# TRAINING SYMBOL PLACEMENT FOR PACKET TRANSMISSIONS UNDER ASYNCHRONOUS INTERFERENCE

Cristian Budianu and Lang Tong†

School of Electrical and Computer Engineering,
Cornell University, Ithaca, NY, USA
{cris, ltong}@ece.cornell.edu

**ABSTRACT**

In a multiple access communication system that uses packet transmissions, the packets of one user might be subject to asynchronous interference from the packets of other users in the system. This paper analyses the influence of the placement of training symbols on the performance of the channel estimator in this scenario. The analysis of the mean square error (MSE) of the minimum mean square error estimator (MMSE) is shown to be equivalent to the analysis of the Fisher Information Matrix (FIM) for mixtures of Gaussian distributions. A complete solution to this problem is hard to find, but the bounds, asymptotics and simulations suggest that the best placement of training symbols is in two clusters of equal or quasi-equal size at the two ends of the data packet.

## 1. INTRODUCTION

The most challenging problems in the design of wireless communication systems are the time-varying wireless channel and the multiple access problem.

Communication through point-to-point time-varying channels under Gaussian noise has been studied extensively. A traditional method used by the receiver is to first obtain an estimate of the channel that is then used when decoding the data symbols. In order to cope with the fast variations of the channel the transmitter needs to insert a significant amount of side information into the data stream. Recent papers [1, 2] analyse the ultimate performance of some communication systems that use training to obtain the Channel State Information (CSI). Moreover, if a communication system relies on the assumption of perfect CSI, then the errors in the CSI can significantly degrade its performance [3]. On the other hand, if the CSI is available at the transmitter then the total throughput of the time-varying channel can be improved by allocating the transmitting power function of the CSI [4].

Similarly, in a multiple access scenario, the availability of CSI at the transmitter and receiver can be used by the medium access control (MAC) layer for power allocation among users and by the receiver for decoding [5]. For example, if the objective function is the sum capacity, then only the user with the best channel is allowed to transmit [6, 5]. Recently, in the framework of cross-layer design, the CSI is used by the MAC layer of a modified slotted ALOHA protocol to vary the transmission probability of each user [7]; the increase in the total throughput of the system is significant.

This paper considers the channel estimation problem in a multiple access scenario. The system under study has a MAC layer that allows collisions among packets. In addition, consider that the system uses frequency hopping, *i.e.*, each data packet of one user is transmitted using another carrier, and the hops are made according to a predetermined pattern. Frequency hopping makes the packets of any fixed pair of users collide very seldom. Thus, the interference that affects the packets of one fixed user comes from a different user each time. The packets of one fixed user are called data packets; the other colliding packets are called interference packets. Our approach considers a receiver that wants to recover as much information as possible from each of the data packets, instead of recovering (with a certain probability) a whole packet involved in a collision and dropping it if recovery is not possible. This approach is useful if the only synchronization among users is at the symbol level ( and not at the packet level). In this case the data and interference packets overlap partially, so that some symbols from the data packet are interference-free. Because the interference packets come from different users, the relative position of the data and interference packets is random from packet to packet. The results can be important for ad-hoc networks that are able to use the CSI (at either transmitter and receiver) to improve the total throughput.

Inserting training symbols in the data stream is one of the simplest but most widely used method of obtaining the CSI. In this paper the channel is assumed to be Rayleigh block flat fading and the receiver uses only training-based estimation. As a consequence, the channel parameter during the transmission of each packet is estimated using only the training symbols in the respective packet.

The effect of training symbol placement in the data stream have been studied previously. Negi and Cioffi [8], Adireddy, Tong and Viswananthan [1], Dong and Tong [9], Ohno and Giannakis [10], Budianu and Tong [11], considered this problem in different frameworks, and using different metrics. The goal of our paper is to investigate how the placement of training symbols within a packet influences the performance of channel estimation in the framework described above.

The paper is organized as follows: section 2 contains the model and is followed by section 3 with the description and analysis of the channel estimation. In 3.2 a genie lower bound on the MSE of the MMSE estimator is introduced and analyzed, and in 3.3 the behavior of the same MSE is analyzed using the FIM. In section 4 the simulations provide additional insight into the problem. We conclude the paper in section 5.

