PROCEEDINGS OF THE CONFERENCE ON MULTIACCESS, MOBILITY AND TELETRAFFIC FOR WIRELESS COMMUNICATIONS

Edited by
JAIGANESH BALAKRISHNAN [1]
MARKUS RUPP [2]
HARISH VISWANATHAN [2]

[1] School of Electrical and Computer Engineering,
Cornell University, Ithaca, NY 14850, USA
Fax: (607) 255 9072, Tel: (607) 254 8819
e-mail:jai@ee.cornell.edu

[2] Bell-Labs, Lucent Technologies, Wireless Research Lab.
791 Holmdel-Keyport Rd., Holmdel, NJ 07733-0400, USA
Fax: (732) 888 7074, Tel: (732) 888 7104
e-mail:{rupp,harishv}@lucent.com

Kluwer Academic Publishers
Boston/Dordrecht/London

Contents

List of Tables

Contributing Authors

Jaiganesh Balakrishnan is a graduate student in the School of Electrical and Computer Engineering at Cornell University. He earned his MS from Cornell University in August 1999 and is currently pursuing his Ph.D.

Markus Rupp is a Member of Technical Staff in the Wireless Research Laboratory at Bell-labs, Lucent Technologies, where he is responsible for rapid prototyping methods in wireless systems. He earned his Ph.D at the Technische Hochschule Darmstadt, Germany.

Harish Viswanathan is a Member of Technical Staff in the Wireless Research Laboratory at Bell-labs, Lucent Technologies, where he is currently involved in research and development of the use of multiple antennas in third generation wireless systems.

OPTIMAL CHANNEL TRAINING FOR MULTIPLE ANTENNA SYSTEMS

Jaiganesh Balakrishnan
School of Electrical and Computer Engineering
Cornell University
jai@ee.cornell.edu

Markus Rupp
Bell-Labs, Lucent Technologies
Wireless Research Lab
rupp@lucent.com

Harish Viswanathan
Bell-Labs, Lucent Technologies
Wireless Research Lab
harishv@lucent.com

Keywords: Training Sequence Design, Optimal Training, Channel Estimation, Multiple Antennas, Delay Diversity.

Abstract The design of optimal training sequences for channel estimation in multiple antenna systems is considered. The optimality criterion for training sequence design is derived and the design tradeoffs associated with the choice of training length is discussed. A few heuristic methods for the search of near-optimal training sequences are proposed. Optimal and near-optimal (binary) training sequences for multiple antenna systems are listed. The design of training sequences for a delay-diversity scheme is discussed. It is shown that the choice of identical training sequences, transmitted with the appropriate delays, is optimal for the delay-diversity scheme.

1. INTRODUCTION

Recent antenna technology advances have made it possible to support multiple transmit and receive antennas in the terminal[1, 2]. Particularly for large size data terminals such as laptops, it is possible to have up to four integrated antennas with sufficient spacing so that the correlation of the transmitted/received signals across the antennas is small. Phased array antennas and widely-spaced diversity antennas are two ways to use multiple antenna to provide improved spectral efficiency. In the first case a narrow beam directed towards the terminal is formed by transmitting the same signal, appropriately weighted in amplitude and phase, from each antenna element, while in the later case different signals are transmitted from the different antennas in order to take advantage of scattering through space-time coding. Space-Time Coding (STC) can be used in different ways: some use the additional antenna elements to provide diversity gain (e.g. [3]), while other techniques, such as BLAST (Bell Labs Layered Space-Time) [4], are practical methods for achieving higher data rates through the use of multi-element antenna arrays.

The multi-path effects of the wireless channel can be characterized as a linear filtering of the original signal causing inter-symbol interference at the receiver. In general, the design of an optimal detector at the receiver requires knowledge of each of the transmitted channels. Training-based estimation, semi-blind estimation and blind estimation are three types of estimators that can be potentially used to estimate the Multi-Input-Multi-Output (MIMO) channel impulse response at the receiver.

Among the first new wireless standards that explores the possibility of exploiting multiple antennas is the Enhanced Data rates through GSM Evolution (EDGE) system. An EDGE burst contains a training sequence and hence the training-based and semi-blind channel estimators are more suitable. Although, semi-blind channel estimators may offer a better performance over purely training-based channel estimators [5], this paper considers the use of a training-based channel estimator. In a training-based channel estimator, the channel impulse response is estimated during the training phase and the quality of the estimate depends on the particular choice of the training sequence.

