcal{H}} \mathbf{R}_{\bar{w}_{s_d}} \mathbf{e}}\right) \\ &\leq \exp\left(-\frac{\sigma_\delta^2 \hat{\mathbf{h}}_{\text{MMSE}}^{\mathcal{H}} \mathbf{V}_d^{\mathcal{H}} \mathbf{D}_e^{\mathcal{H}} \mathbf{D}_e \mathbf{V}_d \hat{\mathbf{h}}_{\text{MMSE}}}{4\beta_{\max}}\right) \qquad (12) \\ &= \exp\left(-\hat{\mathbf{h}}_{\text{MMSE}}^{\mathcal{H}} \mathbf{A}_e' \hat{\mathbf{h}}_{\text{MMSE}}\right) \qquad (13) \end{aligned}$$

where $\beta_{\max} := \left[K_3 \nu + \sigma_w^2\right]$ is the maximum eigenvalue of $\mathbf{R}_{\bar{w}_{s_d}}$, $\nu := \max\{|s_{d_i}|^2\}_{i=1}^{K}$, $\mathbf{D}_e := \mathbf{D}_{s_d} - \mathbf{D}_{s_d'}$, $\mathbf{e} := \sigma_\delta \hat{\mathbf{D}}_{H_d}(\mathbf{s}_d - \mathbf{s}_d') / \| \sigma_\delta \hat{\mathbf{D}}_{H_d}(\mathbf{s}_d - \mathbf{s}_d') \|$, a unit vector, and the inequality in (12) is obtained observing the fact that $\mathbf{e}^{\mathcal{H}} \mathbf{R}_{\bar{w}_{s_d}} \mathbf{e} \leq \beta_{\max}$. Equality in (13) is obtained with $\mathbf{A}_e' := \frac{\sigma_\delta^2 \mathbf{V}_d^{\mathcal{H}} \mathbf{D}_e^{\mathcal{H}} \mathbf{D}_e \mathbf{V}_d}{4\beta_{\max}}$. Thus the average PEP is bounded as

$$
\begin{aligned}
\mathrm{P}\left(\mathbf{s}_d \to \mathbf{s}'_d\right) &\le E_{\hat{\mathbf{h}}_{\text{MMSE}}}\left[\exp\left(-\hat{\mathbf{h}}^{\mathcal{H}}_{\text{MMSE}}\mathbf{A}'_e\hat{\mathbf{h}}_{\text{MMSE}}\right)\right] \\
&= \prod_{i=1}^{r}\frac{1}{1+\sigma^2_{\hat{h}}\lambda'_i} \\
&\le \left(\underbrace{\frac{\sigma^2_\delta K(K_1^2\sigma_h^2+K_2^2)}{4\left[K_3\nu+\sigma_w^2\right]}}_{f(\sigma_\tau,\sigma_\delta)}\right)^{-r}\prod_{i=1}^{r}\left|s_{d_i}-s'_{d_i}\right|^{-2} \quad (14) \\
&\approx \left(\frac{\sigma_h^2 K}{4\sigma_w^2}\right)^{-r}(Mg(\gamma))^{-r}\prod_{i=1}^{r}\left|s_{d_i}-s'_{d_i}\right|^{-2} \quad (15)
\end{aligned}
$$

where $\sigma^2_{\hat{h}} = (K_1^2\sigma_h^2+K_2^2)$, $\left\{\lambda'_i = \frac{\sigma^2_\delta K\left|s_{d_i}-s'_{d_i}\right|^2}{4\left[K_3\nu+\sigma_w^2\right]}\right\}_{i=1}^{r}$ is the i-th eigenvalue of $\mathbf{A}'_e$, $r \le (L+1)$ is the rank of $\mathbf{A}'_e$, γ is the percentage of total power employed for training, i.e., $\gamma := \sigma_\tau^2 K/P$ and $g(\gamma) := \gamma(1-\gamma)/(\nu-\gamma(\nu+1-M))$. Equation (15) can be obtained using $f(\sigma_\tau,\sigma_\delta) \approx \sigma_h^2 KMg(\gamma)/4\sigma_w^2$ at high SNR ($\sigma_w^2 \ll 1$) and is derived in Section 5. We include it here to show that the diversity order is $r \le (L+1)$, the rank of $\mathbf{A}'_e$, which is the same as $\mathbf{D}_e$. When there is no CEE, i.e., with $\hat{\mathbf{h}}_{\text{MMSE}} = \mathbf{h}$, we have, $K_1 = 1, K_2 = 0, K_3 = 0$ and $\sigma^2_{\hat{\delta}} = 1$, i.e., (14) becomes

$$
\mathrm{P}\left(\mathbf{s}_d \to \mathbf{s}'_d\right) \le \left(\frac{\sigma_h^2 K}{4\sigma_w^2}\right)^{-r}\prod_{i=1}^{r}\left|s_{d_i}-s'_{d_i}\right|^{-2} \quad (16)
$$