Notations : the vectors are in bold fonts, $\mathbb{E}_X$ is the expectation with respect to the random variable $X$, $\mathbb{P}\{A\}$ is the probability of the event $A$, $\nabla_{\mathbf{a}} f(\mathbf{a})$ is the gradient operator with respect to vector $\mathbf{a}$, $\mathrm{diag}(\mathbf{A})$ is a column vector formed by the diagonal elements of the square matrix $\mathbf{A}$. Sometimes the same function $f(\cdot)$ is written $f(\cdot; a)$ to emphasize the dependence on the parameter $a$. $a_k$ is the $k$-th element of vector $\mathbf{a}$ and for $f$ vectorial function $(f(x))_k$ is the $k$-th element of $f(x)$. If $A, B$ are square matrices, $A \geq B$ means that $A - B$ is positive semidefinite.

†This work was supported in part by the Multidisciplinary University Research Initiative (MURI) under the Office of Naval Research Contract N00014-00-1-0564, and Army Research Laboratory CTA on Communication and Networks under Grant DAAD19-01-2-0011.

## 2. SYSTEM MODEL AND PROBLEM FORMULATION

Consider a point-to-point one-way communication link. The channel is assumed Rayleigh block-flat-fading, constant during the transmission of one packet, and independent from packet to packet. The symbols of the data packets can be either training or data. The receiver obtains an estimate of the channel based only on the training symbols from the current packet received.

The communication is subject to the usual i.i.d. complex additive white Gaussian noise (CAWGN) with known variance $\sigma^2$. Besides this, the data packets are affected by interference which is modeled as a packet of i.i.d CAWGN with known variance $\sigma_q^2$ that affects a subset of the symbols of the packet depending on its position. The following assumptions about the interference are made: only one interference packet can hit the data packet at a time, the data and interference packets have the same length, the interference packet always hits the data packet, and the relative position of the data and interference packets is distributed uniformly. An important feature of this model is that the relative position mentioned above is not known by the transmitter or the receiver.

Denote by $B$ the total number of symbols in one packet, out of which $N$ are allocated to training. The symbols in one received packet are given by the $B \times 1$ vector

$$\mathbf{y} = \mathbf{s}A + \mathbf{z}. \tag{1}$$

In the equation above $A \sim \mathcal{CN}(0,1)$ is the complex scalar channel parameter, $\mathbf{s}$ a $B \times 1$ vector representing one block of transmitted symbols and $\mathbf{z}$ the total noise vector that includes the CAWGN and the interference; its probability density function (pdf) is a mixture of Gaussians. We assumed the variance of the channel equal to 1 without loss of generality.

A channel estimate for each packet is obtained only from the received training symbols. Denote by $\mathcal{J} = \{t_1, \dots, t_N\} \subset \{1, \dots, B\}$, $1 \le t_1 < t_2 < \dots < t_N \le B$, the ordered set of indexes of the training symbols within the packet. Extracting these symbols into an $N \times 1$ vector $\mathbf{x}$ we have

$$\mathbf{x} = \mathbf{P}\mathbf{y} = \mathbf{1}A + \mathbf{P}\mathbf{z} = \mathbf{1}A + \mathbf{n}. \tag{2}$$

The placement matrix $\mathbf{P}$ is an $N \times B$ matrix, with elements $\mathbf{P}(k, t_k) = 1$, $\forall k \in \{1, \dots, N\}$ and the rest of the elements 0. In other words, for each $k$, the $k$-th row of $\mathbf{P}$ has only one non-zero element in the column $t_k$ that gives the position of the $k$-th training symbol within the data packet. We also supposed that all the training symbols are equal to 1.

To do the channel estimation we need the distribution of the total noise and interference vector $\mathbf{n}$ that affects the training symbols.

From the assumption that the interference packets are of the same length $B$ as the data packets, it follows that the relative position between the two packets can be described by a discrete random variable $V$ distributed uniformly on $\{1, \dots, 2B-1\}$. If $V \in \{1, \dots, B\}$ then the first $V$ symbols of the data packet are hit, if $V \in \{B+1, \dots, 2B-1\}$ then the last $2B - V$ symbols are hit, see Fig.1. Then, we introduce another random variable, $U \in \{1, \dots, 2N\}$ that gives the position of the interference packet with respect to the *training symbols*. Similar to $V$, if $U \in \{1, \dots, N\}$ the first $N$ training symbols are hit, if $U \in \{N+1, \dots, 2N-1\}$ the last $2N - U$ training symbols are hit, and $U = 2N$ corresponds to the case in which no training symbol is hit by the interference. $U$ is not distributed uniformly anymore; given $V$, the value of $U$ follows but there can be more than one value of $V$ leading to the same $U$, see Fig. 1. The distribution of $U$ can be obtained from the placement of the training symbols by taking into account the uniform distribution of $V$. Defining $\varepsilon \triangleq \frac{1}{2B-1}$ and the sets $\mathcal{S}_{N-1} \triangleq \{1, \dots, N-1\}$, $\mathcal{S}_{N+1} \triangleq \{N+1, \dots, 2N-1\}$, the distribution $p_u \triangleq \mathbb{P}\{U = u\}$ of $U$ is given by