The MIMO system model and a Least-Squares (LS) channel estimator are discussed in Section 2. LS estimation for channel sounding has been proposed previously [6] but, typically in the context of single antenna systems. Optimal and near-optimal binary sequences, for various channel lengths, are listed in [6] and [7] for single antenna systems. An optimality criterion for training sequence design for MIMO channel estimation is derived in Section 3. Furthermore, design tradeoffs in the choice of training sequence length are discussed. Section 4 describes a few heuristic methods for the search of near-optimal training sequences. Many optimal sequences for specific alphabets are listed. A few near-optimal binary sequences obtained using these search methods are tabulated in Appendix B. The design of training sequences for a delay-diversity scheme is discussed in Section 5.

2. SYSTEM MODEL AND CHANNEL ESTIMATION

A system with M transmitter antennas and P receiver antennas is considered. Without loss of generality, each transmit antenna can be assumed to transmit different information symbols. The received signal vector at the p-th receive antenna during the n-th symbol period can be expressed by the discrete-time model,

$$r_p(n) = \sum_{m=1}^{M} \sum_{k=0}^{L-1} h_{m,p}(k) s_m(n-k) + w_p(n) \tag{1.1}$$

where $\mathbf{h}_{m,p} \triangleq [h_{m,p}(0), ..., h_{m,p}(L-1)]^T$ is the impulse response of the channel between the m-th transmit antenna and the p-th receive antenna, $s_m(n)$ is the transmitted symbol from the m-th transmit antenna and $w_p(n)$ is the additive noise at the p-th receive antenna. Each of the channel impulse responses is assumed to be of length L taps. During the training phase different training sequences are transmitted from each of the transmit antennas. The training sequences are assumed to be N symbols long, and the channel impulse response is estimated at the receiver based on the knowledge of the training symbols.

The vector of observations at the p-th receive antenna, during the training phase can be written in matrix form as,

$$\mathbf{r}_p = \mathbf{S}\mathbf{h}_p + \mathbf{w}_p \tag{1.2}$$

where $\mathbf{r}_p \triangleq [r_p(N), \ \cdots, \ r_p(L)]^T$ is the vector of observations, $\mathbf{h}_p \triangleq [h_{1,p}(0), \ \cdots, h_{1,p}(L-1), \ \cdots, h_{M,p}(0), \ \cdots, h_{M,p}(L-1)]^T$ is the stacked vector of channel impulse responses, $\mathbf{w}_p \triangleq [w_p(N), \ \cdots, \ w_p(L)]^T$ is the noise vector and $\mathbf{S}$ is an $(N-L+1) \times ML$ block-Toeplitz matrix consisting of the training symbols,

$$\mathbf{S} \triangleq \begin{bmatrix} s_1(N) & \cdots & s_1(N-L+1) & \cdots & s_M(N) & \cdots & s_M(N-L+1) \\ \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\ s_1(L) & \cdots & s_1(1) & \cdots & s_M(L) & \cdots & s_M(1) \end{bmatrix} \tag{1.3}$$

The stacked impulse response vector $\mathbf{h}_p$ is estimated at the receiver for each of the P receiver antennas. An LS channel estimator is considered for that purpose. An LS channel estimator minimizes the squared error between the received signal vector and the reconstructed signal based on the channel estimate. The LS estimate is given by,

$$\hat{\mathbf{h}}_p^{LS} = \arg\min_{\mathbf{h}} \|\mathbf{r}_p - \mathbf{S}\mathbf{h}\|^2 = \left(\mathbf{S}^H\mathbf{S}\right)^{-1}\mathbf{S}^H\mathbf{r}_p. \tag{1.4}$$

In equation (1.4) the auto-correlation matrix $\mathbf{S}^H\mathbf{S}$ is assumed to be invertible. The LS estimate can be expressed in terms of the channel impulse response as, $\hat{\mathbf{h}}_p^{LS} = \mathbf{h}_p + \left(\mathbf{S}^H\mathbf{S}\right)^{-1}\mathbf{S}^H\mathbf{w}_p$. If the additive noise is zero-mean and uncorrelated to the training sequence, the LS estimate is unbiased. The LS channel estimation error is,

$$E\left[\|\mathbf{h}_p - \hat{\mathbf{h}}_p^{LS}\|^2\right] = \mathrm{tr}\left\{\left(\mathbf{S}^H\mathbf{S}\right)^{-1}\mathbf{S}^H\mathrm{E}\left[\mathbf{w}_p\mathbf{w}_p^H\right]\mathbf{S}\left(\mathbf{S}^H\mathbf{S}\right)^{-1}\right\}. \tag{1.5}$$

