HOW MANY KNOWN SYMBOLS ARE REQUIRED FOR LINEAR CHANNEL ESTIMATION IN OFDM?

Shuichi Ohno, Emmanuel Manasseh, and Masayoshi Nakamoto

Hiroshima University, 1-4-1 Kagamiyama Higashi-Hiroshima, 739-8527, JAPAN
E-mail: {ohnom,manassehjc,msy}@hiroshima-u.ac.jp

ABSTRACT

Training sequences for estimation of channel parameters have been well designed under the condition that the number of the unknown channel parameters is not greater than the number of the known symbols. Without the condition, we develop a method to obtain the optimal number of known pilot symbols for pilot-aided linear channel estimation in orthogonal frequency division multiplexing (OFDM). We derive an expression for channel estimation error covariance from which performance measures can be numerically evaluated to obtain the optimal number of pilot symbols. Using capacity lower bound as a performance measure, we demonstrate that for channels with exponential power profiles, it is better to estimate a truncated version of the channel response, which reveals the disadvantage of the existing design based on the condition.

Index Terms— Orthogonal frequency division multiplexing (OFDM), channel estimation, pilot symbols, pilot design.

1. INTRODUCTION

Since orthogonal frequency division multiplexing (OFDM) can achieve an effective high-rate wireless transmission, it has been adopted in many standards (IEEE802.11a/g, HIPERLAN 2, DAB, DVB, and more). In OFDM, training OFDM symbols (or OFDM preambles) or pilot symbols embedded in each OFDM symbol are utilized to obtain the channel state information (CSI). OFDM preambles are transmitted at the beginning of the transmission record, while pilot symbols (or tones) are embedded in each OFDM symbol where they are separated from the information symbols in the frequency-domain [1,2]. If the channel varies, OFDM preambles should be retransmitted to estimate the channel again to obtain reliable channel estimates. On the other hand, to track the relatively fast varying channel, pilot symbols embedded into every OFDM symbol facilitates channel estimation with each OFDM symbol. This is an application of pilot symbol assisted (or aided) modulations (PSAM) [3]. The main drawback of PSAM lies in the reduction of the transmission rate, especially when larger number pilot symbols are inserted in each OFDM symbol. Thus it is desirable to minimize the number of pilot symbols in each OFDM symbol to avoid excessive transmission rate loss.

When all subcarriers are available for transmission, OFDM preambles and pilot symbols have been well designed to enhance the channel estimation accuracy, see e.g., [4] and the references therein. Training sequence can be optimally designed in terms of several performance measures [1,5,6]. It has been found that equally spaced and equally powered pilot symbols are optimal for several cases.

Training sequences have been also designed, e.g., for multiple input multiple output (MIMO) system [7] and MIMO-OFDM system [8]. Even when some subcarriers are unavailable for transmission, which implies that equally spaced pilot symbols may not be available, pilot symbols have been designed [9–11]. However, most of the existing design methods assume that the number of the unknown channel parameters is not greater than the number of the known symbols. Is it really optimal if we set the number of the known symbols equal to the number of the unknown channel parameters?

We revisit the problem in attempt to address the question as to how many known pilot symbols are required for channel estimation in PSAM OFDM. We derive the channel estimation error covariance of linear minimum mean squared error (MMSE) estimator, without the assumption that the number of the unknown channel parameters is not greater than the number of the known symbols. Our expression is not analytically tractable except for some special cases but can be numerically computable if necessary information is provided. To show that the optimality of the pilot symbols whose number is equal to the number of the unknown channel coefficients is not guaranteed, we take the capacity lower bound of equally powered pilot symbols for example. Then, using the channel estimation error covariance, we evaluate the capacity lower bound as a function of the number of pilot symbols together with the power distribution between pilot and information-bearing data symbols to numerically obtain the optimal number of pilot symbols. Simulation results are provided to show that for channels with exponential power profile, the capacity bound for estimating of a truncated version of the channel response has higher capacity than the capacity bound for estimating of the whole channel response.
2. PRELIMINARIES AND SYSTEM MODEL