$$p_u = \begin{cases} \varepsilon(t_{u+1} - t_u) & \text{if} \quad u \in \mathcal{S}_{N-1} \\ \varepsilon(B - (t_N - t_1)) & \text{if} \quad u = N \\ p_{u-N} & \text{if} \quad u \in \mathcal{S}_{N+1} \\ p_N - \varepsilon & \text{if} \quad u = 2N \end{cases}. \tag{3}$$

We'll use $\mathbf{p} \triangleq [p_1, \dots, p_{2N}]^T$ to refer to the distribution of U. Also, denote by $\mathcal{P}$ the set of all distributions $\mathbf{p}$ that satisfy the conditions (3). For channel estimation purposes, all placements with the same $\mathbf{p}$ have the same behavior due to the uniform distribution of $V$, so we'll refer to placements through their corresponding $\mathbf{p}$.

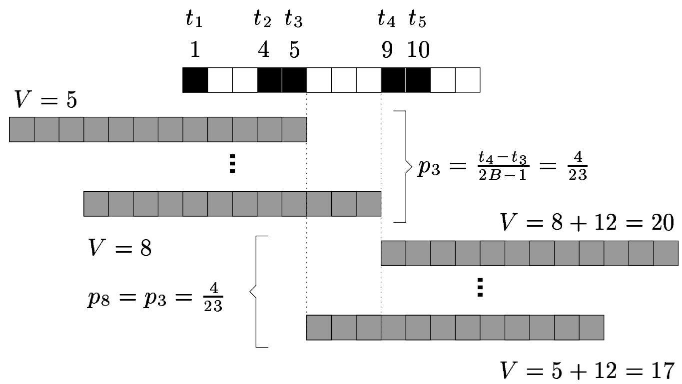



**Fig. 1**: The symmetry property of the distribution of U. The training symbols are colored in black. The interference packets are colored in gray. $B = 12$, $N = 5$, $\mathcal{J} = \{1, 4, 5, 9, 10\}$

The pdf of the total noise $\mathbf{n}$ is given by $p(\mathbf{n}) = \sum_{u=1}^{2N} p_u p(\mathbf{n}|U = u)$, where $p(\mathbf{n}|U = u) = \frac{1}{\pi^N |\mathbf{D}_u|} \exp\left(-\mathbf{n}^H \mathbf{D}_u^{-1} \mathbf{n}\right)$. For any $u \in \{1, \dots, 2N\}$ the elements of the diagonal matrices $\mathbf{D}_u$ take only 2 values, $\sigma^2$ and $\sigma_h^2 \triangleq \sigma^2 + \sigma_q^2$ ( which is the variance of noise and interference corresponding to the symbols that are hit )

$$\mathrm{diag}(\mathbf{D}_u) = \begin{cases} [\sigma_h^2 \mathbf{1}_u^T \;\; \sigma^2 \mathbf{1}_{N-u}^T]^T & \text{if} \quad u \in \mathcal{S}_{N-1} \\ \sigma_h^2 \mathbf{1}_N & \text{if} \quad u = N \\ [\sigma^2 \mathbf{1}_{u-N}^T \;\; \sigma_h^2 \mathbf{1}_{2N-u}^T]^T & \text{if} \quad u \in \mathcal{S}_{N+1} \\ \sigma^2 \mathbf{1}_N & \text{if} \quad u = 2N \end{cases}. \tag{4}$$

Further, the pdf of the received signal is

$$p(\mathbf{x}) = \sum_{u=1}^{2N} p_u p(\mathbf{x}|U = u) \tag{5}$$

$$p(\mathbf{x}|U = u) = \frac{1}{\pi^N |\mathbf{C}_u|} \exp\left(-\mathbf{x}^H \mathbf{C}_u^{-1} \mathbf{x}\right), \tag{6}$$

where $\mathbf{C}_u \triangleq \mathbf{1}\mathbf{1}^T + \mathbf{D}_u$.