Under the assumption that the noise process is white with a variance of σ_w^2, the LS estimation error simplifies to

$$E\left[\|\mathbf{h}_p - \hat{\mathbf{h}}_p^{LS}\|^2\right] = \mathrm{tr}\left\{\sigma_w^2\left(\mathbf{S}^H\mathbf{S}\right)^{-1}\right\}. \tag{1.6}$$

3. TRAINING SEQUENCE DESIGN

The training sequence $\mathbf{S}$ is optimized to minimize the LS estimation error of equation (1.6). The minimum LS error is obtained iff (can be proved by arguments similar to those in [8])

$$\mathbf{S}^H\mathbf{S} = (N - L + 1)\sigma_s^2\mathbf{I}_{ML}, \tag{1.7}$$

where σ_s^2 is the variance of the training symbols. The minimum value of the estimation error is

$$\min E\left[\|\mathbf{h}_p - \hat{\mathbf{h}}_p^{LS}\|^2\right] = \frac{ML\sigma_w^2}{(N - L + 1)\sigma_s^2}. \tag{1.8}$$

The above result is analogous to the single transmit antenna scenario and is equivalent to choosing the training sequences to be temporally white and spatially uncorrelated (i.e., across transmit antennas).

A critical parameter in training sequence design is the length of the sequence. The training sequence needs to be long enough for the channel to be identified. A longer training sequence has the added advantage of reducing the channel estimation error. However, an increase in training sequence length results in a decrease in the useful data rate of the transmission. The design tradeoffs associated with the choice of training sequence length are discussed in this section.

Identifiability: For the channel impulse response to be identifiable, the auto-correlation matrix $\mathbf{S}^H\mathbf{S}$ of equation (1.4) has to be invertible. Hence, the training sequence matrix $\mathbf{S}$ has to be of full column rank, i.e.,

$$(N - L + 1) \geq ML \tag{1.9}$$

Loss due to Channel Estimation: Any error resulting from channel estimation can be incorporated into the noise process and can be quantified as a loss in effective SNR. However, the statistics of the noise process will now be different. If the noise process is assumed to be uncorrelated with the source symbols the mean-squared error is given by,

$$\mathrm{MSE} = \|\mathbf{h}_p - \hat{\mathbf{h}}_p^{LS}\|^2\sigma_s^2 + \sigma_w^2 = \sigma_w^2\left\{1 + \frac{ML}{N-L+1}\right\}. \tag{1.10}$$

Equation (1.10) assumes that the source variance equals the variance of the training symbols. The increase in MSE due to channel estimation error can be interpreted as a loss in effective SNR at the receiver.

Loss in Throughput: The throughput of a system is the product of the data rate and the probability of successful transmission of a packet. Clearly, from equation (1.10), the longer the training sequence, the lesser the channel estimation error. A smaller channel estimation error results in a decrease in packet error rate. However, an increase in the training sequence length reduces the number of information bits that can be transmitted in a packet and hence the data rate. Since the training symbols come at the cost of the data symbols, the throughput depends on the length of the training sequence. Hence, a good criterion for the design of training length would be to maximize the achievable throughput.

4. OPTIMAL SEQUENCES

Once the training sequence length has been decided, it becomes essential to search for training sequences with good properties. Equation (1.7) specifies the optimality criterion for the search of such sequences. The sequences have to be of simple alphabets in order to guarantee low complexity realization. Two types of sequences are common: aperiodic and periodic sequences. While aperiodic sequences exist for many lengths, periodic ones are much harder to find. However, due to an expansion theorem[9], short periodic sequences can be concatenated to very large sequences preserving their orthogonal properties. Table 1 lists known periodic sequences of simple alphabets. Originally applied to single antenna systems, the periodic sequences can be used for multiple antenna systems. Construction of multiple antenna training sequences from the periodic sequence is described later in this section under "*Cyclic Shift Search*". The QPSK sequence of length $L = 16$ was proposed to extend existing OFDM systems to four transmit and receive antennas [10].