Let us consider point-to-point wireless orthogonal frequency division multiplexing (OFDM) transmissions over frequency-selective fading channels. Our discrete-time baseband equivalent channel is assumed to have a finite impulse response of maximum length $L$ and remains constant in at least one OFDM symbol, i.e., quasi-static. We denote the channel impulse response as $\{h_0, h_1, \ldots, h_{L-1}\}$. Since we deal with one OFDM symbol, we omit the OFDM symbol number for notational simplicity. The number of subcarriers of the OFDM symbol is assumed to be $N$.

At the transmitter, a serial symbol sequence $\{s_0, s_1, \ldots, s_{N-1}\}$ undergoes serial-to-parallel (S/P) conversion to be stacked into one OFDM symbol. Then, an $N$-point inverse discrete Fourier transform (IDFT) follows to produce the $N$ dimensional data, which is parallel-to-serial (P/S) converted. A cyclic prefix (CP) of length $N_p$ is appended to mitigate the multipath effects. The discrete-time baseband equivalent transmitted signals $u_n$ can be expressed in the time-domain as $u_n = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} s_k e^{-j2\pi nk/N}, \quad n \in [0, N-1]$.

We assume that $N_p > L$ so that there is no inter-symbol interference (ISI) between OFDM symbols. We also assume perfect timing synchronization between the transmitter and the receiver.

At the receiver, after removing CP, we apply discrete Fourier transform (DFT) to the received time-domain signal $y_n$ for $n \in [0, N-1]$ to obtain for $k \in [0, N-1]$ $Y_k = H_k s_k + W_k$, where $H_k$ is the channel frequency response at frequency $2\pi k/N$ given by $H_k = \sum_{l=0}^{L-1} h_l e^{-j2\pi kl/N}$ and the noise $W_k$ is assumed to be i.i.d. circular Gaussian with zero mean and variance $\sigma_w^2$, which is independent of the channel coefficients.

For coherent detection, we need to estimate the channel. In this paper, we adopt pilot symbol assisted (or aided) modulation (PSAM) [3], which embeds a few pilot symbols known to the receiver to enable channel estimation [1,2].

For channel estimation, we place $N_p \leq N_c$ non-zero pilot symbols (or tones) $\{p_1, \ldots, p_{N_p}\}$ at subcarriers $k_1, k_2, \ldots, k_{N_p} \in [0, N-1]$ ($k_1 < k_2 < \cdots < k_{N_p}$), which are known at the receiver. We denote the index set of pilot symbols as $K_p = \{k_1, \ldots, k_{N_p}\}$.

Let $\text{diag}(a)$ be a diagonal matrix with the vector $a$ on its main diagonal. Collecting the received signals having pilot symbols as $Y_p = [Y_{k_1}, \ldots, Y_{k_{N_p}}]^T$, we obtain

$$Y_p = D_{hp}p + W_p \quad (1)$$

where $D_{hp}$ is a diagonal matrix with its $n$th diagonal entry being $H_{k_n}$ such that $D_{hp} = \text{diag}(H_{k_1}, \ldots, H_{k_{N_p}})$ and $p$ is a pilot vector defined as $p = [p_1, \ldots, p_{N_p}]^T$.

From $Y_p$, we would like to estimate the channel frequency responses at information-bearing data subcarriers for equalization and decoding. Let us define an index set specifying the channel frequency responses to be estimated as $\mathcal{K}_d$.

In other words, $\mathcal{K}_d$ is the index set of data subcarriers.

If the number of unknown channel coefficients $L$ is not greater than the number of known pilot symbols $N_p$, the unknown channel coefficients can be easily estimated. Moreover, pilot symbols can be designed to enhance the system performance. If one can adopt equally spaced pilot symbols with equal power when $N_p < L$, then it can be analytically shown that the channel mean squared estimation error is minimized [1] and the lower bound on channel capacity is maximized [5, 6]. However, it has not been well studied what will happen if the number of unknown channel coefficients is greater than the number of the known symbols, i.e., $N_p < L$.