## 3. CHANNEL ESTIMATION

### 3.1. The Bayesian Estimator

The Bayesian MMSE estimator for the model considered is

$$\begin{aligned} \hat{A}(\mathbf{x}; \mathbf{p}) &= \mathbb{E}\{A|\mathbf{x}\} = \mathbb{E}_U \, \mathbb{E}_A \{A|\mathbf{x}, U\} \\ &= \mathbb{E}_U \{\mathbf{1}^T \mathbf{C}_U^{-1} \mathbf{x} | \mathbf{x}, U\} \\ &= \mathbf{1}^T \mathbb{E}_{U|\mathbf{x}} \{\mathbf{C}_U^{-1}\} \mathbf{x}. \end{aligned} \tag{7}$$

Conditioned on $U$, the model becomes the well-known Gaussian model, and relation (7) follows easily. Writing the expectation explicitly we have

$$\hat{A}(\mathbf{x};\mathbf{p}) = \mathbf{1}^T \left( \sum_{u=1}^{2N} \frac{p_u p(\mathbf{x}|U=u)}{p(\mathbf{x})} \mathbf{C}_u^{-1} \right) \mathbf{x}. \tag{8}$$

The performance of the Bayesian MMSE estimator is given by the MSE

$$MSE(\mathbf{p}) = \mathbb{E}[|A|^2] - \mathbb{E}[|\hat{A}|^2]. \tag{9}$$

The MMSE estimator and its MSE are function of the distribution of the random variable $U$, *i.e.,* the vector $\mathbf{p}$. Our goals are to characterize the dependence of the performance of the MMSE estimator on the placement, and to find the placement(s) $\mathbf{p}_0$ that minimizes the MSE (9) under the conditions (3) imposed by the physical model

$$\arg\min_{\mathbf{p}\in\mathcal{P}} \{MSE(\mathbf{p})\}. \tag{10}$$

Given a placement defined by the set $\mathcal{J}$ and having the distribution $\mathbf{p}$, define its mirror reflection by $\mathcal{J}^{\leftarrow} \triangleq \{N+1-t_N, \ldots, N+1-t_1\}$ and the corresponding $\mathbf{p}^{\leftarrow}$. Note that we can have $\mathbf{p} = \mathbf{p}^{\leftarrow}$. Because of the left-right symmetry of the model, the mirror reflection $\mathbf{p}^{\leftarrow}$ has the same performance as $\mathbf{p}$. Thus if $\mathbf{p}_0$ is a solution of (10), so is $\mathbf{p}_0^{\leftarrow}$.

In our case the MMSE estimator and its performance are nonlinear functions of $\mathbf{p}$. This makes their analysis a hard problem.

### 3.2. The Genie Lower Bound on the MSE

A lower bound on the MSE can be obtained by considering the performance of a receiver helped by a genie who provides the current value of $U$, *i.e.,* the position of the interference packet with respect to the training symbols. In this case, for each value $U = u$ we have a Gaussian model, for which the MMSE and its MSE are well known. Consider the following estimator, that assumes the random variable $U$ known:

$$\tilde{A}(\mathbf{x}, U) \triangleq \mathbb{E}\{A|\mathbf{x}, U\} = \mathbf{1}^T \mathbf{C}_U^{-1} \mathbf{x}.$$

Its MSE for each $U = u$ is given by $m(u)$ and the averaged MSE by $\xi(\mathbf{p})$

$$\begin{aligned} m(u) &\triangleq \mathbb{E}\{|\tilde{A}(\mathbf{x}, U) - A|^2 | U = u\} = 1 - \mathbf{1}^T \mathbf{C}_u^{-1} \mathbf{1} \\ \xi(\mathbf{p}) &\triangleq \mathbb{E}\{|\tilde{A}(\mathbf{x}, U) - A|^2\} \\ &= \sum_{u=1}^{2N} p_u \mathbb{E}\{|\tilde{A}(\mathbf{x}, U) - A|^2 | U = u\} = \sum_{u=1}^{2N} p_u m(u). \end{aligned} \tag{11}$$

Since $\tilde{A}$ is the MMSE estimator given the state $U$, for any other estimator $\hat{A}_0(\mathbf{x}, U)$ we have

$$m(u) \leq \mathbb{E}\{|\hat{A}_0(\mathbf{x}, U) - A|^2 | U = u\}.$$

The relation above is true for $\hat{A}_0 = \hat{A}$ as well, for each $u$, so we have a lower bound for the MSE of the MMSE estimator $\hat{A}$

$$\xi(\mathbf{p}) \leq \mathbb{E}\{|\hat{A}(\mathbf{x};\mathbf{p}) - A|^2\}.$$

The genie lower bound can be optimized with respect to $\mathbf{p}$. The result is given below.