which is the familiar expression for the known channel case [6]. Comparing (15) with (16), it is apparent that although the diversity order r is not affected, there is loss of coding gain in the presence of imperfect channel estimates. However, the code design criterion that ensures maximum diversity, i.e, $r = L+1$ and minimum loss of coding gain remains the same and is given by the maximization of the product distance $\min_{\forall \mathbf{s}_d \neq \mathbf{s}'_d}\prod_{i=1}^{(L+1)}\left|s_{d_i}-s'_{di}\right|^2$. Thus the design of effective codes requires designing $\mathbf{D}_e$ such that it has full rank. Group linear constellation precoder (GLCP) $\boldsymbol{\Theta}$, designed in [9] is a good choice and with which $\mathbf{s}_d = \boldsymbol{\Theta}\tilde{\mathbf{s}}_d$, where $\tilde{\mathbf{s}}_d$ comes from a finite alphabet of constellation.

5. OPTIMAL TRAINING AND LOSS OF PERFORMANCE

We assume that the transmitter can allocate different power to training symbol and data, i.e., the σ_τ^2 and σ_δ^2 may not be equal. From (15), it is apparent that the loss in coding gain due to error in channel estimates can be optimized by choosing γ which allocates the total power to training and data. Thus, for a fixed total budget of power for an OFDM symbol, the power allocation among the training and data subgroups affect the BER significantly. If we allocate too much power for training, CEE will be reduced but detectability of the data will be susceptible to noise due to weak data SNR. On the other hand, lower power for training deteriorates the estimated channel quality that results in poor detection even at high data SNR. From (14), it is evident that PEP can be minimized by maximizing $f(\sigma_\tau,\sigma_\delta)$ with respect to σ_τ and σ_δ. Thus, the following optimization problem can be formulated to obtain the optimal value of σ_τ and σ_δ.

$$
\{\sigma_{\tau_{\text{opt}}},\sigma_{\delta_{\text{opt}}}\} = \arg\max_{\sigma_\tau,\sigma_\delta}\underbrace{\left(\frac{\sigma^2_\delta K(K_1^2\sigma_h^2+K_2^2)}{4\left[K_3\nu+\sigma_w^2\right]}\right)}_{f(\sigma_\tau,\sigma_\delta)} \quad (17)
$$

with the power constraint of (4). By using

$$
\sigma_\tau^2 = \frac{\gamma P}{K}\ ,\ \sigma_\delta^2 = \frac{(1-\gamma)P}{(M-1)K}, \quad (18)
$$

$f(\sigma_\tau,\sigma_\delta)$ can be maximized over $0 \le \gamma \le 1$ and the optimal $\gamma = \gamma_{\text{opt}}$ can be obtained. This can be done numerically. However, a closed form approximate expression of γ_{opt} for high SNR can be obtained. With the assumption $\sigma_w^2 \ll 1$ we get $K_1 \approx 1$, $K_2 \approx \frac{\sigma_w}{\sigma_\tau\sqrt{K}}$, $K_3 \approx \frac{\sigma_\delta^2\sigma_w^2}{\sigma_\tau^2}$. Thus,

$$
f(\sigma_\tau,\sigma_\delta) \approx \frac{\sigma_\delta^2 K\left(\sigma_h^2+\frac{\sigma_w^2}{\sigma_\tau^2 K}\right)}{4\left(\frac{\sigma_\delta^2\sigma_w^2}{\sigma_\tau^2}\nu+\sigma_w^2\right)} \approx \frac{\sigma_\delta^2\sigma_h^2 K}{4\sigma_w^2\left(\frac{\sigma_\delta^2}{\sigma_\tau^2}\nu+1\right)} \quad (19)
$$

and using (18) in (19), we get $f(\gamma) \approx \sigma_h^2 KMg(\gamma)/4\sigma_w^2$, where $g(\gamma) = \gamma(1-\gamma)/(\nu-\gamma(\nu+1-M))$. Differentiating $g(\gamma)$ the following expression of γ_{opt} at high SNR denoted as γ_∞ is obtained:

$$
\gamma_\infty = \frac{\sqrt{\nu}}{\sqrt{\nu}+\sqrt{M-1}}. \quad (20)
$$

This expression of optimal power allocation at high SNR turns out to be identical to the one in [7],[8] with $\nu = 1$, e.g., uncoded OFDM with unit energy PSK constellation, although their optimality criterion was a lower bound of average training-based capacity.