It is possible that optimum training sequences may not exist for a particular choice of training length and channel delay spread. In that case training sequences with near optimal properties can be searched for. A few heuristic methods for the search of such sequences are discussed in this section. A few near-optimal binary training sequences are listed in Appendix B. The search for these sequences are based on the methods described in this section.

Full search: Near-optimal sequences can be obtained by searching over all possible sequences and choosing those which have the minimum value of $\mathrm{tr}\{(\mathbf{S}^H\mathbf{S})^{-1}\}$. However, the search has to be done over $|\mathcal{A}|^{MN}$ sequences, where $|\mathcal{A}|$ is the number of points in the source constellation. This search is computationally prohibitive. Hence, heuristic methods that search over a reduced set of sequences are of special interest.

Random Search: From equation (1.7) it is clear that near-optimal sequences should have good auto-correlation and cross-correlation properties, i.e., small non-peak auto-correlation terms over a window of size $L-1$ on either side of the peak location and small cross-correlation terms for a window of length $2L-1$. To begin with, sequences with good auto-correlation properties can be determined by searching over all the $|\mathcal{A}|^N$ possible sequences. The number of such sequences can be expected to be much smaller than $|\mathcal{A}|^{MN}$. This is followed by a search for M sequences with good cross-correlation properties from this reduced set of sequences.

Cyclic Shift Search: Consider the sequence $t_1 = [s(1) \ \cdots \ s(N')]$ of length N', where $N' = N - L + 1$. The sequences $t_2, \ \cdots, \ t_M$ are now constructed by cyclic-shifts of the sequence t_1. For example, the sequence $t_{k+1} = [s(k\delta + 1) \ \cdots \ s(N') \ s(1) \ \cdots \ s(k\delta)]$ is obtained by a cyclic-shift of $k\delta$ of the sequence t_1, where $\delta = \lfloor \frac{N'}{M} \rfloor$. New sequences $s_1, \ \cdots, \ s_M$ are constructed by adding a cyclic-prefix of length $L-1$ to the sequences $t_1, \ \cdots, \ t_M$. For example, $s_1 = [s(N' - L + 2) \ \cdots \ s(N') \ s(1) \ \cdots \ s(N')]$ is one such sequence derived from the original sequence. Note that the new sequences s_k are of length N.

If the sequence t_1 has a cyclic auto-correlation function with zero off-peak terms and if $\delta \geq L$, then it is easy to see that equation (1.7) will be satisfied for the choice of training sequences $s_1, \cdots, s_M$. However, when searching for near-optimal training sequences the restriction of zero off-peak terms for the cyclic auto-correlation function can be relaxed and small off-peak cyclic auto-correlation terms can be allowed. This restricts the search space to a size $|\mathcal{A}|^{N-L+1}$.

5. DELAY-DIVERSITY SCHEME

Consider the use of delay diversity when two antennas are available at the transmitter. In delay diversity technique the same information symbols are transmitted from the two transmit antennas with a single symbol delay on the second antenna. This has the advantage that an optimized equalizer is sufficient to decode the delay-diversity code. For the delay diversity scheme, the received signal at the p-th receive antenna at the n-th symbol period is,

$$\begin{aligned} r_p(n) &= \sum_{m=1}^{2} \sum_{k=0}^{L-1} h_{m,p}(k) s(n-k-m+1) + w_p(n) \\ &= \sum_{k=0}^{L} \{\tilde{h}_{1,p}(k) + \tilde{h}_{2,p}(k)\} s(n-k) + w_p(n) \end{aligned} \tag{1.11}$$

where $\tilde{\mathbf{h}}_{1,p}$ and $\tilde{\mathbf{h}}_{2,p}$ are augmented channel impulse response vectors of length $L+1$ taps such that $\tilde{\mathbf{h}}_{1,p} = [h_{1,p}(0) \ \cdots \ h_{1,p}(L-1) \ 0]^T$ and $\tilde{\mathbf{h}}_{2,p} = [0 \ h_{2,p}(0) \ \cdots \ h_{2,p}(L-1)]^T$. The equivalent channel impulse response $\mathbf{h}_p^{eq}$ is a sum of these augmented channel impulse response vectors and it is sufficient to estimate $\mathbf{h}_p^{eq}$ at the receiver.