**Theorem 1** *Let $\bar{\mathbf{p}}$ be the probability distribution given by*

$$\bar{p}_u = \begin{cases} \varepsilon & \text{if} \quad u \notin \{\lfloor \frac{N}{2} \rfloor, N + \lfloor \frac{N}{2} \rfloor, 2N\} \\ (B - N + 1)\varepsilon & \text{if} \quad u \in \{\lfloor \frac{N}{2} \rfloor, N + \lfloor \frac{N}{2} \rfloor\} \\ 0 & \text{if} \quad u = 2N \end{cases}. \tag{12}$$

*Then $\bar{\mathbf{p}}$ and $\bar{\mathbf{p}}^{\leftarrow}$ are the only placements that minimize the genie lower bound $\xi(\mathbf{p})$ subject to the conditions (3):*

$$\{\bar{\mathbf{p}}, \bar{\mathbf{p}}^{\leftarrow}\} = \arg\min_{\mathbf{p}\in\mathcal{P}} \{\xi(\mathbf{p})\}. \tag{13}$$

*Note that $\bar{\mathbf{p}} = \bar{\mathbf{p}}^{\leftarrow}$ if $N$ is even.*

*Proof*: Since the MSE of the genie estimator is the same if the interference hits the first $u$ symbols or the last $u$, for $u = 1, \ldots, 2N-1$ we have $m(u) = m(2N-u)$. Then, taking into account the definitions of $\mathbf{C}_u$ and $\mathbf{D}_u$, for $u = 1, \ldots, N$ we have

$$m(u) = \frac{1}{1 + \sum_{k=1}^{N} \frac{1}{D_u(k,k)}} = \frac{1}{1 + \frac{u}{\sigma^2 + \sigma_q^2} + \frac{N-u}{\sigma^2}}$$

Replacing $u$ with a continuous variable $x$, for $x \in [0, N]$, $m(x)$ is a convex function of $x$. Thus the function $m(x) + m(N-x)$ is strictly decreasing on $\left[0, \frac{N}{2}\right]$. It follows that under the conditions (3), $\bar{\mathbf{p}}$ given by (12) minimize the function $\xi(\mathbf{p}) = \sum_{u=1}^{N-1} p_u(m(u) + m(N-u)) + p_N(m(2N) + m(N)) - \frac{1}{2B-1} m(2N)$.

The distributions given in the previous theorem correspond to the placement of training symbols in two clusters of equal (or nearly equal) length at the edges of the packet. This can be easily seen from the definition (3) of $p_u$ ( $t_1 = 1$ and $t_N = B$ follow from $p_{2N}$, then the other $t_k$).

Next, one can observe that for any placement $\mathbf{p}$ the genie bound is tight when the power of the interference is high. This is stated in the next theorem.

**Theorem 2** *As denoted before, let $\xi(\mathbf{p};\sigma_q)$ and $MSE(\mathbf{p};\sigma_q)$ be the genie bound and the MSE of the MMSE estimator respectively, where the dependence on the power of the interference has been shown explicitly. Then we have*

$$\lim_{\sigma_q \to \infty} (MSE(\mathbf{p};\sigma_q) - \xi(\mathbf{p};\sigma_q)) = 0. \tag{14}$$

*Proof*: The proof is rather long and it will be available in [12]. The result can be briefly explained by observing that if the power of the interference is high, then $U$, the position of the interference, can be detected accurately. We consider a suboptimal estimator that uses hard detection of $U$, then linear estimation based on the detected value. For high values of $\sigma_q^2$ the probability of detection error is close to zero and the corresponding estimation error is bounded. This shows that the performance of the suboptimal estimator is close to the genie bound which proves the theorem.

From theorem 2 it follows that increasing $\sigma_q^2$ we can make the MSE of the MMSE estimator be as close to the genie bound as wanted. Observing that the function $m(x)$ preserves its strict convexity when $\sigma_q^2 \longrightarrow \infty$, we see that even in this case the placements $\bar{\mathbf{p}}$ given by (12) and $\bar{\mathbf{p}}^{\leftarrow}$ are the only solutions of the optimization problem stated in Theorem 1. We have the following corollary.

