

OPTIMAL PLACEMENT OF TRAINING FOR CHANNEL ESTIMATION AND TRACKING

Min Dong and Lang Tong

School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853
 {mdong, ltong}@ee.cornell.edu

ABSTRACT

The problem of training symbol placement for the estimation of block-fading channel is first considered. The channel is frequency selective with correlated random taps. The Cramér-Rao Bound (CRB) on the mean square error (MSE) of channel estimators is derived and minimized with respect to the pilot symbols and their placement. It is shown that, among all orthogonal pilot sequence placements, the minimum CRB is achieved by concentrating pilot and data powers in the middle of the packet. The placement of training for tracking time-varying channels is then addressed. Both flat and frequency selective fading environments are considered. We optimized training by minimizing the MSE of channel estimators chosen from optimal and sub-optimal tracking algorithms.

1. INTRODUCTION

One of the major challenges in achieving reliable transmission over time varying fading channels is channel estimation. Typically, pilot symbols are embedded in data packets to facilitate channel state acquisition. It has been shown recently that the placement of training can affect overall system performance [1, 2, 3].

For block fading channels with independent and identically distributed (i.i.d.) taps, we considered the problem of optimal design and placement of pilot symbols for the estimation of random channels [4, 5]. The performance criterion used was the Cramér-Rao Bound (CRB) on the mean-square error (MSE). It turns out that the optimal placement is to maximize the total pilot and data power at the midamble positions while satisfying the pilot symbol orthogonality condition. In practice, however, the use of pulse-shaping filters and the presence of scatters render the assumption of i.i.d. fading taps unrealistic. This motivates us to search for optimal placement for channels with correlated fading taps.

A more challenging problem, perhaps more important for mobile applications, is the placement of training symbols for channel tracking. Although the problem of channel tracking has been well studied for a long time [6], little attention has been paid on the optimal placement of training symbols in a data stream or in a packet. Recently, in [7], the authors compared capacities in adaptive and non-adaptive coding schemes for a time-varying Rayleigh fading channel with pilot symbol assisted modulation. However, how to place training symbols for optimal channel tracking is still an open problem.

This work was supported in part by the Army Research Office under Grant ARO-DAAB19-00-1-0507 and the Multidisciplinary University Research Initiative (MURI) under the Office of Naval Research Contract N00014-00-1-0564.

In this paper, we first consider the problem of optimal placement of pilot symbols for block fading channels. We obtain the expression of CRB as a function of pilot symbols and their placement. It is then shown that among all orthogonal sequences, the CRB is minimized by concentrating data and pilot powers in the midamble of the packet. Next, for Rayleigh fading channels, we present analytical and numerical results on the tracking performance as a function of the placement of training symbols and fading bandwidth.

2. PLACEMENT FOR SLOW FADING CHANNELS

2.1. Problem Statement

2.1.1. The Model

We consider the estimation of an SISO linear channel with random impulse response. A multipath fading environment is considered and the block-fading channel is assumed, *i.e.*, channel remains constant for one data packet length and changes to another independent value for the next packet length. We further assume that the estimation is performed within one transmitted packet.

Within one data packet, the channel is modeled by an FIR linear system with order L : $y_k = \sum_{i=0}^L h_i s_{k-i} + n_k$, where y_k is the received signal, $\mathbf{h} = [h_0, \dots, h_L]^T$ is the channel vector, s_k the input symbol and n_k the i.i.d. circular complex Gaussian noise with zero mean and variance σ_n^2 . We assume that each data packet consists of N data symbols and P pilot symbols, denoted as $\mathbf{s}_d = [s_d[1], \dots, s_d[N]]^T$ and $\mathbf{s}_p = [s_p[1], \dots, s_p[P]]^T$ respectively. The vector channel model is used for the entire packet corresponding to the $(N+P)$ symbols. Denoting $\mathbf{y} = [y_{N+P-L}, \dots, y_1]^T$, $\mathbf{s} = [s_{N+P}, \dots, s_1]^T$, we have

$$\mathbf{y} = \mathcal{T}(\mathbf{h})\mathbf{s} + \mathbf{n} = \mathcal{H}(\mathbf{s})\mathbf{h} + \mathbf{n}, \quad (1)$$

