TRAINING PLACEMENT FOR TRACKING FADING CHANNELS

Min Dong and Lang Tong

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

Brian M. Sadler

Army Research Laboratory
Adelphi, MD 20783
bsadler@arl.army.mil

ABSTRACT

The problem of training symbol placement in data packets for channel tracking is considered, where the channel is time-varying Rayleigh flat fading. We use the minimum mean-square error (MMSE) estimator for channel tracking. A minmax approach is considered. We optimize the placement by minimizing the maximum MSE over a packet. It is shown that training symbols should be scattered throughout the packet with equal space to achieve the optimal performance.

1. INTRODUCTION

To facilitate channel acquisition and tracking, training symbols (also referred to as pilot symbols) known to the receiver are often embedded in the data stream. Traditionally, the design of training scheme, especially the placement of training symbols, is based primarily on the simplicity of receiver implementation. In [2], the author analyzed the pilot symbol assisted modulation (PSAM) under flat Rayleigh fading in terms of bit error rate and discussed effects of pilot symbol spacing and Doppler spread. In [5], the authors compared capacities in adaptive and non-adaptive coding schemes for a time-varying Rayleigh fading channel under PSAM. Both papers assumed a specific placement of training scheme where pilot symbols are inserted periodically in the data.

It has been shown recently that the optimized placement of training symbols enhances overall system performance from both information theoretical and estimation theoretical perspectives [3, 1, 7, 8]. The gain is especially evident for time varying channels. In [4], we have optimized the training placement for the transmission of infinite data streams. *The Kalman Filter* and the Least Square algorithms are considered as channel tracking schemes. It is shown that single pilot periodic placement is optimal for both tracking strategies. In [6], from the channel capacity viewpoint, the optimal training symbol spacing and power allocation were analyzed in PSAM at high SNR and the packet length being infinity.

In this paper, we consider the problem of optimal training placement for the transmission of finite data packets over time-varying flat fading channels. The minimum mean-square error (MMSE) channel estimator is used for channel tracking. The presence of training symbols in the data stream makes MMSE of the channel estimator time varying. Here we consider a minmax approach where the maximum MSE as a function of placement is minimized.

This criterion is particularly relevant for receivers using symbol-by-symbol techniques. At high SNR, it is shown that packets should start and end with data symbols, and the Quasi Periodic Placement (QPP) of training symbols with cluster size one (QPP-1) is optimal.

2. PROBLEM STATEMENT

2.1. The Model

We consider the estimation of a time-varying Rayleigh flat fading channel. The discrete-time system model is given by

$$y_k = h_k s_k + n_k, \qquad (1)$$

where y_k is the received signal, s_k is the transmitted symbol, $n_k \overset{i.i.d.}{\sim} \mathcal{CN}(0, \sigma_n^2)$ the additive complex Gaussian noise and $h_k \sim \mathcal{CN}(0, \sigma_h^2)$ the channel state at time k.

To characterize the dynamics of the channel, we assume that the fading process is a first-order Gauss-Markov process:

$$h_k = \alpha h_{k-1} + u_k, \qquad (2)$$

where $u_k \overset{i.i.d.}{\sim} \mathcal{CN}(0, (1-\alpha^2)\sigma_h^2)$ is the driving noise, and α the correlation coefficient that may vary between zero to one according to the fading channel bandwidth f_m (Doppler spread).

We assume each packet consists of of N data symbols and P training symbols. Within one data packet, the received signals can be separated into data part $\mathbf{y}_d$ and training part $\mathbf{y}_t$, correspondingly. The system equations for these two parts are then given by

$$\mathbf{y}_t = \mathbf{S}_t \mathbf{h}_t + \mathbf{n}_t \qquad (3)$$
$$\mathbf{y}_d = \mathbf{S}_d \mathbf{h}_d + \mathbf{n}_d, \qquad (4)$$

where $\mathbf{h}_d$ and $\mathbf{h}_t$ are channel state vectors associated with data and training symbols, respectively.

We further assume that the data, the channel and noise are independent. Finally, we assume that the estimation is training based and is performed independently for each transmitted packet.