**Corollary 1** *The placements $\bar{\mathbf{p}}$ given by (12) and $\bar{\mathbf{p}}^{\leftarrow}$ are the solutions of the general problem (10) for $\sigma_q^2$ high enough.*

It can be observed that the placement that maximizes the genie bound (*i.e.,* the worst ) can be found as well, and it corresponds to placing all the training symbols into one cluster. The position of the training cluster within the data packet does not modify the channel estimation performance (because it does not modify $\mathbf{p}$). Using theorem 2 we can conclude that placing the training symbols in one cluster provides the worst performance at high values of the interference power.

### 3.3. A partial characterization of the MSE

In this section the connection between the MSE (9) and the FIM is established and used to show that the optimal placement $\mathbf{p}_0$ (solution of (10)) belongs to a certain subset with $\lfloor \frac{N}{2} \rfloor + 1$ elements.

For a random complex vector $\mathbf{X}$ with pdf $f$, the Fisher Information Matrix (FIM) is defined as

$$\mathbf{J}(f) = \mathbb{E}\left\{ \nabla_{\mathbf{x}^*} \{\log f(\mathbf{x})\} (\nabla_{\mathbf{x}^*} \{\log f(\mathbf{x})\})^H \right\}. \qquad (15)$$

Some regularity conditions on $f$ are necessary for the FIM to exist, see [13] for details. These conditions are satisfied by the distributions considered in our problem.

**Lemma 1** *The MSE (9) of the MMSE estimator can be written as*

$$MSE(\mathbf{p}) = 1 - \mathbf{1}^T \mathbf{J}(p(\cdot; \mathbf{p}))\mathbf{1}. \qquad (16)$$

*The pdf $p(\cdot; \mathbf{p})$ is given by (5); here we indicated the dependence on $\mathbf{p}$ explicitly.*

*Proof* : From the properties of Gaussian pdf's we have

$$\nabla_{\mathbf{x}^*} p(\mathbf{x}|U = u) \quad = \quad (-1) p(\mathbf{x}|U = u) \mathbf{C}_u^{-1} \mathbf{x}. \qquad (17)$$

Using this to express $\hat{A}$ given by (8), and then substituing in (9) the lemma follows.

This lemma allows us to use some properties of the FIM of random vectors in our problem. The next lemma is a convexity property of FIM, that is an extension to complex vectors of a weaker form found in [14],[15].

**Lemma 2** *The FIM of a random complex vector $\mathbf{X}$ with pdf $f$ is a convex functional of $f$. The convexity holds in the sense of positive definiteness.*

Proof: The proof follows the one given in [14] for real random variables.

More detailed, let $\mathbf{X}_1$ and $\mathbf{X}_2$ be random vectors with densities $f_1$ and $f_2$ respectively, and $a \in [0, 1]$ an arbitrary number. Then, the following inequality holds :

$$a\mathbf{J}(f_1) + (1 - a)\mathbf{J}(f_2) \geq \mathbf{J}(af_1 + (1 - a)f_2). \qquad (18)$$

The convexity property given above allows us to reduce the number of possible solutions of the optimization problem (10).

Using the symmetry properties in (3), we can rewrite the pdf in (5)

$$\begin{aligned} p(\mathbf{x}; \mathbf{p}) \quad = \quad & \sum_{u=1}^{N-1} 2p_u \frac{1}{2} \left( p(\mathbf{x}|U = u) + p(\mathbf{x}|U = N + u) \right) \\ & + 2p_{2N} \left( \frac{1}{2} p(\mathbf{x}|U = N) + \frac{1}{2} p(\mathbf{x}|U = 2N) \right) \\ & + \varepsilon p(\mathbf{x}|U = N). \qquad (19) \end{aligned}$$

Using the convexity property of the FIM functional the following theorem is easy to prove.

**Theorem 3** *Define the set $\mathcal{V}$ that contains all those vectors $\mathbf{p} \in \mathcal{P}$ for which one of the elements has its maximum possible value; once this element is chosen, the rest of the elements have the values given by (3).*

$$\begin{aligned} \mathcal{V} \quad \triangleq \quad & \{\mathbf{p} \in \mathcal{P} | \, \exists m \in \{1, \ldots, N\} : \\ & p_m = (B - N + 1)\varepsilon; \\ & p_u = \varepsilon \, , \; \forall u = 1, \ldots, N, u \neq m\}. \qquad (20) \end{aligned}$$

*Any solution $\mathbf{p}_0$ of (10) satisfies $\mathbf{p}_0 \in \mathcal{V}$.*

From (3) it follows that all the placements in $\mathcal{V}$ have all the training symbols placed in exactly two clusters at the two ends of the packet, or all the training symbols placed in one cluster. Although the set $\mathcal{V}$ has $N$ elements, since the mirror pairs have the same MSE, the optimal solution should be searched among only $\lfloor \frac{N}{2} \rfloor + 1$ elements.