where $\mathcal{T}(\mathbf{h})$ is a Toeplitz matrix generated from \mathbf{h} and $\mathcal{H}(\mathbf{s})$ a Hankel matrix generated from input \mathbf{s}

$$\mathcal{T}(\mathbf{h}) = \begin{pmatrix} h_0 & \cdots & h_L & & \\ & \ddots & & \ddots & \\ & & h_0 & \cdots & h_L \end{pmatrix}_{(N+P-L) \times (N+P)}, \quad (2)$$

$$\mathcal{H}(\mathbf{s}) = \begin{pmatrix} s_{N+P} & \cdots & s_{N+P-L} \\ \vdots & \text{Hankel} & \vdots \\ s_{L+1} & \cdots & s_1 \end{pmatrix}. \quad (3)$$

We assume that data symbols are drawn from an i.i.d. sequence with zero mean and variance σ_d^2 . The power of the pilots is defined as $\sigma_p^2 = \frac{1}{P} \sum_{i=1}^P |s_p[i]|^2$.

The channel \mathbf{h} can be represented by

$$\mathbf{h} = \mathbf{G}\mathbf{v}, \quad (4)$$

where $\mathbf{G} = (\mathbf{g}_1, \dots, \mathbf{g}_q) \in \mathcal{C}^{(L+1) \times q}$ have q orthonormal columns and vector \mathbf{v} consists of q i.i.d. zero mean random variables with PDF $p_v(\cdot)$ and variance σ_v^2 . When $q < L + 1$, taps of channel \mathbf{h} are correlated with covariance $E[\mathbf{h}\mathbf{h}^H] = \sigma_v^2 \mathbf{G}\mathbf{G}^H$.

Specifically, the channel is the combination of the pulse shaping filter and propagation channel. Although propagation channel appears random changes from packet to packet, channel taps are correlated to each other in general due to the pulse shaping filter. Thus, channel \mathbf{h} appears random but is restricted in a certain subspace. Note that in a special case when $q = L + 1$, channel taps are i.i.d. with $E[\mathbf{h}\mathbf{h}^H] = \sigma_v^2 \mathbf{I}$. The optimal placement in this case was described in [4, 5].

Finally, the data \mathbf{s} , channel \mathbf{h} and noise \mathbf{n} are jointly independent and we assume that the channel is to be estimated using data \mathbf{y} for the entire packet, i.e., the estimation is semi-blind.

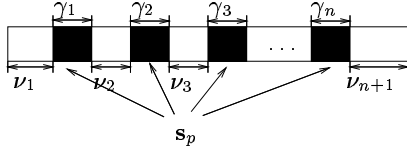


Fig. 1: An input sequence with multiple clusters

2.1.2. Pilot Symbol Placement

In general, the placement of n clusters of pilot symbols can be described by $\mathcal{P} = (\boldsymbol{\nu}, \boldsymbol{\gamma})$, where $\boldsymbol{\nu} = [\nu_1, \dots, \nu_{n+1}]$ is the data block length vector and $\boldsymbol{\gamma} = [\gamma_1, \dots, \gamma_n]$ the pilot cluster length vector, as illustrated in Figure 1. Constrained by the total number of data and pilot symbols, we have $\sum_{i=1}^{n+1} \nu_i = N$ and $\sum_{i=1}^n \gamma_i = P$. Moreover, for those placements starting with pilot symbols, $\nu_1 = 0$, and those end with pilot symbols, $\nu_{n+1} = 0$.

The input symbol vector can be decomposed into the pilot and data parts $\mathbf{s} = \bar{\mathbf{s}}_p + \bar{\mathbf{s}}_d$, where $\bar{\mathbf{s}}_p$ is obtained by setting the data part of \mathbf{s} to zero and similarly is $\bar{\mathbf{s}}_d$ obtained. This introduces a similar decomposition in the input symbol matrix

$$\mathcal{H}(\mathbf{s}) = \mathcal{H}(\bar{\mathbf{s}}_p) + \mathcal{H}(\bar{\mathbf{s}}_d), \quad (5)$$

and for convenience, we define their “autocorrelation” matrices as

$$\mathbf{R}_s \triangleq \mathcal{H}^H(\mathbf{s})\mathcal{H}(\mathbf{s}), \quad \mathbf{R}_{s_p} \triangleq \mathcal{H}^H(\bar{\mathbf{s}}_p)\mathcal{H}(\bar{\mathbf{s}}_p). \quad (6)$$

Note that quantities $\bar{\mathbf{s}}_p$, $\bar{\mathbf{s}}_d$ and their corresponding autocorrelation matrices are functions of the placement \mathcal{P} .

