Channel Tracking for Multiple Input, Single Output Systems using EM algorithm

Carlos H. Aldana and John Cioffi
STAR Lab, Electrical Engineering Dept.
Stanford University

Abstract—**This paper investigates the problem of blindly acquiring the channel gains for a synchronized multiuser system using the expectation maximization (EM) algorithm. The EM algorithm takes advantage of the finite alphabet property of the transmitted signal. It also provides MMSE estimates of the transmitted data that can be used by the receiver for decoding purposes. The algorithm has been applied to a multicarrier system and results show that the application of EM in the high SNR case provides significant improvement over traditional channel estimation techniques.**

I. INTRODUCTION

Traditionally, channel estimation has been done by periodically sending training sequences through the channel. A new technique that both reduces the amount of training needed and finds the maximum likelihood (ML) estimate of the channel in the high SNR case is proposed.

This paper considers the problem of channel identification in an ISI (intersymbol interference) free communications system with K synchronized transmitters. Each transmits digital signals at the same time and in the same bandwidth using the same symbol period T, as shown in Fig. 1. H_j denotes the scalar gain for the jth user. The output at the receiver is a superposition of signal waveforms plus additive white Gaussian noise. An example of such system is transmission in a wireless system from the base stations to the mobile where a narrowband signal is transmitted through an unknown multipath propagation environment. Another example is the wireline case where the central office employing Discrete Multitone (DMT) modulation wishes to find the crosstalk coupling functions of synchronized remote units.

We apply a training-aided EM (Expectation Maximization) algorithm to a multicarrier system, which has no ISI. Training data provides the initial condition for EM. The principle behind multicarrier modulation (MCM) is to superimpose several carrier-modulated waveforms in parallel subchannels [1]. A MISO (Multiple Input, Single Output) system consists of a single modem or antenna which receives signals from multiple sources. In a MISO system employing MCM, each user must have a long enough cyclic prefix so that data is transmitted on orthogonal subchannels. On each subchannel l, the received subsymbol can be written as $Y(l) = \sum_{j=1}^{K} H_j(l)X_j(l)+N(l)$, where $H_j(l)$ is the subchannel gain on the lth tone for the jth user, $X_j(l)$ is the transmitted signal for the jth user, and $N(l)$ is the noise (assumed to be block-stationary Gaussian noise). Each subchannel looks like Fig. 1.

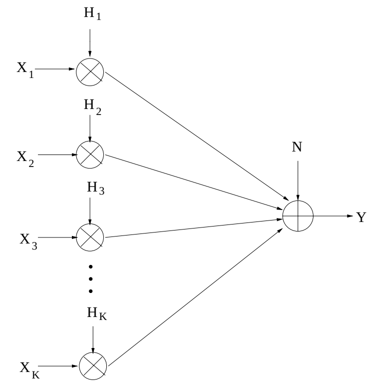

Fig. 1. Multiple input, single output system

For the system described above, the EM algorithm is applied in the frequency domain on each tone. In [2], it was shown that in order to acquire an exact estimate of the channel, at least $2^{\lceil \log_2 \nu \rceil}$ EM algorithms need to be run in parallel. ν is the number of taps of the channel and $\lceil . \rceil$ denotes the ceiling operator. The algorithm is able to exploit the finite alphabet property and the distribution of the transmitted signals. It acts on blocks of data in which the channel is assumed to remain stationary, or constant. This assumption is valid in the wireline scenario and also holds for high data rate wireless applications. This is not true in general. The validity should be checked against the value of the Doppler bandwidth B_D of the multipath channel. If one takes a block of length L containing G time samples spaced at time interval T_s, then the multipath channel can be considered block stationary if $B_D LGT_s < 0.01$.

Previous applications of the EM algorithm for the Single Input Single Output (SISO) case include [3], [4], [5], [6], [7], [8], [9]. Iterative solutions that exploit the finite alphabet property have been found in [10], [11]. In addition, the cyclic prefix information is used in [12] while precoding the input to acquire channel estimates for the SISO case is done in [13]. Here we solve the finite alphabet MISO case.