Comparing (15) and (16), we can quantify the loss of performance due to CEE for a particular value of γ at high SNR.

$$
\frac{\left[\mathrm{P}\left(\mathbf{s}_d \to \mathbf{s}'_d\right)\right]_{\text{perfect}}}{\left[\mathrm{P}\left(\mathbf{s}_d \to \mathbf{s}'_d\right)\right]_{\text{imperfect}}} \le \left[\frac{\sigma_\delta^2}{\frac{\sigma_\delta^2}{\sigma_\tau^2}\nu+1}\right]^r = \left[\frac{M\gamma(1-\gamma)}{\nu-\gamma(\nu+1-M)}\right]^r \quad (21)
$$

Therefore for equal power allocation, i.e., for $\gamma = \frac{1}{M}$, with $\sigma_\tau^2 = \sigma_\delta^2 = 1$, we have the following bound.

$$
\frac{\left[\mathrm{P}\left(\mathbf{s}_d \to \mathbf{s}'_d\right)\right]_{\text{perfect}}}{\left[\mathrm{P}\left(\mathbf{s}_d \to \mathbf{s}'_d\right)\right]_{\text{imperfect}}} \le \left[\frac{1}{\nu+1}\right]^r \quad (22)
$$

Similarly, for γ_∞, we reach the following bound.

$$
\frac{\left[\mathrm{P}\left(\mathbf{s}_d \to \mathbf{s}'_d\right)\right]_{\text{perfect}}}{\left[\mathrm{P}\left(\mathbf{s}_d \to \mathbf{s}'_d\right)\right]_{\text{imperfect}}} \le \left[M\left(\sqrt{\nu}+\sqrt{M-1}\right)^{-2}\right]^r \quad (23)
$$

Thus, using the union bounds on BEP, (22) and (23), it can be shown that channel estimation incurs no more than $10\log_{10}(\nu+1)$ dB and $10\log_{10}(M\left(\sqrt{\nu}+\sqrt{M-1}\right)^{-2})$ dB loss with equal power and optimal power allocation respectively from the performance when the channel is perfectly known. Notice that for uncoded PSK constellations, $\nu = 1$ and the loss is about 3 dB for the equal power training.

6. SIMULATION EXAMPLE

We demonstrate the effect of channel estimation and optimal power allocation with GLCP-OFDM [9] and uncoded OFDM. Here, the parameters are $P = 16, L = 1, K = L+1$ [9]. The channel is complex Gaussian with $\sigma_h^2 = 1/(L+1)$ and are fixed for one OFDM symbol. Both $\mathbf{s}_t$ and $\tilde{\mathbf{s}}_d$ are chosen from the normalized unit energy QPSK constellation. The precoder $\boldsymbol{\Theta}$ is obtained from Table I in [9] and normalized such that $\mathrm{tr}\{\boldsymbol{\Theta}\boldsymbol{\Theta}^{\mathcal{H}}\} = K$. Through simulations, ν is found to be 1.7071 for each of the code words, $\mathbf{s}_d$, in GLCP-OFDM. $f(\gamma) = f(\sigma_\tau,\sigma_\delta)$ in (17) and simulated BER graphs are plotted as a function of γ and shown in Fig. 1 for different SNR values. We observe that BER graphs are minimum around γ_{opt}, where $f(\gamma)$ is maximum. Thus, the effectiveness of the proposed optimal power distribution based on training-based PEP is verified. From (20) we get $\gamma_\infty = 0.3306$ that matches closely the γ_{opt} obtained numerically for different SNR validating the expression in (20). Fig. (2) shows the BER performance with respect to SNR, using the ML1 in (10) and ML2 in (11) for different values