2.2. Pilot Symbol Placement

In general, the placement of n clusters of training symbols can be described by $\mathcal{P} = (\boldsymbol{\nu}, \boldsymbol{\gamma})$, where $\boldsymbol{\nu} = [\nu_1, \cdots, \nu_{n+1}]$ is the data block length vector and $\boldsymbol{\gamma} = [\gamma_1, \cdots, \gamma_n]$ the training cluster length vector, as illustrated in Figure 1. Constrained by the total number of data and training 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 ending with pilot symbols, $\nu_{n+1} = 0$.
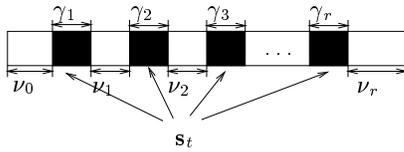
Fig. 1: An input sequence with multiple clusters

Fig. 2: Two types of data blocks in a packet.

2.3. MMSE Channel Estimator

We should only be concerned about the MSE of channel state (denoted by $\mathbf{h}_d$) associated with data symbols. Because the received data $\mathbf{y}_t$ and the channel $\mathbf{h}_d$ are jointly Gaussian, the MMSE estimator of $\mathbf{h}_d$, denoted as $\hat{\mathbf{h}}_d$, is linear. Let $\tilde{\mathbf{h}}_d \triangleq \hat{\mathbf{h}}_d - \mathbf{h}_d$ be the estimation error, the resulting minimum MSE is then given by

$$E\{\tilde{\mathbf{h}}_d \tilde{\mathbf{h}}_d^H\} = \mathbf{R}_{\mathbf{h}_d} - \mathbf{R}_{\mathbf{h}_{dt}} \mathbf{S}_t^H (\mathbf{S}_t \mathbf{R}_{\mathbf{h}_t} \mathbf{S}_t^H + \sigma_n^2 \mathbf{I})^{-1} \mathbf{R}_{\mathbf{h}_{dt}}^H. \quad (5)$$

where $\mathbf{R}_{\mathbf{h}_{dt}} = E\{\mathbf{h}_d \mathbf{h}_t^H\}$, $\mathbf{R}_{\mathbf{h}_t} = E\{\mathbf{h}_t \mathbf{h}_t^H\}$ and $\mathbf{R}_{\mathbf{h}_d} = E\{\mathbf{h}_d \mathbf{h}_d^H\}$. Note that the quantities $\mathbf{R}_{\mathbf{h}_{dt}}$, $\mathbf{R}_{\mathbf{h}_d}$ and $\mathbf{R}_{\mathbf{h}_d}$ are functions of placement $\mathcal{P}$.

Let $\mathcal{E}_i(\mathcal{P})$ denote the MSE for each channel state over data symbols, *i.e.,*

$$\mathcal{E}_i(\mathcal{P}) \triangleq E\{\tilde{\mathbf{h}}_d \tilde{\mathbf{h}}_d^H\}_{ii}, i = 1, \cdots, N. \quad (6)$$

When symbol-by-symbol detection is performed, it is $\mathcal{E}_i(\mathcal{P})$ that affects the symbol error probability. Therefore, it is reasonable to minimize the maximum MSE defined by

$$\mathcal{E}_{max}(\mathcal{P}) = \max_{1 \le i \le N} \mathcal{E}_i(\mathcal{P}). \quad (7)$$

Our objective is to find the optimal placement $\mathcal{P}_*$ that minimizes the maximum channel estimation error

$$\mathcal{P}_* = \arg \min_{\mathcal{P}} \mathcal{E}_{max}(\mathcal{P}). \quad (8)$$

3. OPTIMAL PLACEMENT FOR CHANNEL TRAINING

It can be seen from (5) that noise variance and the positions of the P training symbols affect the estimation performance of channel states for the training part, hence further affect the tracking performance at the data part. Thus, in general the resulting MSE is a complicated function of the placement $\mathcal{P}$, the channel correlation coefficient α and the SNR. Searching for the optimal placement turns out to be a difficult problem.

At high SNR, however, as channel learning becomes effective by inserting training symbols, only the tracking performance is influenced by training placement. In the following, we will consider the placement problem for the high SNR case, where perfect channel estimation is obtained during training, *i.e.,* $\hat{h}_{t_i} = h_{t_i}, i = 1, \cdots, P$.