This paper is organized as follows. Section II describes the system model. Section III describes a typical channel estimation procedure with training data available. Section IV introduces and solves the EM formulation when the transmitted data is not available at the receiver. Section V provides simulation results that compare training only with EM and Section VI provides concluding remarks.

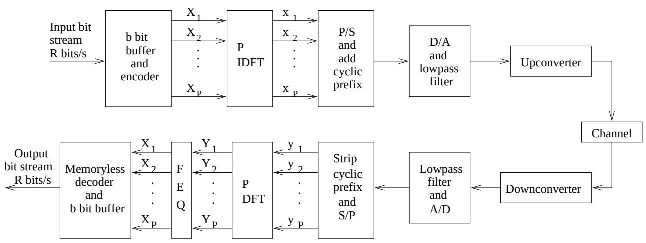

Fig. 2. Multicarrier System

II. System Model

In a multicarrier modulation system, each user loads data, as in Fig. 2, by means of an IDFT onto each subcarrier. A cyclic prefix is appended to make the transmit signal appear periodic. This prefix is necessary to maintain orthogonality of the subchannels and to prevent intersymbol interference (ISI). We assume that K synchronized users are transmitting at the same time with the same symbol period. Over one block of data transmission, the channel output in the time domain can be written as

$$y(t) = \sum_{j=1}^{K} h_j(t) * x_j(t) + n(t) \tag{1}$$

where $*$ denotes convolution, and $h_j(t)$ is the sampled impulse response of the channel corresponding to the jth user. We assume that all the channel impulse responses to be of finite length less than or equal to ν. Let $\mathcal{H}_j$ be the matrix representation of this block transmission so that

$$\mathbf{y} = \sum_{j=1}^{K} \mathcal{H}_j \mathbf{x}_j + \mathbf{n} \tag{2}$$

where $\mathbf{x}_j = [x_j(0) \ldots x_j(P-1)]^T$, $\mathbf{y} = [y(0) \ldots y(P-1)]^T$, and $\mathbf{n} = [n(0) \ldots n(P-1)]^T$ are the block of P data symbols for the jth user, the block of received symbols and block of P additive white Gaussian noise samples. The notation $[\,.\,]^T$ denotes the transpose operation.

Assuming the first ν symbols of $\mathbf{x}_j$ form the cyclic prefix that is added to the transmitted data block, the $P \times P$ channel matrix is given by

$$\mathcal{H}_j = \begin{bmatrix} h_j(0) & 0 & \cdots & h_j(\nu) & \cdots & h_j(1) \\ \vdots & h_j(0) & 0 & 0 & \cdots & h_j(2) \\ h_j(\nu) & \vdots & & & & \vdots \\ \vdots & h_j(\nu) & & \ddots & & \vdots \\ \vdots & \vdots & \ddots & & & \\ 0 & 0 & \cdots & h_j(\nu)\cdots & h_j(1) & h_j(0) \end{bmatrix} \tag{3}$$

The eigenvalue decomposition of $\mathcal{H}_j$ is given by $\mathcal{H}_j = Q^* \Lambda_j Q$ where Q and Q^* are the FFT and IFFT matrices, and Λ_j is a diagonal matrix. Λ_j is given by $diag(H_j(0), H_j(1), \ldots, H_j(P-1))$. If $\mathbf{Y} = Q\mathbf{y}$ and $\mathbf{x}_j = Q^*\mathbf{X}_j$, where $\mathbf{Y} = [Y(0)\, Y(1) \ldots Y(P-1)]^T$ and $\mathbf{X}_j = [X_j(0)\, X_j(1) \ldots X_j(P-1)]^T$, then

$$Y(l) = \sum_{j=1}^{K} H_j(l) X_j(l) + N(l) \qquad l = 0, 1, \ldots, P-1 \tag{4}$$

where $H_j(l)$ is the lth subchannel gain for the jth user, $X_j(l)$ is the transmitted symbol on the lth subchannel for the jth user, and $N(l)$ is assumed to be zero mean white Gaussian noise.