A pilot sequence with placement \mathcal{P} is called orthogonal if, given n clusters of pilot symbols and their placement \mathcal{P} , the autocorrelation matrix of pilot symbols \mathbf{R}_{s_p} is diagonal, i.e., \mathbf{R}_{s_p} satisfies

$$\mathbf{R}_{s_p} = \begin{pmatrix} \kappa_1^2 & & \\ & \ddots & \\ & & \kappa_{L+1}^2 \end{pmatrix}, \quad (7)$$

where κ_i^2 is the square of 2-norm of the i th column of $\mathcal{H}(\bar{\mathbf{s}}_p)$.

Finally we define the midamble and edge positions within one packet as shown in Figure 2. Midamble positions are defined as those within interval $[L + 1, N + P - L]$ while the rest part belong to edge positions.

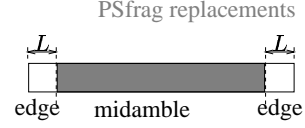


Fig. 2: Edge and midamble positions of one data packet

The CRB on the MSE is used as a performance measure for the placement of pilot symbols. Our objective is to find the optimal placement that minimizes the sum of the CRB for each channel tap.

2.2. The Optimal Placement of Pilot Sequence

2.2.1. The Cramér-Rao Bound

The CRB for random channels is used as a performance measure for the design and placement of pilot symbols [5]. The CRB for transformations of the random parameters is given by the following lemma.

Lemma 1 Let $\boldsymbol{\theta}$ be a random vector. Let $\mathbf{f} = \mathbf{g}(\boldsymbol{\theta})$, where \mathbf{f} , as a function of $\boldsymbol{\theta}$, is a random parameter vector to be estimated. Then the complex CRB for \mathbf{f} is

$$\boldsymbol{\Lambda}_{\mathbf{f}} = E\left[\frac{\partial \mathbf{g}}{\partial \boldsymbol{\theta}} \boldsymbol{\Lambda}_{\boldsymbol{\theta}} E^H\left[\frac{\partial \mathbf{g}}{\partial \boldsymbol{\theta}}\right]\right], \quad (8)$$

where $\boldsymbol{\Lambda}_{\boldsymbol{\theta}}$ is the complex CRB for $\boldsymbol{\theta}$.

Given the channel model in (4), the following theorem provides the expression of the CRB as a function of pilot symbols and their placement.

Theorem 1 Under the assumed model in Section 2.1 and regularity conditions [8, 9], the MSE of any channel estimator $\hat{\mathbf{h}}(\mathbf{y})$, defined as

$$\mathcal{M}(\hat{\mathbf{h}}) \triangleq E\{[\hat{\mathbf{h}}(\mathbf{y}) - \mathbf{h}][\hat{\mathbf{h}}(\mathbf{y}) - \mathbf{h}]^H\}. \quad (9)$$

satisfies the following inequality

$$\mathcal{M}(\hat{\mathbf{h}}) \geq \boldsymbol{\Lambda}(\mathcal{P}, \mathbf{s}_p) \triangleq \mathbf{G} \left(\frac{1}{\sigma_v^2} \mathbf{G}^H E\{\mathbf{R}_s\} \mathbf{G} + \rho_v^2 \mathbf{I} \right)^{-1} \mathbf{G}^H, \quad (10)$$

where $\boldsymbol{\Lambda}(\mathcal{P}, \mathbf{s}_p)$ is the complex CRB and $\rho_v^2 = E\left\{\left|\frac{\partial \ln p_v(\cdot)}{\partial v^*}\right|^2\right\}$.

From (10) we note that, for the channel with correlated taps, the CRB for channel estimators is a function of \mathbf{G} , but independent of specific channel realizations. Furthermore, for fixed \mathbf{G} , the CRB is only a function of pilot symbols and their placement \mathcal{P} , which makes the minimization of CRB with respect to \mathbf{s}_p and \mathcal{P} channel independent.