During the training phase, the transmitter, however, has the option of transmitting different training sequences from each of the two transmitter antennas. The individual channel impulse responses can then be estimated and summed up with the appropriate delay to obtain the equivalent channel impulse response. A second possibility would be to transmit the same training sequence from both the transmitter antennas, one of them with a single symbol delay, and estimate the equivalent channel response directly. Intuitively, the idea of using the same training sequence is appealing and it will be shown that this in fact is a better choice.

First, we consider the use of the same training sequence from both the transmit antennas. From equation (1.11), it is clear that the LS estimate of the equivalent channel impulse response $\mathbf{h}_p^{eq}$ can be determined like that of a SISO system. Since, the same training sequence is used from both the antennas, the Toeplitz matrix $\mathbf{S}$ has a size $(N-L)\times(L+1)$. Based on analysis similar to ([8]), it can be shown that the minimum possible LS estimation error is,

$$\min \mathrm{E}\left[\|\mathbf{h}_p^{eq}-\hat{\mathbf{h}}_p^{eq}\|^2\right] = \frac{\sigma_w^2(L+1)}{\sigma_s^2(N-L)}, \tag{1.12}$$

and is obtained iff

$$\mathbf{S}^H\mathbf{S} = (N-L)\sigma_s^2\mathbf{I}_{L+1}. \tag{1.13}$$

Now we consider the use of different training sequences from the two transmit antennas. Note that the training sequence from the second antenna is transmitted with a symbol delay. The received sequence at the p-th receive antenna, during the training phase, can be expressed as,

$$\begin{aligned} r_p(n) &= \sum_{m=1}^{2}\sum_{k=0}^{L-1} h_{m,p}(k)s_m(n-k-m+1)+w_p(n) \\ &= \sum_{k=0}^{L}\{\tilde{h}_{1,p}(k)s_1(n-k)+\tilde{h}_{2,p}s_2(n-k)\}+w_p(n) \end{aligned} \tag{1.14}$$

and the stacked channel impulse response vector $\tilde{\mathbf{h}}_p$ can be estimated as in equation (1.4). The only difference is that $\tilde{\mathbf{h}}_p$ has $2(L+1)$ taps and $\mathbf{S}$ is an $(N-L)\times 2(L+1)$ block-Toeplitz matrix. The equivalent channel impulse response $\mathbf{h}_p^{eq}$ can be expressed in terms of $\tilde{\mathbf{h}}_p$ as, $\mathbf{h}_p^{eq} = \mathbf{A}\tilde{\mathbf{h}}_p$, where the $(L+1)\times 2(L+1)$ matrix $\mathbf{A}$ is constructed as,

$$\mathbf{A} = \begin{bmatrix} \mathbf{I}_L & \mathbf{0}_{L\times 2} & \mathbf{0}_{1\times L} \\ \mathbf{0}_{1\times L} & \mathbf{0}_{1\times 2} & \mathbf{I}_L \end{bmatrix} \tag{1.15}$$

The LS error for the equivalent channel impulse response is,

$$E\left[\|\mathbf{h}_p^{eq}-\hat{\mathbf{h}}_p^{eq}\|^2\right] = \mathrm{tr}\left\{\sigma_w^2\mathbf{A}\left(\mathbf{S}^H\mathbf{S}\right)^{-1}\mathbf{A}^T\right\}. \tag{1.16}$$

The training sequence $\mathbf{S}$ has to be optimized to minimize the LS channel estimation error of equation (1.16). A lower bound for the LS error is derived in Appendix A and is given by,

$$E\left[\|\mathbf{h}_p^{eq}-\hat{\mathbf{h}}_p^{eq}\|^2\right] \geq \frac{\sigma_w^2(L+1)}{\sigma_s^2(N-L)}. \tag{1.17}$$

Comparing equation (1.12) with the lower bound obtained in equation (1.17), it is clear that the choice of identical training sequences for the two transmit antennas, provided equation (1.13) is satisfied, is indeed optimal.

6. CONCLUSION

In this paper, optimal training sequences for wireless systems with multiple antennas were discussed. After deriving an optimality criterion for training sequence design for MIMO systems, many aperiodic and periodic sequences were presented. A few heuristic methods for the search of near-optimal sequences were proposed. It was further shown that, the choice of identical training sequences, transmitted with the appropriate delays, is optimal for the delay-diversity scheme.