III. Channel Gain and Noise Variance Estimates using Training Sequence

Due to the subchannel orthogonality, training and EM analysis in Sections III and IV apply for each tone, allowing (l) to be neglected. From now on, the superscript n denotes time. With the aid of training, the channel gains on each tone can be easily acquired. If for each tone or subchannel, we collect L_{tr} samples and form the column vector $\mathbf{Y} = [Y^1, Y^2, \ldots Y^{L_{tr}}]^T$, we can write (4) as

$$\mathbf{Y} = \mathcal{X}\mathbf{H} + \mathbf{N} \tag{5}$$

where $\mathcal{X} = [\mathbf{X}^1 \mathbf{X}^2 \ldots \mathbf{X}^{L_{tr}}]^T$, $\mathbf{X}^n = [X_1^n, X_2^n, \ldots X_K^n]^T$, $\mathbf{H} = [H_1 H_2 \ldots H_K]^T$, and $\mathbf{N} = [N^1, N^2, \ldots N^{L_{tr}}]^T$. $\mathbf{X}^n$ consists of training data from K users at time n. We assume that the noise $\mathbf{N}$ is white, zero-mean Gaussian with covariance matrix $\sigma^2\mathbf{I}$, where $\mathbf{I}$ is the identity matrix. The maximum likelihood estimate of the parameters $\mathbf{H}$ and σ^2 is given by $(\mathbf{H}_{ML}, \sigma^2_{ML}) = \arg\max f(\mathbf{Y}|\mathbf{H}, \sigma^2, \mathcal{X})$, where

$$f(\mathbf{Y}|\mathbf{H}, \sigma^2, \mathcal{X}) = \frac{1}{\sqrt{(2\pi\sigma^2)^{L_{tr}}}} \times \exp\{-\frac{1}{2\sigma^2}(Y - \mathcal{X}\mathbf{H})^T (Y - \mathcal{X}\mathbf{H})\}. \tag{6}$$

If $\mathcal{X}$ has full column rank, then it can be easily shown that the maximum-likelihood (ML) estimate of $\mathbf{H}$ is given by the least squares solution

$$\mathbf{H}_{ML} = (\mathcal{X}^T \mathcal{X})^{-1} \mathcal{X}^T \mathbf{Y}. \tag{7}$$

The error in this estimate, denoted by $\delta = \mathbf{H}_{ML} - \mathbf{H}$, is a zero mean Gaussian process with covariance given by $\sigma^2(\mathcal{X}^T\mathcal{X})^{-1}$. The complexity of (7) is of order $\mathcal{O}(L_{tr}K^2)$. We can also obtain a ML estimate of the noise variance by computing

$$\sigma^2_{ML} = \frac{1}{L_{tr}}(\mathbf{Y} - \mathcal{X}\mathbf{H}_{ML})^T (\mathbf{Y} - \mathcal{X}\mathbf{H}_{ML}) \tag{8}$$

Note that when $(\mathcal{X}^T\mathcal{X})^{-1} = \frac{\mathbf{I}}{L_{tr}E[X^2]}$, this estimate has expected value equal to $\sigma^2(1 - \frac{K}{L_{tr}})$. This estimate of the noise variance, along with the channel gains, can be used as initial conditions for the Expectation Maximization Algorithm described in the next section.

IV. Expectation Maximization Algorithm

In this section, the equiprobable input symbols belong to a finite alphabet and are considered unknown. The EM algorithm [14], [15] is one way of acquiring maximum likelihood (ML) estimates when evaluation of the likelihood is made difficult by the absence of certain data. The EM algorithm first finds its objective function, the conditional expectation of the joint log-likelihood, using the current estimate of the parameters. The maximization step then provides a new estimate of the parameters by maximizing this new objective function. This process is repeated until convergence is achieved. The initialization of EM is provided by using a training sequence.