2.2.2. The Placement

In this section, we optimize the placement by minimizing the CRB. Our objective is to find \mathcal{P}_* and \mathbf{s}_{p*} , among all orthogonal pilot sequences, such that for all \mathcal{P} and \mathbf{s}_p satisfying the power constraint,

$$\text{tr} \boldsymbol{\Lambda}(\mathcal{P}_*, \mathbf{s}_{p*}) = \min_{\mathcal{P}, \mathbf{s}_p: \|\mathbf{s}_p\|^2 = P \sigma_p^2} \text{tr} \boldsymbol{\Lambda}(\mathcal{P}, \mathbf{s}_p). \quad (11)$$

Given n clusters of pilot symbols, as shown in Figure 1, with the placement of the clusters specified by $\mathcal{P} = (\boldsymbol{\nu}, \boldsymbol{\gamma})$, we have the following theorem about the optimal placement for orthogonal pilot sequences defined in (7).

Theorem 2 Assume $P \geq 2L + 1$. Among all orthogonal pilot sequences, i.e., $\{\mathbf{s}_p : \mathbf{R}_{s_p} \text{ is diagonal}\}$, the optimal placement is given by

$$\mathcal{P}_* = \{(\nu, \gamma) : n > 2; \nu_1 = \nu_{n+1} = 0; \gamma_1 = \gamma_n = L\}. \quad (12)$$

$$\mathbf{s}_{p_*} = \{\mathbf{s}_p : s_p[i] = 0, \forall i \in [1, L] \cup [P - L + 1, P]; \mathbf{R}_{s_p} = P\sigma_p^2 \mathbf{I}\}. \quad (13)$$

The minimum CRB is given by

$$\begin{aligned} \text{tr}\mathbf{\Lambda}(\mathcal{P}_*, \mathbf{s}_{p_*}) &= \min_{\substack{\mathcal{P}, \mathbf{s}_p : \|\mathbf{s}_p\|^2 = P\sigma_p^2, \\ \mathbf{R}_{s_p} \text{ diagonal}}} \text{tr}\mathbf{\Lambda}(\mathcal{P}, \mathbf{s}_p) \\ &= \frac{q\sigma_n^2}{N\sigma_d^2 + P\sigma_p^2 + \rho_v^2\sigma_n^2}. \end{aligned} \quad (14)$$

Proof: Refer to [4].

Theorem 2 states that the optimal placement of orthogonal pilot sequences calls for setting zeros to symbols at two ends of the packet and putting other pilots in the middle in such a way that they satisfy the orthogonality condition. For a sufficient number of pilot symbols, i.e., $P \geq 2L + 1$, Theorem 2 indicates that, for channels with correlated taps, again, concentrating all the data and pilot powers in the midamble positions leads to the minimum CRB. Notice that although there is no specification of placement for those midamble pilots, as long as the orthogonality condition $\mathbf{R}_{s_p} = P\sigma_p^2 \mathbf{I}$ is satisfied, it is easier to meet this requirement when multiple clusters are used.

In many communication systems, pilot symbols with Constant Modulus property are used, i.e., $|s_p[i]|^2 = \sigma_p^2, i = 1, \dots, P$. Following Theorem 2 and using the similar proof, we see that, for $\sigma_p^2 > \sigma_d^2$, among all orthogonal pilot sequences with their placement \mathcal{P} , the optimal placements are given by placing all pilot symbols in the midamble positions:

$$\mathcal{P}_* = \{(\nu, \gamma) : \nu_1, \nu_{n+1} \geq L\}, \quad (15)$$

and the minimum CRB is given by

$$\begin{aligned} \text{tr}\mathbf{\Lambda}(\mathcal{P}_*, \mathbf{s}_{p_*}) &= \min_{\substack{\mathcal{P}, \mathbf{s}_p : |s_p[i]|^2 = \sigma_p^2, \\ \mathbf{R}_{s_p} \text{ diagonal}}} \text{tr}\mathbf{\Lambda}(\mathcal{P}, \mathbf{s}_p) \\ &= \frac{q\sigma_n^2}{(N - L)\sigma_d^2 + P\sigma_p^2 + \rho_v^2\sigma_n^2}. \end{aligned} \quad (16)$$