Given an n-cluster training with placement $\mathcal{P}$, the data symbols are divided into $n+1$ blocks. It is easy to see that there are two types of data blocks in a packet: one type of blocks are those between two training symbol clusters. The second type are those at the two ends, next to only one training cluster. We first derive the maximum MSE and its position in these two types of data blocks, then find the maximum MSE in a packet, and finally we optimize the placement to minimize the maximum MSE.
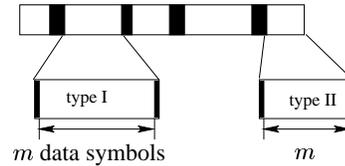
3.1. Data Blocks Between Training Symbols

Consider a data block of size m between two consecutive training symbols, as shown in Fig 2. We define it as a type-I block. The channel states associate with these m data symbols are denoted in a vector form $\mathbf{h}_d(m)$, and $\mathbf{h}_{t_i} = [h_{t_i}, h_{t_{i+1}}]^t$ is the channel state vector for the two training symbols at two ends of the block, Given $\mathbf{h}_{t_i}$, due to the Markov property of the channel model in (2), $\mathbf{h}_d(m)$ is independent of channel states associated with the rest of training symbols, *i.e.,* $h_{t_j}, j \ne i, i+1$. Therefore, $\hat{\mathbf{h}}_d(m)$ is only a function of $\mathbf{h}_{t_i}$. Denoting the received data according to training as $\mathbf{y}_{t_i} = [y_{t_i}, y_{t_{i+1}}]^t$, we have

$$\hat{\mathbf{h}}_d(m) = \mathbf{F} \mathbf{y}_{t_i} = \tilde{\mathbf{F}} \mathbf{h}_{t_i}, \quad (9)$$

where $\mathbf{F}$ is the optimal interpolator, *i.e., the Wiener Filter* coefficient matrix. The minimum MSE in (5), in this case, can be rewritten as

$$E\{\tilde{\mathbf{h}}_d(m) \tilde{\mathbf{h}}_d^H(m)\} = \mathbf{R}_{\mathbf{h}_d}(m) - \mathbf{R}_{\mathbf{h}_{dt}}(m) \mathbf{R}_{\mathbf{h}_t}^{-1}(m) \mathbf{R}_{\mathbf{h}_{dt}}^H(m), (10)$$

where

$$\mathbf{R}_{\mathbf{h}_d}(m) = \sigma_h^2 \begin{bmatrix} 1 & \alpha & \cdots & \alpha^{m-1} \\ \vdots & & \ddots & \vdots \\ \alpha^{m-1} & \alpha^{m-2} & \cdots & 1 \end{bmatrix}, \quad (11)$$

$$\mathbf{R}_{\mathbf{h}_{dt}}(m) = \sigma_h^2 \alpha \begin{bmatrix} 1 & \alpha^m \\ \vdots & \vdots \\ \alpha^m & 1 \end{bmatrix}, \quad (12)$$

$$\mathbf{R}_{\mathbf{h}_t}(m) = \sigma_h^2 \begin{bmatrix} 1 & \alpha^{m+1} \\ \alpha^{m+1} & 1 \end{bmatrix}. \quad (13)$$

It follows that the MSE of each channel state over the data block is

$$E\{\tilde{\mathbf{h}}_d(m) \tilde{\mathbf{h}}_d^H(m)\}_{ii} = (1 - \frac{\alpha^{2i} - 2\alpha^{2(m+1)} + \alpha^{2(m-i+1)}}{1 - \alpha^{2(m+1)}}) \sigma_h^2, \quad (14)$$

for $i = 1, \cdots, m$. Therefore, the position that gives the maximum MSE in the block is

$$\begin{aligned} i^* &= \arg \max_{1 \le i \le m} E\{\tilde{\mathbf{h}}_d(m) \tilde{\mathbf{h}}_d^H(m)\}_{ii} \\ &= \lceil \frac{m+1}{2} \rceil \, (\text{ or } \lfloor \frac{m+1}{2} \rfloor), \end{aligned} \quad (15)$$

and the maximum MSE is given by

$$\begin{aligned} \mathcal{E}_{max}^I(m) &= \max_{1 \le i \le m} E\{\tilde{\mathbf{h}}_d(m) \tilde{\mathbf{h}}_d^H(m)\}_{ii} \\ &= \begin{cases} \frac{1 - \alpha^{m+1}}{1 + \alpha^{m+1}} & m \text{ odd}; \\ \frac{2 - \alpha^m - \alpha^{m+2}}{1 - \alpha^{2(m+1)}} - 1 & m \text{ even} \end{cases} \end{aligned} \quad (16)$$

Intuitively, we expect that the interpolation performance is the worst for the middle position. The above confirms that the maximum error indeed appears in the middle of the data block, and is only a function of the data block size m for a fixed α. Thus, for any data block between training symbols, the maximum MSE can be calculated using (16).