In this paper, we develop an expression for the mean squared channel estimation error covariance when $N_p < L$ and then resort to numerical evaluations to obtain the optimal number of equally powered pilot symbols as well as the optimal power distribution to pilot and data symbols.

3. MMSE IN PSAM OFDM

Let us define an $N \times N$ DFT matrix as $F$, whose $(m+1, n+1)$th entry is $e^{-j2\pi mn/N}$, an $N_p \times N$ matrix $F_p$ having $f_{kp}$ for $k_n \in K_p$ as its $n$th row, where $(\cdot)^T$ is the complex conjugate transpose. Then, we can express (1) as

$$Y_p = D_{p} F_p h + W_p \quad (2)$$

where the diagonal matrix $D_{p}$ and channel vector $h$ are respectively defined as $D_{p} = \text{diag}(p_1, \ldots, p_{N_p})$ and $h = [h_0, \ldots, h_{N-1}]^T$. In this section, we allow $h_l$ for $l \in [N_p, N-1]$ to take non-zero value for notational simplicity, although we have assumed that $h_l = 0$ for $l \in [N_p, N-1]$ to completely remove ISI in the previous section.

Let us consider the linear minimum mean squared error (MMSE) estimator, assuming that $E\{h_l\} = 0$ and $\sigma_{h_l}^2 = E\{|h_l|^2\}$ for $l \in [0, N-1]$, where $E\{\cdot\}$ stands for the expectation operator.

For $N_p < L$, one can estimate only $N_p$ channel coefficients with the linear MMSE estimation. On the other hand, it is often the case that the channel power profile is exponentially decreasing so that it can be parameterized as $\sigma_{h_l}^2 = ae^{-cl}$ for positive constants $a$ and $c$. Thus, it is reasonable to estimate the first $N_p$ coefficients of $h$, which is expressed as $h_l = [h_0, \ldots, h_{N_p-1}]^T$. Accordingly, we define an $N_p \times N_p$ matrix $F_{p, l}$ having the first $N_p$ columns of $F_p$ and an $N_p \times (N - N_p)$ matrix $F_{p, b}$ so that we can express $F_p = [F_{p, l} F_{p, b}]$. Then, with $h_b = [h_{N_p}, \ldots, h_{N-1}]^T$, we can re-express (2) as

$$Y_p = D_{p} F_{p, l} h_l + V_p \quad (3)$$

where $V_p = D_{p} F_{p, b} h_b + W_p$.

From (3), one can estimate $h_l$ since $D_{p} F_{p, l}$ is a non-singular matrix known to the receiver. Let a vector having
channel responses to be estimated, i.e., \( H_{k_n} \) for \( k_n \in \mathcal{K}_d \), be \( \hat{H}_d = [H_{k_1}, \ldots, H_{k_N-N_p}]^T \). Similar to \( F_p \), we define an \((N - N_p) \times N \) matrix \( F_d \) having \( f_{k_n}^H \) for \( k_n \in \mathcal{K}_d \) as its nth row, where \( k_n < k_n' \) if \( n < n' \). We also define two submatrices of \( F_d \) associated with \( h_t \) and \( h_b \) such as \( H_d = F_d h_t = F_d h_t h_b + F_d h_b \).