Note that our results of optimal placement are confined in searching among all possible orthogonal pilot sequences. It does not imply that this placement minimizes $\text{tr}\mathbf{\Lambda}(\mathcal{P}, \mathbf{s}_p)$ for all choice of pilot sequences. Indeed, in general, the placement that gives the minimum $\text{tr}\mathbf{\Lambda}(\mathcal{P}, \mathbf{s}_p)$ depends heavily on each specific realization of pilot sequence.

3. PLACEMENT FOR CHANNELS TRACKING

3.1. Problem Statement

3.1.1. The Model

Within one data packet, the frequency selective fading channel with order L can be modeled as an FIR filter with time-varying coefficients:

$$y_k = \sum_{i=0}^L h_k[i] s_{k-i} + n_k, k = 1, \dots, N + P,$$

where $\mathbf{h}_k = [h_k[0], \dots, h_k[L]]^T$ is the channel state vector at time k . Elements in \mathbf{h}_k and n_k are independent complex circular Gaussian random variables with zero mean and variance σ_h^2 and σ_n^2 , respectively. Note that, as a special case when $L = 0$, we have the Rayleigh flat fading model.

To characterize the dynamics of the channel, we assume a first-order Gauss-Markov process, and each channel tap fades independently with the same correlation coefficient. Therefore, fading process can be assumed as a vector Gauss-Markov model:

$$\mathbf{h}_k = a\mathbf{h}_{k-1} + \mathbf{u}_k, \quad (17)$$

where $\mathbf{u}_k \sim \mathcal{CN}(0, (1 - a^2)\sigma_h^2 \mathbf{I})$ is the driving noise with the \mathbf{u}_k 's independent, and a the correlation coefficient that may vary between zero to one according to the fading channel bandwidth f_m (Doppler spread). For transmission over bandwidth W and decorrelation level ϕ that defines the coherence time $T_c (= 1/f_m)$, a is determined by $a^{\frac{W}{f_m}} = \phi$ [7]. Finally, the data power σ_d^2 is defined the same as before. We assume pilot symbols to be of the same energy, i.e., $|s_p[i]|^2 = \sigma_p^2$.

3.1.2. The Placement of Training

Since the channel state changes from symbol to symbol, it is fair to confine placements to the type of periodical placement with equal training cluster sizes, as shown in Figure 3. The placement of training can be characterized by period T , where pilot clusters are inserted at the beginning of each interval of length T . We assume that $N = rP$, where r is an integer. Note that $1/r$ is the ratio of pilot to data symbols. For fixed r , it is not hard to see that given the period T , the pilot cluster size P_T is then given by $P_T = \frac{T}{r+1}$. Note also that when P_T varies from 1 to P , the placement changes from single pilot periodical placement to single cluster placement.

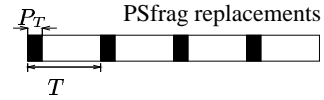


Fig. 3: An input sequence with period T

3.1.3. Channel Tracking Algorithms

The problem of channel tracking has been widely studied. Among all linear tracking algorithms, the Kalman filter has long been known to be optimal in terms of minimizing the Bayesian MSE for each estimator of channel state \mathbf{h}_k . However, this algorithm needs the knowledge of fading correlation coefficient a , which may not be available at the receiver end. On the other hand, the Least-Square (LS) algorithm as a sub-optimal deterministic method requires little knowledge of channel fading characteristics. We consider both of them as optimal and sub-optimal algorithms for channel tracking.