3.2. Data Blocks at Ends of Packets

We now consider the second type of data block of size m which only next to one training cluster, shown in Fig 2. This type of data block appears only at two ends of a packet. We define it as a type-II block. The corresponding channel state vector only depends on the channel state associated to the training symbol next to the block. Following the similar derivation in the last section, the MSE for each channel state over the block of size m is given by

$$E\{\tilde{\mathbf{h}}_d(m)\tilde{\mathbf{h}}_d^H(m)\}_{ii} = (1-\alpha^{2i})\sigma_h^2, \quad i = 1, \cdots, m. \quad (17)$$

The maximum MSE and the position attaining it is then given by

$$\mathcal{E}_{max}^{II}(m) = \max_{1 \le i \le m} E\{\tilde{\mathbf{h}}_d(m)\tilde{\mathbf{h}}_d^H(m)\}_{ii} = 1 - \alpha^{2m}, \quad (18)$$

$$i^* = m. \quad (19)$$

For this type of data block, the maximum error is obtained at the furthest position from the training symbols, *i.e.,* the beginning and end of the packet. Again, it is only a function of data block size m for a given fading correlation coefficient α.

3.3. The Optimal Placement

3.3.1. Packet starting and ending with training

If we constrain transmitting packets starting and ending with training symbols, then only the type-I data blocks in Fig 2 are presented in each packet. It can be seen that the optimal placement minimizing the maximum MSE is obtained by minimize the size of the longest data block in a packet. The result is concluded in the following and shown in Fig 3

Proposition 1 *Assume each data packet starts and ends with training symbols, i.e., $\nu_1 = \nu_{n+1} = 0$. Under the assumed Rayleigh flat fading model, at high SNR, the optimal placement $\mathcal{P}_*$ is given by*

1. for $N = r(P-1), r \in N$:

$$\mathcal{P}_* = \{n = P; \gamma_i = 1; \nu_i = \frac{N}{P-1}, i = 2, \cdots, P\}; \quad (20)$$

2. for $N \ne r(P-1)$:

$$\mathcal{P}_* = \{n = P; \gamma_i = 1; \nu_i \in \{\lceil \frac{N}{P-1} \rceil, \lceil \frac{N}{P-1} \rceil - 1\}, 2 \le i \le P\} \quad (21)$$

The maximum MMSE under the optimal placement is given by

$$\mathcal{E}_{\max\min}^* = \min_{\mathcal{P}} \max_{1 \le i \le N} \mathcal{E}_i(\mathcal{P})$$

$$= \begin{cases} (\frac{1-\alpha^{\lceil \frac{N}{P-1} \rceil+1}}{1+\alpha^{\lceil \frac{N}{P-1} \rceil+1}})\sigma_h^2, & \lceil \frac{N}{P-1} \rceil \text{ odd;} \\ \\ (\frac{2-\alpha^{\lceil \frac{N}{P-1} \rceil}-\alpha^{\lceil \frac{N}{P-1} \rceil+2}}{1-\alpha^{2(\lceil \frac{N}{P-1} \rceil+1)}} - 1)\sigma_h^2, & \lceil \frac{N}{P-1} \rceil \text{ even.} \end{cases} \quad (22)$$

Fig. 3: The optimal placement under the constraint.

QPP placement is a family of placement strategy introduced in [8]. In a QPP-α scheme, training symbols are divided into as many clusters as possible provided that each of them is no less than α, and data blocks are divided as equal as possible. Proposition 1 shows that at high SNR, under the training symbol constraint, QPP-1 placement is optimal for channel tracking. It also shows that the optimal placement is invariant under channel fading characteristics α.