Appendix: A

The auto-correlation matrix $\mathbf{R} = \mathbf{S}^H\mathbf{S}$ is obviously Hermitian and positive definite. We assume that the maximum energy points of the source constellation are used as the training symbols. Hence the diagonal entries of $\mathbf{R}$ are equal to $\sigma_s^2(N-L)$. Let $\mathbf{a}_1, \cdots, \mathbf{a}_{L+1}$ be the row vectors of the augmentation matrix $\mathbf{A}$. Each of this row vectors $\mathbf{a}_k$ has non-zero entries, namely unity, at the k-th location and the $(L+1+k)$-th location, except for $\mathbf{a}_1$ and $\mathbf{a}_{L+1}$. The vectors $\mathbf{a}_1$ and $\mathbf{a}_{L+1}$ are vectors with an entry of one at the 1-st and the $2(L+1)$-th locations, respectively. The LS error of equation (1.16) can be expressed as a function of the row vectors $\mathbf{a}_k$ as,

$$E\left[\|\mathbf{h}_p^{eq} - \hat{\mathbf{h}}_p^{eq}\|^2\right] = \sigma_w^2 \sum_{k=1}^{L+1} \mathbf{R}^{-1}\mathbf{a}_k, \mathbf{a}_k \quad . \tag{1.A.1}$$

<u>*Kantorovich inequality:*</u> If $\mathbf{R} \in \mathcal{C}^{n\times n}$ is a positive definite Hermitian matrix and $x \in \mathcal{C}^n$ is a vector, then

$$\langle x, x\rangle^2 \leq \langle \mathbf{R}x, x\rangle\langle \mathbf{R}^{-1}x, x\rangle. \tag{1.A.2}$$

For a k value of 2, ..., L one can see that $\langle \mathbf{a}_k, \mathbf{a}_k\rangle = 2$ and

$$\langle \mathbf{R}\mathbf{a}_k, \mathbf{a}_k\rangle = r_{k,k} + r_{k,k+L} + r_{k+L,k} + r_{k+L,k+L} \tag{1.A.3}$$

where $r_{i,j}$ are the entries of the auto-correlation matrix $\mathbf{R}$. Auto-correlation matrices have the property that the off-diagonal entries are no larger than the diagonal entries. Hence,

$$\langle \mathbf{R}\mathbf{a}_k, \mathbf{a}_k\rangle \leq 4\sigma_s^2(N-L). \tag{1.A.4}$$

From equations (1.A.2) and (1.A.4), for the particular choice of k between 2 to L,

$$\langle \mathbf{R}^{-1}\mathbf{a}_k, \mathbf{a}_k\rangle \geq \frac{1}{\sigma_s^2(N-L)}. \tag{1.A.5}$$

It is easy to see that equation (1.A.5) is also satisfied when $k=1$ and $k=L+1$. Hence equation (1.A.1) reduces to,

$$\mathrm{E}\left[\|\mathbf{h}_p^{eq} - \hat{\mathbf{h}}_p^{eq}\|^2\right] \geq \frac{\sigma_w^2(L+1)}{\sigma_s^2(N-L)}. \tag{1.A.6}$$

Table 1 Periodic sequences for channel estimation of length $L(A = \{-2,-1,1,2\}, B = \{-8,-4,-2,-1,1,2,4,8\})$.