In this paper, we are interested in optimizing the placement of training using above two tracking algorithms. The performance criterion is the average (steady-state) MSE for one period T , defined as

$$\mathcal{E}(T) = \lim_{t \rightarrow \infty} \frac{1}{T} \sum_{i=1}^T E\{\|\hat{\mathbf{h}}_{t+i} - \mathbf{h}_{t+i}\|^2\}, \quad (18)$$

where $\mathbf{h}_i, i = 1, \dots, T$ is the i th channel state vector in the period T . Our objective is to find an optimal period T (therefore, P_T) that minimizes the average MSE of channel estimators

$$T_* = \arg \min_T \mathcal{E}(T). \quad (19)$$

3.2. The Optimal Placement

3.2.1. Rayleigh Flat Fading Channels

We first optimize the placement for channel tracking by the LS method. For the placement of training as shown in Figure 3 with period T , we have the following estimation procedures:

1. During training period within each pilot cluster, the receiver performs a LS estimation procedure yielding

$$\hat{h}_{lT+k} = \frac{y_{lT+k}}{s_p^{lT}[k]}, \quad k = 1, \dots, P_T; l = 0, 1, \dots, \frac{N+P}{T}-1. \quad (20)$$

where $s_p^{lT}[k]$ is the k th pilot symbols in the l th cluster.

2. During data transmission period, the receiver assume the same channel state as the last updated estimate from the training period, i.e.,

$$\hat{h}_{lT+k} = \hat{h}_{lT+P_T}, \quad k = P_T + 1, \dots, T. \quad (21)$$

We now optimize the period T by minimizing $\mathcal{E}(T)$, and have the following theorem.

Theorem 3 *For the Rayleigh flat fading channel model in (17) with the LS channel tracking method, under the assumption that $N = rP$,*

1. $\mathcal{E}(T)$ monotonically increases with T .
2. The single pilot periodical placement is optimal, i.e.,

$$T_* = 1 + r, \quad \text{and} \quad P_{T_*} = 1. \quad (22)$$

The minimum $\mathcal{E}(T)$ is given by

$$\mathcal{E}(T_*) = \frac{\sigma_n^2}{\sigma_p^2} + \frac{2r\sigma_u^2}{(1+r)(1-a^2)} \left[1 - \frac{a}{1-a} \frac{1-a^r}{r} \right]. \quad (23)$$

Theorem 3 indicates that for flat fading channels, pilot symbols should evenly scatter through the packet. This allows the receiver keeps updating channel information as frequently as possible, which reduces the performance degradation to the minimum.

We next consider the optimal tracking method: the Kalman filter. We assume the fading correlation coefficient a is known. The channel estimation is performed by the following two step procedure:

1. During training period within each pilot cluster, the receiver uses the Kalman filter update, yielding the MMSE update for each channel state

$$M[lT+k] = \frac{\sigma_n^2(a^2 M[lT+k-1] + \sigma_u^2)}{\sigma_n^2 + (a^2 M[lT+k-1] + \sigma_u^2)\sigma_p^2}, \quad k = 1, \dots, P_T, \quad (24)$$

where $M[lT+k]$ is the MMSE for \hat{h}_{lT+k} based on all the previous observations.

2. During data transmission period, the receiver predicts the channel state based on the last Kalman filter update in the training period

$$\hat{h}_{lT+P_T+k} = a^k \hat{h}_{lT+P_T}, \quad k = 1, \dots, T - P_T. \quad (25)$$

It can be seen that, although initial value $M[0]$ will affect $M[k]$ for the beginning updates, after several periods of iterations, the system tracking procedure resembles steady-state process. $M[lT+P_T]$ can then be viewed as constant for all l , and we denote it as M_{P_T} . The $\mathcal{E}(T)$ for period T is then given by

$$\mathcal{E}(T) = \frac{\sigma_u^2 r}{(1-a^2)(r+1)} - \frac{a^2}{1-a^2} \frac{1-a^{\frac{2r}{r+1}T}}{T} \left(\frac{\sigma_u^2}{1-a^2} - M_{P_T} \right) + \frac{1}{T} \sum_{k=1}^{P_T} M[lT+k]. \quad (26)$$

Since it is difficult to obtain the closed form for the transient state of the Kalman filter MSE update, we numerically calculate $\mathcal{E}(T)$ and compare it for different period T .