3.3.2. General case

In general, without constraint on training symbols, both type-I and type-II blocks are presented in a packet. $\mathcal{E}_{max}(\mathcal{P})$ is obtained by comparing the maximum MSEs of $n+1$ data blocks. Notice that both (16) and (18) are increasing functions. Therefore in each type, the block with the largest size gives the maximum MSE. We also have the following relation between a type-I block and a type-II block with size m_I and m_{II} respectively,

$$\mathcal{E}_{max}^I(m_I) \ge \mathcal{E}_{max}^{II}(m_{II}) \iff m_{II} \le g(m_I), \quad (23)$$

where

$$g(m) = \begin{cases} \frac{m+1}{2} + \frac{\log 2 - \log(1+\alpha^{m+1})}{2\log\alpha} & m \text{ odd;} \\ \\ \frac{m+1}{2} + \frac{\log((\alpha^{-1}+\alpha-2\alpha^{m+1})/(1-\alpha^{2(m+1)}))}{2\log\alpha} & m \text{ even.} \end{cases} \quad (24)$$

Optimizing the placement requires searching among all possible sizes of type-I and type-II blocks to minimize the maximum MSE. We seek the optimal placement such that the sizes of type-I blocks are as equal as possible, and between type-I and type-II blocks, the relation in (23) is as close to the equality as possible. This is described in the following and illustrated in Fig 4.

Theorem 1 *Under the assumed Rayleigh flat fading model, at high SNR, the following placement is optimal:*

$$\mathcal{P}_* = \left\{ n = P; \gamma_i = 1; \nu_1 = \lceil \frac{N - r^* - (m^*-1)(P-1)}{2} \rceil, \right.$$

$$\left. \nu_{P+1} = \lfloor \frac{N - r^* - (m^*-1)(P-1)}{2} \rfloor; \nu_i \in \{m^*, m^*-1\}, 2 \le i \le P. \right\}, \quad (25)$$

and m^*, r^* are given by

$$(m^*, r^*) = \arg\min\{\mathcal{E}_{max}^I(m_I^*), \mathcal{E}_{max}^{II}(\lceil \frac{N - r_{II}^* - (m_{II}^*-1)(P-1)}{2} \rceil)\} \quad (26)$$

where

$$(m_I^*, r_I^*) = \arg \min_{\substack{1 \le m \le \lceil \frac{N}{P-1} \rceil \\ 1 \le r \le P-1}} f(m, r), \quad \text{subject to } f(m, r) \ge 0;$$

$$(m_{II}^*, r_{II}^*) = \arg \max_{\substack{1 \le m \le \lceil \frac{N}{P-1} \rceil \\ 1 \le r \le P-1}} f(m, r), \quad \text{subject to } f(m, r) \le 0, \quad (27)$$

where

$$f(m, r) = g(m) - \lceil \frac{N - r - (m-1)(P-1)}{2} \rceil. \quad (28)$$

Fig. 4: The optimal placement in general cases.

Theorem 1 shows that in general, at high SNR, the optimal placement requires that each packet starts and ends with data symbols, between which training symbols comply with QPP-1 placement. The relation between the optimal sizes of type-I and type-II blocks is a function of the channel fading correlation coefficient α and the percentage of training. Notice that when $\alpha \to 1$, $g(m) \to \frac{m+1}{4}$. This shows that when channel fading is slow, under the optimal placement, the size of the blocks at two ends is about $\frac{1}{4}$ of that of blocks between training symbols.

4. NUMERICAL RESULTS

We compared the performance of tracking the fading channel under different placement strategies. The channel was Gaussian with variance $\sigma_h^2 = 1$. Fig 5 shows the maximum MSE vs. channel correlation coefficient α for different placement schemes at high SNR. Training percentage is 33%. Note that that when α is close to 0 and 1, all the placement schemes achieve almost the same performance. Intuitively, when $\alpha \to 0$, the channel changes independently, and inserting training provides little information for tracking no matter under what scheme. When $\alpha \to 1$, the channel becomes constant, and at high SNR training symbols provide almost perfect information about the channel. The efficiency of the optimal placement becomes apparent for α between 0 and 1. We see that there is a significant gain by placing training symbols optimally. Also, further performance improvement can be obtained by using the optimal placement in Theorem 1, comparing with the placement in Proposition 1. However, we also notice that when α is close to 1, the performance by the optimal placements in the two cases is very close. This indicates that QPP-1 scheme at $\alpha \to 1$ is an optimal placement scheme. Finally, we also plotted the performance of the optimal placement in [4]. In that paper, we showed that the single pilot periodical placement is optimal in an infinite data stream under the *Kalman Filter* channel tracking method. For finite packets, we see that the optimal placement under *Kalman Filter* tracking performs worse than that using the MMSE estimator. We also plotted the variation of the maximum MSE with α at $SNR = 30dB$, shown in Fig. 6, we see that a gain can still be obtained by placing training optimally. Fig 7 shows $\mathcal{E}_{max}(\mathcal{P})$ vs. percentage of training η for different placement schemes at high SNR. The channel Doppler bandwidth is fixed at $f_m = 100Hz(a = 0.97)$. We see that the efficiency of optimal placement shows its significance for tracking performance at low percentage of training. Again, at a close to 1, we see that QPP-1 scheme is in general optimal for different percentages of training.