Notice that the solution (12) obtained using the genie bound corresponds to the choice $p_{\lfloor \frac{N}{2} \rfloor} = (B - N + 1)\varepsilon$. Unfortunately, we were not able to show that the choice given above is the solution of the problem for any choice of the parameters. Besides the genie bound solution, the simulations suggest the following conjecture.

**Conjecture 1** *The problems (10) and (13) have the same solutions that are $\bar{\mathbf{p}}$ given by (12) and $\bar{\mathbf{p}}^{\leftarrow}$.*

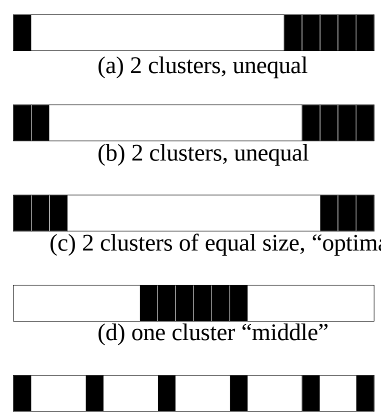
As noticed before, the distribution (12) implies the placement of the training symbols in two clusters of equal or quasi-equal sizes at the two ends of the packet.

From the approach above, it is clear that the genie bound discussed in the previous section is the bound given by the convexity relation, where the pdf of the received signal is splitted into its Gaussian components.

## 4. SIMULATION RESULTS

As it was mentioned previously, a complete solution of the problem is hard to find. The simulations available in this section are aimed to give some evidence in favor of the conjecture.

The simulations were done for the following parameters : $B = 80, N = 6, \sigma^2 = 1/49 = -16.9dB, \sigma_A^2 = 1$. Taking into account theorem 3 and the remark that follows, the optimal scheme should be searchd in a set with 4 elements. The corresponding placements are the first 4 placements in Fig.2. Fig.3 shows that we can gain


(a) 2 clusters, unequal

(b) 2 clusters, unequal

(c) 2 clusters of equal size, "optimal"

(d) one cluster "middle"

(e) training symbols "spread"

**Fig. 2**: The training schemes compared in Figs. 3 and 4

more than 10dB by using the 'optimal' placement over the placement that uses one cluster (in the middle of the packet). This is the maximum gain that can be obtained; according to the genie bound the placement of all the training symbols in the middleoffers the worst performance. Once we have two clusters placed at the edges of the packet, the performance gain by using the optimal placement is smaller, up to 2dB. In fig.4 we compared the actual MSE with the genie bound for three of the placements represented in fig.2 *i.e.*, "optimal", "middle" and "spread" schemes. For the "optimal" and the "middle" placements, the genie bounds are relatively tight. The interesting fact is that the MSE of the "spread" placement scheme has a bell shape and the genie bound in not tight. This can be explained by thinking that the coefficients $p(U = u|\mathbf{x})$ in the expression (8) of the MMSE estimator $\hat{A}$ act like an embedded soft

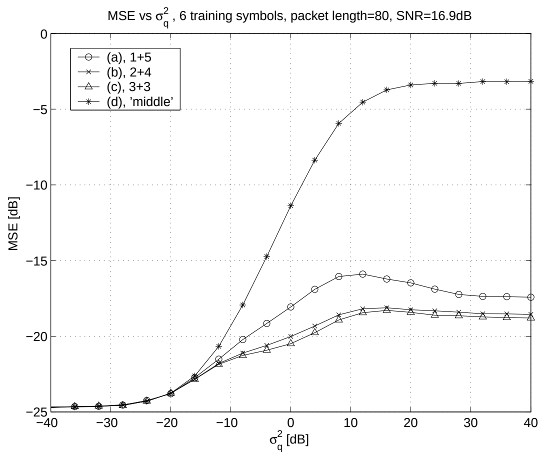



**Fig. 3**: The performance of the first 4 training schemes from Fig. 2. The legend shows hows the number of training symbols in each cluster.

detector. The detection can be done better if there are fewer events with high *a priori* probabilities. This happens if the symbols are grouped into two clusters placed at the edges or in one cluster; in these cases the MSE is close to the genie bound.