L=2	1,-j	QPSK
L=3	-2,-2,1	A
L=3	-3, 3+3j,3+3j	V29
L=4	1,1,1,-1	BPSK
L=5	-3+3j,-3j,-3+3j,3+3j,3+3j	V29
L=6	-1,1,-1,1,-1,-2	A
L=6	-3+3j,-1+3j,-1-j,1-3j,-1+3j,-1-j	16QAM
L=6	3-3j,-3,3-3j,3+3j,3j,3+3j	V29
L=7	-2,-2,-1,1,1,-2,1	A
L=7	1-j,1-j,1-j,1-j,1-j,1-j,5j	V29
L=8	1,-j,1,-1,-1,-j,-1,-1	QPSK
L=9	-2,-8,1,-2,1,1,-2,1,1	B
L=9	-3-3j,-3+3j,3,-3+3j,-3-3j,3+3j,3+3j,3+3j,3+3j	V29
L=10	3-j,3+j,3-j,-3+3j,1+3j,-1-j,-3+j,-1-j,1+3j,-3+3j	16QAM
L=12	-2,-2,-2,-1,1,-2,-2,2,-2,1,1,2	A
L=14	1-j,1-j,1-j,1-j,1-j,1-j,5j,1+j,1+j,1+j,1+j,1+j,1+j,-5	V29
L=15	2,-2,-2,1,-2,2,1,1,-2,1,2,1,1,1,1	A
L=16	1,1,1,1,1,j,-1,-j,1,-1,1,-1,1,-j,-1,j	QPSK
L=18	2,-2,1,-2,1,1,-1,1,1,-2,-2,1,2,1,1,1,1,1	A
L=19	-2,-2,1,2,-2,1,1,-2,-2,-2,-2,1,-2,1,-2,1,1,1,1	A
L=21	-2,1,-2,1,1,1,1,-2,-2,1,1,-2,1,1,1,1,1,1,1,1,1	A

Table 2 Training Sequences for $M = 2$.

Antenna 1	**Antenna 2**
0FB5D8F	293BE29
0391483	251F725
3785377	0BB9F4B
3BB287B	0B4188B
1D2F9DD	21135E1
11182D1	21EB221
2F0A6EF	1773E97
3DD943D	05A0C45

Appendix: B

A few near-optimal binary training sequences for a system, incorporating multiple antennas, are shown in this section. A training length of $N = 26$ was chosen for the

Table 3 Training Sequences for $M = 4$

Antenna 1	**Antenna 2**	**Antenna 3**	**Antenna 4**
0A7076510	7076510A7	76510A707	510A70765
2F9291822	9291822F9	91822F929	822F92918
517A46305	7A4630517	4630517A4	30517A463
C2D45980C	D45980C2D	5980C2D45	80C2D4598
2D8B8E402	8B8E402D8	8E402D8B8	402D8B8E4
B6E05238B	E05238B6E	5238B6E05	38B6E0523
59B80A8E5	B80A8E59B	0A8E59B80	8E59B80A8
CC876AEBC	876AEBCC8	6AEBCC876	EBCC876AE

2 transmit antenna case. This choice was made based on the various design metrics described in Section 3. The delay spread L of the channel impulse response was assumed to be 7 for the training sequence design. The training symbols were determined based on the random search method described in Section 4 and were restricted to a BPSK constellation. A few pairs of these sub-optimal training sequences are shown in hexadecimal format Table 2. The most-significant-bit (MSB) of the hexadecimal representation corresponds to the first symbol of the training sequence. The bit 1 corresponds to the symbol "+1" and the bit 0 to the symbol "-1". Figure 1 illustrates the loss incurred by the sub-optimal training sequences over the ideal training sequences.

For the four transmit antenna case, namely $M = 4$, a choice of $N = 36$ was made for the training length. For the training sequence design, a delay spread of $L = 5$ was assumed. A few sets of sub-optimal BPSK training sequences are tabulated in hexadecimal format in Table 3. These training sequences were obtained using the cyclic shift search method described in Section 4. Figure 2 illustrates the loss incurred by the sub-optimal training sequences over the ideal training sequences.

Figure 1 Loss due to channel estimation, M=2

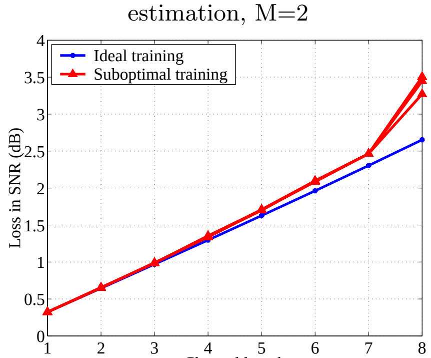
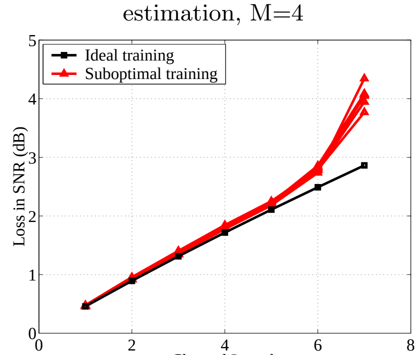


Figure 2 Loss due to channel estimation, M=4



References

[1] G.J. Foschini, M.J. Gans, (1998). *On limits of wireless communication in a fading environment when using multiple antennas,* Wireless Personal Communications, vol 6. No. 3, pp 311-335.