5. CONCLUSION

In this paper, we considered the placement of training symbols for a time-varying Rayleigh flat fading channel. We tackled this problem by using a minmax approach. The placement is optimized by minimizing the maximum MSE in a packet. Our results have shown that training symbols should be placed periodically to maximize the tracking performances.

6. REFERENCES

[1] B.Hassibi and B.Hochwald. "How much Training is Needed in Multiple-Antenna Wireless Links". *Submitted to IEEE Trans. Information Theory*, August 2000.

[2] James K. Cavers. "An Analysis of Pilot Symbol Assisted Modulation for Rayleigh Fading Channels". *IEEE Trans. on Veh. Tech.*, 40:686–693, November 1991.

[3] M. Dong and L. Tong. "Optimal Design and Placement of Pilot Symbols for Channel Estimation". *Submitted to IEEE Trans. Signal Processing, April, 2001*.

[4] M. Dong and L. Tong. "Optimal Placement of Training for Channel Estimation and Tracking". In *MILCOM*, Vienna, Virginia, October 2001.

[5] 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.

[6] S. Ohno and G. B. Giannakis. "Average-Rate Optimal PSAM Transmissions over Time-Selective Fading Channels". *submitted to IEEE Trans. on Wireless Communications*, July 2001.

[7] S.Adireddy and L.Tong. "Optimal Placement of Known Symbols for Nonergodic Broadcast Channels". *Submitted to IEEE Trans. Info. Theory*, August 2001.

[8] S.Adireddy, L.Tong, and H.Viswanathan. "Optimal Placement of Known Symbols for Unknown Channels". *Submitted to IEEE Trans. Info. Theory*, March 2001.
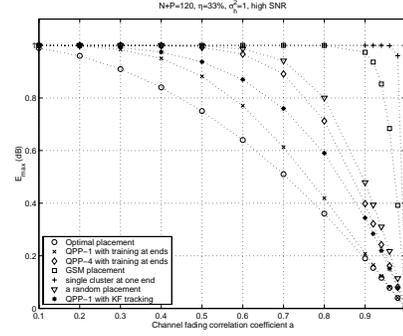
Fig. 5: $\mathcal{E}_{max}(\mathcal{P})$ vs. α at high SNR. $N + P = 120$, $\eta = 33\%, \sigma_h^2 = 1$.
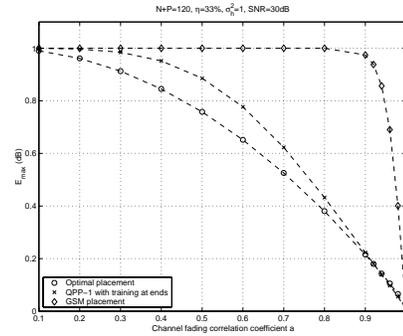

Fig. 6: $\mathcal{E}_{max}(\mathcal{P})$ vs. α. SNR=30dB. $N + P = 120$, $\eta = 33\%, \sigma_h^2 = 1$.
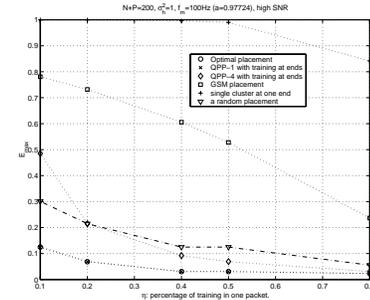

Fig. 7: $\mathcal{E}_{max}(\mathcal{P})$ vs. η at high SNR. $N + P = 200$, $f_m = 100Hz$, $\sigma_h^2 = 1$.