The subscript k will now denote the iteration number. For a given set of observations $Y^1, Y^2, \ldots, Y^L$, it is desirable to find $\mathbf{H}$ such that the likelihood $\mathcal{L}(\mathbf{H}) = \prod_{n=1}^{L} f_Y(Y^n; \mathbf{H})$ is maximized. For each tone, white noise and a block stationary channel are assumed. For simplicity, it is assumed that each user transmits from an M point constellation. In Subsection IV-A we apply the EM algorithm to the case where the noise variance σ^2 is known for the duration of the block length while in Subsection IV-B, we apply the EM algorithm to the case where the noise variance is unknown as well.

A. EM Algorithm for Channel Gain Estimation with Known Noise Variance

Using the notation found in [15], let the complete data be denoted by $Z^n = (Y^n, \mathbf{X}^n)$ and form the vectors $\mathbf{Z} = (Z^1, Z^2, \ldots Z^L)$, $\mathbf{Y} = (Y^1, Y^2, \ldots Y^L)$. The objective function is $\log \mathcal{L}(\mathbf{H}) = \log f_{\mathbf{Y}}(\mathbf{Y}; \mathbf{H}) = \log f_{\mathbf{Z}}(\mathbf{Z}; \mathbf{H}) - \log f_{\mathbf{Z}|\mathbf{Y}}(\mathbf{Z}|\mathbf{Y}; \mathbf{H})$. $f_{\mathbf{Z}}(\mathbf{Z}; \mathbf{H})$ denotes the pdf of $\mathbf{Z}$ given $\mathbf{H}$. From the definition of complete data [15] and Bayes rule, one can write $f_{\mathbf{Z}|\mathbf{Y}}(\mathbf{Z}|\mathbf{Y}; \mathbf{H}) = \frac{f_{\mathbf{Z}}(\mathbf{Z};\mathbf{H})}{f_{\mathbf{Y}}(\mathbf{Y};\mathbf{H})}$. Since $\mathbf{Z}$ is unobservable, $\log f_{\mathbf{Z}}(\mathbf{Z}; \mathbf{H})$ is replaced with its conditional expectation given $\mathbf{Y}$, and $\mathbf{H}$ with $\mathbf{H}_k$ (k denotes the iteration number). The new objective function is

$$\begin{aligned} E[\log \mathcal{L}(\mathbf{H})|\mathbf{Y}, \mathbf{H}_k] = E[\log f_{\mathbf{Z}}(\mathbf{Z}; \mathbf{H})|\mathbf{Y}, \mathbf{H}_k] \\ -E[\log f_{\mathbf{Z}|\mathbf{Y}}(\mathbf{Z}|\mathbf{Y}; \mathbf{H})|\mathbf{Y}, \mathbf{H_k}]. \end{aligned} \tag{9}$$

This can be rewritten as

$$E[\log \mathcal{L}(\mathbf{H})|\mathbf{Y}, \mathbf{H}_k] = Q(\mathbf{H}, \mathbf{H}_k) - R(\mathbf{H}, \mathbf{H}_k), \tag{10}$$

where $Q(\mathbf{H}, \mathbf{H}_k) = E[\log f_{\mathbf{Z}}(\mathbf{Z}; \mathbf{H})|\mathbf{Y}, \mathbf{H}_k]$ and $R(\mathbf{H}, \mathbf{H}_k) = E[\log f_{\mathbf{Z}|\mathbf{Y}}(\mathbf{Z}|\mathbf{Y}; \mathbf{H})|\mathbf{Y}, \mathbf{H}_k]$. By Jensen's inequality, it can be shown that $R(\mathbf{H}_k, \mathbf{H}_k) \geq R(\mathbf{H}, \mathbf{H}_k)$ for any $\mathbf{H}$ [15]. To increase the objective function, it suffices to pick $\mathbf{H}_{k+1}$ such that $Q(\mathbf{H}_{k+1}, \mathbf{H}_k) \geq Q(\mathbf{H}, \mathbf{H}_k)$ for all $\mathbf{H}$. Thus, our objective is to maximize

$$Q(\mathbf{H}, \mathbf{H}_k) = E[\log f_{\mathbf{Z}}(\mathbf{Z}; \mathbf{H})|\mathbf{Y}, \mathbf{H}_k]. \tag{11}$$