Now, we will derive the LMMSE estimate for \( \hat{H}_{d,t} = F_d h_t \), assuming that \( h_t \) and \( h_b \) with \( E\{h_t h_t^H\} = \mathbf{R}_t \) and \( E\{h_b h_b^H\} = \mathbf{R}_b \) are not correlated. We also assume that \( \mathbf{R}_t > 0 \) for simplicity of presentation. Then, since (2) is linear, the LMMSE estimate \( \hat{H}_{d,t} \) of \( H_{d,t} \) is obtained by

\[
\hat{H}_{d,t} = E\{H_{d,t} Y_p^H\} \left( E\{Y_p Y_p^H\}\right)^{-1} Y_p.
\]

If we define the estimation error vector \( E_{d,t} \) as \( E_{d,t} = \hat{H}_{d,t} - H_{d,t} \), then the correlation matrix \( \mathbf{R}_e \) of \( E_{d,t} \) can be expressed as

\[
\mathbf{R}_e = \mathbf{F}_{d,t} \left[ \mathbf{R}_t^{-1} + \mathbf{F}_{p,t} \hat{\mathbf{R}}_v^{-1} \mathbf{F}_{p,t}^H \right]^{-1} \mathbf{F}_{d,t}^H
\]

(4)

where \( \hat{\mathbf{R}}_v = \mathbf{F}_{p,d} \mathbf{R}_b \mathbf{F}_{p,b}^H + \sigma_v^2 (\mathbf{D}_p^H \mathbf{D}_p)^{-1} \). For given pilot symbols, and channel and noise statistics, one can numerically compute the LMMSE estimate and evaluate its error covariance matrix.

The estimation error covariance matrix (4) can be simplified for some special cases. Let us consider equally spaced and equally powered pilot symbols from their practical importance, which is formally defined as:

**C1** The \( N \) number of subcarriers in one OFDM symbol is a multiple of \( N_p \) such that \( N = N_p N_q \) for some non-zero integer \( N_q \); the index set \( \mathcal{K}_p \) for pilot symbols is \( \{n_q + N_p n_p|n_q \in [0, N_q - 1]\} \) for some \( n_p \in [0, N_p - 1]\); and \( |p_n|^2 = \sigma_p^2 (\geq 0) \) for all \( n \in [1, N_p] \).

Under C1, if the channel coefficients are independent of each other, then using properties of the DFT matrix, one can show that the estimation error variances at data subcarriers, which are the diagonal entries of \( \mathbf{R}_e \), are the same with

\[
\sigma_e^2 = \sum_{l=0}^{N_q-1} \left( \frac{\sum_{l=0}^{N_q} |s_{l+k}|^2}{N_p \sigma_p^2} + \frac{\sum_{l=0}^{N_q} |s_{l+n}|^2}{N_q \sigma_q^2} \right).
\]

(5)

We utilize the channel estimate \( \hat{H}_{d,t} \) to equalize the received signal. Thus, for \( k_n \in \mathcal{K}_d \) and \( n \in [1, N - N_p] \), we can express the received signal as

\[
Y_{k_n} = \hat{H}_{d,t} s_{k_n} + V_{k_n}
\]

(6)

with \( \hat{H}_{k_n} = \hat{H}_{d,t,n} \), where \( \hat{H}_{d,t,n} \) is the nth entry of \( \hat{H}_{d,t} \) and \( V_{k_n} \) is the effective noise expressed as \( V_{k_n} = (H_{k,n} - \hat{H}_{d,t,n}) s_{k_n} + W_{k_n} \) with \( H_{k,n} \) and \( H_{b,n} \) being respectively the nth entry of \( H_{d,t} \) and \( H_{d,b} \).

In general, \( E\{(H_{k,n} - \hat{H}_{d,t,n}) H_{b,n}^*\} \neq 0 \). However, if the noise variance is small enough, then from the orthogonal principle, \( E\{(H_{k,n} - \hat{H}_{d,t,n}) H_{b,n}^*\} \) is nearly zero. Thus, in the following, we consider the case when \( E\{(H_{k,n} - \hat{H}_{d,t,n}) H_{b,n}^*\} = 0 \) holds approximately, assuming that the data symbols are independent of channel coefficients and \( E\{|s_{k_n}|^2\} = \sigma_s^2 \) for all \( k_n \in \mathcal{K}_d \). Then, the variance of the effective noise \( V_{k_n} \) is found to be

\[
\sigma_{v_{k_n}}^2 = \left( |R_e|_{n,n} + \sum_{l=N_p}^{N-1} \sigma_{h_l}^2 \right) \sigma_s^2 + \sigma_w^2
\]

(7)

where \( |R_e|_{n,n} \) is the nth diagonal entry of \( R_e \). For a given \( H_{k_n} \), the instantaneous effective SNR at subcarrier \( k \) is \( |H_{k_n}|^2 \sigma_s^2 / \sigma_{v_{k_n}}^2 \), which can be utilized to design pilot symbols, e.g., to minimize bit error rate (BER).