Numerical Result

Figure 4 shows the variation of $\mathcal{E}(T)$ with training cluster size P_T at SNR = 10dB for the Rayleigh flat fading channel of different correlation coefficient a (thus, f_m) under the Kalman filter(left) and LS method(right). Transmission bandwidth $W = 10^4$ Hz. The ratio of pilot to data symbols ($1/r$) is fixed at 25% for total amount of 1000 symbols in one packet. $\sigma_d^2 = \sigma_p^2 = 1$, the channel variance is 1. It should be noted that for some values of T , it might not be possible to place training periodically, we exclude these cases and only consider those possible values. Each curve corresponding to different fading level according to the channel bandwidth f_m from 10Hz to 400Hz, which corresponding to channel changes slow to fast. For both algorithms, we see that $\mathcal{E}(T)$ monotonically decreases with T , the minimum $\mathcal{E}(T)$ is achieved at $P_T = 1$ ($T = 5$). This numerical result shows that the single pilot periodical placement is again optimal for the Kalman filter method. Furthermore, there is a significant gain by placing training optimally, especially when channel changes relatively slow. On the other hand, for a fixed $\mathcal{E}(T)$ level, we compared the maximum allowable pilot cluster size under different f_m . It is shown that bigger f_m (fast changing) mandates smaller pilot cluster size and more frequent training, while smaller f_m (slow changing) allows larger pilot cluster with low training frequency.

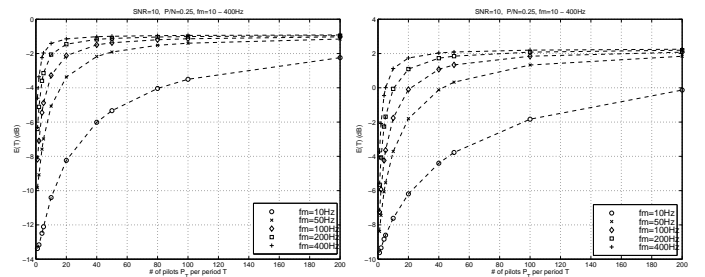


Fig. 4: $\mathcal{E}(T)$ vs. pilot cluster size P_T for the Rayleigh flat fading channel. SNR = 10dB. $P = 200$, $N = 800$. $W = 10^4$ Hz. Left: the Kalman filter; Right: the LS algorithm.

3.2.2. Frequency Selective Fading Channels

We next consider the placement of training for tracking the channel with order L . All channel taps are assumed to have the same correlation coefficient a . Assuming the receiver has no knowledge about a , for each period, the channel state is updated using the training at

the beginning of the period by the LS algorithm:

$$\hat{\mathbf{h}}_{lT} = [\mathcal{H}(\mathbf{s}_p^{lT})^H \mathcal{H}(\mathbf{s}_p^{lT})]^{-1} \mathcal{H}(\mathbf{s}_p^{lT})^H \mathbf{y}_{lp}, \quad l = 0, 1, \dots, \frac{N+P}{T}-1, \quad (27)$$

where $\mathcal{H}(\mathbf{s}_p^{lT})$ is the pilot symbol matrix of the l th cluster and \mathbf{y}_{lp} is the received data based on the pilot cluster; the channel estimate $\hat{\mathbf{h}}_{lT}$ is used for the entire l th period. In this case, the pilot cluster size should be no smaller than $2L + 1$.

For the optimal tracking method, the vector Kalman filter is used during training periods. The estimation procedure is similar as in (24) and (25). Note that in this case, to conduct the Kalman filter update of the channel, the minimum pilot cluster size should be $L + 1$.

Again, due to the difficulty of finding a closed form expression of $\mathcal{E}(T)$, numerical evaluation is used to find optimal placement.

Numerical Result

Figure 5 shows the placement result for the Kalman filter. The left of the Figure gives the $\mathcal{E}(T)$ vs. pilot cluster size P_T for different f_m at $SNR = 20dB$. The channel order is $L = 3$, and 20% training in one packet. From the figure, we see that the optimal cluster size is $P_{T*} = 7$. We also searched P_{T*} for different L , the results show that the optimal size P_{T*} is chosen between two values: $P_{T*} \in \{2L + 1, 2L + 2\}$. It appears that for the frequency selective channel, with low percentage of training, the optimal cluster size is connected to L with pattern $2L + 1$ or $2L + 2$. The right of the Figure shows the $\mathcal{E}(T)$ vs. pilot cluster size P_T for different percentage of training. It is shown that for higher percentage of training, the optimal size P_{T*} reduces to $L + 1$, the minimum required cluster size for tracking.