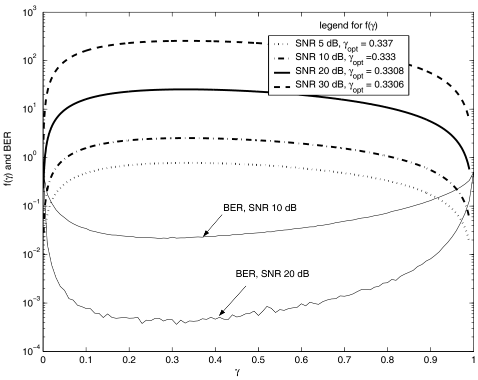

Fig. 1. $f(\gamma)$ and BER (marked with arrow) vs. γ for GLCP-OFDM

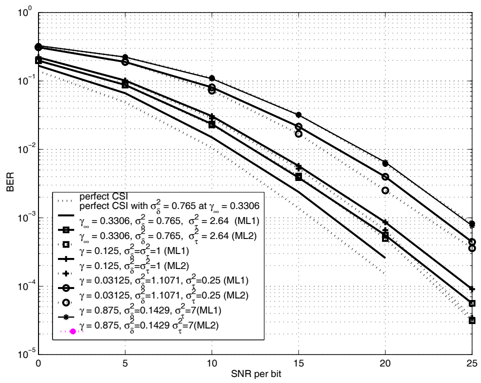

Fig. 2. Effect of channel estimation and power allocation on GLCP-OFDM

of γ. We observe that training with higher values as well as lower values of γ results in high BER. Training corresponding to γ_∞ yields the near-optimal performance. Comparing with the performance having perfect CSI, it is evident that the cost of acquiring CSI with optimally powered pilot is about 1.13 dB and the inaccurate CSI further incurs about 1.8 dB loss. Therefore, the total loss with optimal training is about 2.9 dB and for equal power training the loss is about 3.9 dB. The derived bounds in (22) and (23) give us about 3 dB and 4.3 dB loss respectively. Thus, the bounds are tight. It is also evident that at high SNR the true ML gives slight improvement over ML1 when channel estimate is relatively poor, e.g., with $\gamma = 0.03125$. For good quality channel estimates, e.g., with $\gamma = 0.875$, performance of ML1 and ML2 are similar. Fig. 3 illustrates the case of uncoded OFDM where $\nu = 1$ and consequently $\gamma_\infty = 0.274$. Both ML1 and ML2 perform the same when $\mathbf{s}_d$ has constant modulus entries. There is about 2 dB loss for optimal power allocation, whereas about 2.8 dB loss for equal power allocation. The bounds for these power allocation schemes as obtained from (23) and (22) are 2.2 dB and 3 dB respectively, demonstrating the tightness of the bound.

7. CONCLUSION

We have considered PEP analysis for OFDM in the presence of channel estimation error. It is observed that the use of LMMSE channel estimator does not reduce the diversity but causes loss of coding gain. For a given power budget, this loss is reduced, by optimal power allocation between equi-spaced pilot symbols and equi-spaced data symbols based on the upper bound on average training-based PEP expression. Also upper bounds of SNR (in dB) loss due to training are obtained for optimal and equal power

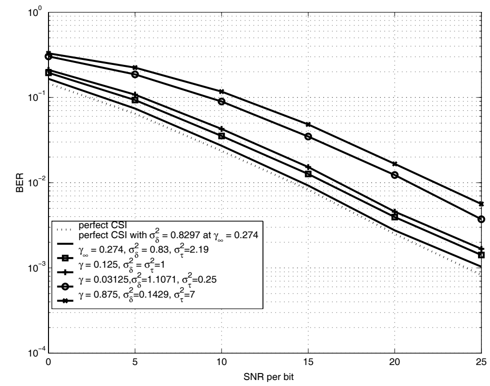

Fig. 3. Effect of channel estimation and power allocation on uncoded OFDM

allocation at high SNR.

8. REFERENCES

[1] M. Chang and Y. T. Su, "Performance analysis of equalized OFDM systems in Rayleigh fading," *IEEE Trans. Wireless Commun.*, vol. 1, no. 4, pp. 721–732, Oct. 2002.

[2] A. Leke and J. M. Cioffi, "Impact of imperfect channel knowledge on the performance of multicarrier systems," in *IEEE Global Telecommun. Conf.*, vol. 2, Sydney, NSW, Australia, Nov. 08-12 1998, pp. 951–955.

[3] H. Cheon and D. Hong, "Effect of channel estimation error in OFDM-based WLAN," *IEEE Commun. Lett.*, vol. 6, no. 5, pp. 190–192, May 2002.

[4] S. Furrer and D. Dahlhaus, "Mean bit-error rates for OFDM transmission with robust channel estimation and space diversity reception," in *Int. Zurich Seminar on Broadband Commun.*, Zurich, Switzerland, Feb. 19-21 2002, pp. 47–1–47–6.

[5] M. Sandell, S. Wilson, and P. Borjesson, "Performance analysis of coded OFDM on fading channels with non-ideal interleaving and channel knowledge," in *Veh. Technol. Conf.*, vol. 3, Phoenix, AZ, May 04-07 1997, pp. 1380–1384.

[6] Z. Wang and G. B. Giannakis, "Linearly precoded or coded OFDM against wireless channel fades?" in *Proc. of 3rd IEEE Workshop on SPAWC*, Taoyuan, Taiwan, R.O.C, Mar. 20-23 2001, pp. 267–270.

[7] 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.

[8] S. Ohno and G. B. Giannakis, "Capacity maximizing pilots for wireless OFDM over rapidly fading channels," in *Proc. Int. Symp. Signals, Syst., Electron*, Tokyo, Japan, July 24-27 2001, pp. 246–249.

[9] Z. Liu, Y. Xin, and G. B. Giannakis, "Linear constellation precoding for OFDM with maximum multipath diversity and coding gains," *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 416–427, Mar. 2003.

[10] S. Ohno and G. B. Giannakis, "Optimal traning and redundant precoding for block transmissoins with application to wireless OFDM," *IEEE Trans. Commun.*, vol. 50, no. 12, pp. 2113–2123, Dec. 2002.