[2] J.H. Winters, J. Salz, R.D. Gitlin (1999). *The impact of antenna diversity on the capacity of wireless communication systems,* IEEE Transactions on Communications, vol 47, no. 7, pp 1073-1083.

[3] V. Tarokh, N. Seshadri and A. R. Calderbank (1998). *Space-time codes for high data rate wireless communications: performance Analysis and code construction,* IEEE Transactions on Information Theory, vol 44, no. 2, pp 744-765.

[4] G.J. Foschini, G.D. Golden, R.A. Valenzuela and P.W. Wolaniansky (1999). *Simplified processing for high spectral efficiency wireless communication employing multi-element arrays,* IEEE Journal on Selected areas in communications, vol. 17, no. 11, pp 1841-1852.

[5] E. de Carvalho, D.T.M. Slock (1997). *Cramer-Rao bounds for semi blind, blind and training sequence based channel estimation,* Proc. Signal Processing Advances in Wireless Comm., Paris, France.

[6] S.N. Crozier, D.D. Falconer, S.A. Mahmoud, (1991). *Least sum of squared errors (LSSE) channel estimation,* IEE Proc. F, vol. 138, no. 4, pp. 371-378.

[7] C. Tellambura, M.G.Parker, Y.Jay Guo, S.J. Sheperd, S.K. Barton (1999). *Optimal sequences for channel estimation using discrete Fourier transform techniques,* IEEE Trans. Communications, vol. 47, no. 2, pp. 230-237.

[8] C. Tellambura, Y.Jay Guo, S.K. Barton (1998). *Channel estimation using aperiodic training sequences*, IEEE Communication Letters, vol. 2, no.5, pp. 140-142.

[9] M. Rupp (2000). *FAST Implementation of the LMS Algorithm,* Proc. Eusipco 2000, Tampere, September 2000.

[10] R. Van Nee, R. Prasad (2000). *OFDM for Wireless Multimedia Communications,* Artech House Publishers.

Jaiganesh Balakrishnan received his B. Tech in Electrical Engineering from Indian Institute of Technology, Madras in 1997. He received his M.S degree in Electrical Engineering from Cornell university in August 1999 and is currently pursuing his Ph.D. In Summer 1998 and 2000 he stayed with the Bell-Labs, Lucent Technologies, Wireless Research Lab in Crawford Hill, NJ, for an internship. His research interests include equalization, detection and estimation, adaptive signal processing and wireless communications.

Markus Rupp received his Diploma in Electrical Engineering from FHS Saarbrücken, Germany and Universität of Saarbrücken in 1984 and 1988, respectively. During this time he was also a lecturer in Digital Signal Processing and High Frequency Techniques at the FHS. He finished his Doctoral degree in 1993 with summa cum laude at the TU Darmstadt in the field of acoustical echo compensation. He was awarded a DAAD postdoctoral fellowship and spend the time from November 1993 until September 1995 in the Dept. of Electrical and Computer Engineering of the University of California, Santa Barbara working with Dr. Ali Sayed on new robustness theories for adaptive filters. Since October 1995 he has been with Bell-Labs, Lucent Technologies (formerly part of AT&T), Wireless Research Lab in New Jersey, working on wireless phones and implementation issues of wireless modems. In May 1999 he moved to the Netherlands joining Bell-Labs efforts in wireless LANs in Europe. He is currently involved in Rapid Prototyping methods for wireless systems. Dr. Rupp served at session chair for adaptive filters in communications at the 1998 Asilomar conference and has more then 60 publications and patents in the fields of adaptive filters and wireless systems.

Harish Viswanathan received his B. Tech in Electrical Engineering from Indian Institute of Technology, Madras in 1992. He received an M.S. and the Ph.D. degrees in Electrical Engineering from Cornell University, Ithaca, NY in January 1995 and August 1997, respectively. He was awarded the Cornell Sage Fellowship during the academic year 1992-93. Since October 1995 he has been with Bell-Labs, Lucent Technologies (formerly part of AT&T), Wireless Research Lab in New Jersey working on various physical layer and systems engineering issues for wireless communication systems. His research interests include information theory, coding theory, communication networks and signal processing