Figure 6 shows the variation of $\mathcal{E}(T)$ as T changes when using the LS method. Each curve is corresponding to a different f_m . The channel order is $L = 2$ and 20% of training is used in each packet. We see that there exists an optimal period T_* for each f_m , however the value of T_* does not have a fixed pattern, unlike in the Kalman filter case.

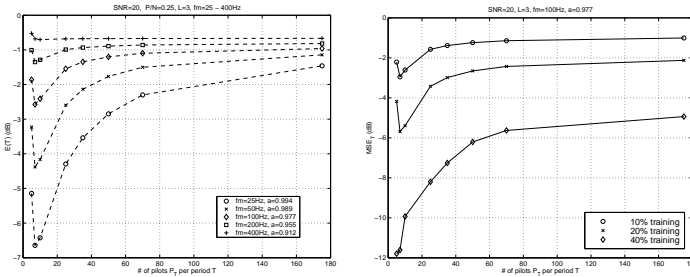


Fig. 5: $\mathcal{E}(T)$ vs. P_T for the frequency selective fading channel using the Kalman filter. $SNR = 20dB, W = 10^4 Hz$, $L = 3$. Left: under different f_m , $P = 200$, $N = 800$. Right: under different percentage of training.

4. CONCLUSION

In this paper, we considered the problem of optimal placement of training symbols for channel estimation and tracking. We extended previous results to channels with correlated taps. We have obtained the optimal placement of orthogonal pilot sequences for by minimizing the CRB of channel coefficients. Our results have shown

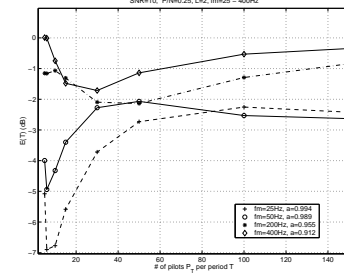


Fig. 6: $\mathcal{E}(T)$ vs. P_T for the frequency selective fading channel using the LS method. $SNR = 20dB, W = 10^4 Hz$, $L = 2$.

that for sufficient amount of pilot symbols, concentrate the powers in the midamble positions of the data packet leads to the minimum CRB. We then considered the problem of optimal pilot placement for channel tracking. In the flat fading scenario, from both analytical and numerical results, it is shown that single pilot periodical placement provides the best channel tracking performance. For the frequency selective channel, our numerical results shows that the optimal pilot cluster size has closed connection with channel order L . Moreover, the percentage of training affects the choice of the optimal cluster size.

5. REFERENCES

- [1] S. Adireddy and L. Tong. "Detection with Embedded Known Symbols: Optimal Symbol Placement and Equalization". In *Proc. ICASSP'00 Conf.*, Istanbul, Turkey, June 2000.
- [2] S. Adireddy and L. Tong. "Optimal Embedding of Known Symbols". In *Proc. 34th Conf. Information Sciences and Systems*, volume 1, Princeton, NJ, March 2000.
- [3] B. Hassibi and B. Hochwald. "How much Training is Needed in Multiple-Antenna Wireless Links". *Submitted to IEEE Trans. Information Theory*, August 2000.
- [4] M. Dong and L. Tong. "Optimal Design and Placement of Pilot Symbols for Channel Estimation". *Submitted to IEEE Trans. Signal Processing*.
- [5] M. Dong and L. Tong. "Optimal Design and Placement of Pilot Symbols for Channel Estimation". In *Proc. 2001 International Conference on Acoustics, Speech, and Signal Processing*, Salt Lake City, Utah, May 2001.
- [6] S. Haykin. *Adaptive Filter Theory*. Prentice-Hall, Englewood Cliffs, NJ., 1996.
- [7] M. Medard, I. Abou-Faycal, and U. Madhowand. "Adaptive Coding with Pilot Signals". In *38th Annual Allerton Conference on Communication, Control, and Computing*, October 2000.
- [8] H.L. Van Trees. *Detection, Estimation and Modulation Theory*, volume 1. Wiley, New York, 1968.
- [9] E. Weinstein and A. Weiss. "A General Class of Lower Bounds in Parameter Estimation". *IEEE Trans. Inform. Theory*, 34(2):338–342, March 1988.