Here, using the capacity lower bound as in [5, 6], we numerically obtain the optimal number of pilot symbols and the power distribution between pilot and data symbols as follows.

Let us assume that channel coefficients are complex Gaussian and normalize the sum of their variances to be one such that \( \sigma_h^2 = e^{-c l} / \left( \sum_{l=0}^{N_p} e^{-cl} \right) \) for \( l \in [0, N_p - 1] \) and \( \sigma_h^2 = 0 \) for \( l \in [N_p, N - 1] \). Then, \( H_{k,n} \) and \( H_{d,t,n} \) are also complex Gaussian. As in [5, 6], we can show that the capacity bound is given as \( \frac{1}{N + N_p} \sum_{k_n \in \mathcal{K}_d} \log \left( 1 + \rho_{k_n} |g|^2 \right) \), where the expectation is taken with respect to a complex Gaussian random variable \( g \) having zero mean and unit variance, and \( \rho_{k_n} = (1 - |R_e|_{n,n}) / \sigma_{v_{k_n}}^2 \).

**4. NUMERICAL EXAMPLES**

To validate our claim, we consider OFDM systems having different number of pilot symbols with subcarriers \( N = 64 \) and CP length \( N_{cp} = 16 \). Equally powered pilot symbols are inserted to be almost equally spaced. The signal-to-noise ratio (SNR) is defined as \( (N - N_p) \sigma_s^2 + N_p \sigma_p^2 / (N \sigma_w^2) \). For all simulations, SNR is fixed at 20dB. We consider the exponential power profiles having \( c = 1 \) and \( c = 2 \). Then, by

![Fig. 1. Maximum of capacity bounds for c = 1 and c = 2.](image-url)
using $\alpha (0 < \alpha < 1)$, the powers of pilot and data symbols are determined by $(N - N_p)\sigma_d^2 : N_p\sigma_p^2 = \alpha : 1 - \alpha$.

In Fig. 1, for each $N_p$, we plot the maximum of capacity bounds for $c = 1$ as well as $c = 2$. For $c = 1$, $N_p = 8$ is optimal, while for $c = 2$, $N_p = 4$. This is intuitively justified because the former has a longer tail than the latter. If one adopts the pilot designs in [5,6], then $N_p = 16$. Fig. 1 clearly shows that the pilot designs in [5,6] are not optimal at least in the number of pilot symbols.

To see the effects of the number of pilot symbols and $\alpha$ on BER performance, we compute empirical average BER, where BER for $10^6$ channel realizations are averaged and $(64 - 16) \cdot 10^4$ data symbols drawn from BPSK constellation are utilized. The received signals are equalized by one-tap ZF equalizers constructed from the estimated channels.

Fig. 2 shows the optimal BER for different numbers of pilot symbols. The BER curves are decreasing monotonically with the number of pilot symbols, because BER does not take care of the loss of the bandwidth due to the insertion of pilot symbols.

In Fig. 3, we evaluate the channel capacities of binary symmetric channels (BSC) having the error rates given in Fig. 2. We have the maximum at $N_p = 5$ and $N_p = 4$ for $c = 1$ and $c = 2$ respectively. The former is different from the maximum of capacity bounds at $N_p = 8$, while the latter is the same.

From the results above, it can be concluded that the number of pilot symbols and the power distribution between pilot and data symbols should be carefully set, depending on channel and noise statistics and modulation type.

5. REFERENCES