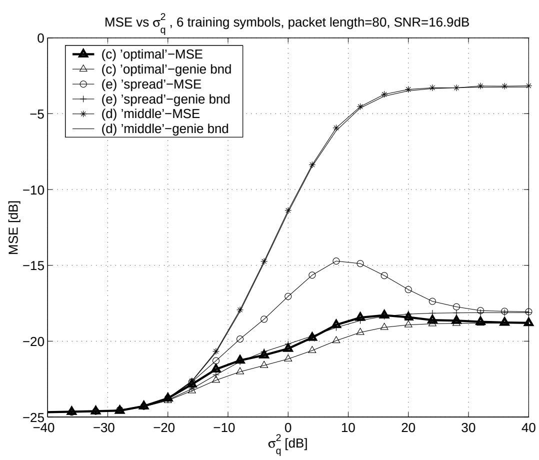



**Fig. 4**: The performance of the “optimal”, “middle” and “spreaded” placements and their genie bounds

## 5. CONCLUSIONS

In this paper we considered the channel estimation in the presence of an asynchronous interference. The model is equivalent to an estimation problem where the noise distribution is a mixture of Gaussians. It was shown that the MSE of the nonlinear MMSE estimator is related to the FIM of the received signal. Further, the convexity property of the FIM allows a characterization of the performance of different placements of the training symbols. It was shown that the optimal placement should be searched in a small set of placement schemes. To find the optimal placement scheme one needs to solve a complicated optimization problem; however the bounds and the simulations suggest that placing the training symbols in two clusters of equal or nearly equal sizes optimizes the MSE of the Bayesian MMSE estimator.

## 6. REFERENCES


[1] S. Adireddy, L. Tong, and H. Viswanathan, “Optimal placement of known symbols for frequency-selective block-fading channels,” *IEEE Trans. Info. Theory*, vol. 48, pp. 2338–2353, August 2002.

[2] B. Hassibi and B. Hochwald, “High-rate codes that are linear in space and time,” *IEEE Transactions on Information Theory*, vol. 47, pp. 1804–1824, July 2002.

[3] A. Lapidoth and S. Shamai, “Fading channels: how perfect need ‘perfect side information’ be,” *IEEE Trans. Information Theory*, vol. 48, pp. 1118–1134, May 2002.

[4] A. Goldsmith and P. Varaiya, “Capacity of fading channels with channel side information,” *IEEE Trans. Info. Theory*, vol. 43, pp. 1986 –1992, Nov 1997.

[5] D.N.C.Tse and S.V.Hanly, “Multiaccess Fading Channels : I Polymatroid Structure, Optimal Resource Allocation and Throughput Capacities ,” *IEEE Trans. Information Theory*, vol. 44, pp. 2796–2815, November 1998.

[6] R.Knopp and P.A.Humblet, “ Information capacity and power control in single cell multi-user communications ,” in *Proc. Intl Conf. Comm.*, (Seattle, WA), pp. 331–335, June 1995.

[7] S. Adireddy and L. Tong, “Exploiting decentralized channel state information for random access,” *Submitted to IEEE Trans. Info. Theory*, November 2002. `http://acsp.ece.cornell.edu/pubJ.html`.

[8] R. Negi and J. Cioffi, “Pilot tone selection for channel estimation in a mobile OFDM System ,” *IEEE Trans. on Consumer Electronics*, vol. 44, pp. 1122–1128, Aug. 1998.

[9] M. Dong and L. Tong, “Optimal Design and Placement of Pilot Symbols for Channel Estimation,” *IEEE Trans. on Signal Processing*, vol. 50, pp. 3055–3069, December 2002.

[10] S. Ohno and G. B. Giannakis, “Optimal Training and Redundant Precoding for Block Transmissions with Application to Wireless OFDM,” *IEEE Trans. on Communications*, vol. 50, December 2002.

[11] C. Budianu and L. Tong, “Channel Estimation for Space-Time Block Coding Systems,” *IEEE Trans. Signal Processing*, vol. 50, pp. 2515–2528, Oct 2002.

[12] C. Budianu and L. Tong, “Channel Estimation under Asynchronous Packet Interference,” Tech. Rep. ACSP-TR-03-03, Cornell University, March 2003. `http://acsp.ece.cornell.edu/pubJ.html`.

[13] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ 07632: Prentice Hall, 1993.

[14] B. Frieden, A. Plastino, A. Plastino, and B. Soffer, “Fisher-based thermodynamics: Its Legendre transform and concavity properties,” *Physical Review E*, vol. 60, pp. 48–53, July 1999.

[15] A. Dembo, T. Cover, and J. Thomas, “Information Theoretic Inequalities,” *IEEE Transactions on Information Theory*, vol. 37, pp. 1501–1518, June 1991.