At iteration time $k+1$, the function in (11) is maximized by setting the gradient $\nabla_{\mathbf{H}} Q(\mathbf{H}, \mathbf{H}_k)$ to 0,

$$\nabla_{\mathbf{H}} Q(\mathbf{H}, \mathbf{H}_k) = \sum_{n=1}^{L} E[\nabla_{\mathbf{H}} \log f_Z(Z^n; \mathbf{H})|\mathbf{Y}, \mathbf{H}_k] = 0. \tag{12}$$

This becomes an iterative procedure in which the root found at iteration $k+1$ takes the place of $\mathbf{H}_k$ and the root is found again until a stationary point is reached (i.e. $\mathbf{H}_{k+1} \approx \mathbf{H}_k$). The pdf of the complete data $f_Z(Z^n; \mathbf{H})$ can be derived using Bayes rule as

$$f_Z(Y^n, \mathbf{X}^n; \mathbf{H}) = f_{Y|\mathbf{X},\mathbf{H}}(Y^n|\mathbf{X}^n; \mathbf{H}) f_{\mathbf{X}}(\mathbf{X}^n) \tag{13}$$

where

$$f_{Y|\mathbf{X},\mathbf{H}}(Y^n|\mathbf{X}^n; \mathbf{H}) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(Y^n - \mathbf{H}^T \mathbf{X}^n)^2}{2\sigma^2}} \tag{14}$$

and $f_{\mathbf{X}}(\mathbf{X}^n) = \frac{1}{M^K}$, where M is the constellation size. Therefore,

$$\begin{aligned} \nabla_{\mathbf{H}} [\log f_Z(Y^n, \mathbf{X}^n; \mathbf{H})] = \\ -\frac{1}{2\sigma^2}(-2Y^n \mathbf{X}^n + 2\mathbf{X}^n (\mathbf{X}^n)^T \mathbf{H}). \end{aligned} \tag{15}$$

Using (15) in (12) and solving for $\mathbf{H}$, the update equation for the channel gains becomes

$$\hat{\mathbf{H}}_{k+1} = \left(\sum_{n=1}^{L} \overline{\mathbf{X}^n \mathbf{X}^{nT}} \right)^{-1} \sum_{n=1}^{L} Y^n \overline{\mathbf{X}^n}. \tag{16}$$

$\overline{\mathbf{X}^n} = E[\mathbf{X}^n | Y^n, \mathbf{H}_k]$ and $\overline{\mathbf{X}^n \mathbf{X}^{nT}} = E[\mathbf{X}^n \mathbf{X}^{nT} | Y^n, \mathbf{H}_k]$. Equation (16) is equivalent to

$$\hat{\mathbf{H}}_{k+1} = (\overline{\mathcal{X}^T \mathcal{X}})^{-1} \overline{\mathcal{X}}^T \mathbf{Y} \tag{17}$$

where $\overline{\mathcal{X}^T \mathcal{X}} = E[\mathcal{X}^T \mathcal{X} | \mathbf{Y}, \mathbf{H}_k]$ and $\overline{\mathcal{X}} = E[\mathcal{X} | \mathbf{Y}, \mathbf{H}_k]$. Equation (17) is the EM equivalent of (7). $\overline{\mathcal{X}}$ provides the minimum mean-squared error estimate of the data sequence based on the observations and can be used by the receiver to acquire soft estimates of the transmitted symbols. The expected value is taken over $f_{\mathbf{X}|Y,\mathbf{H}}$, which can be easily found using Bayes rule. For example,

$$\overline{\mathbf{X}^n} = \frac{\sum_{\mathbf{x} \in \Omega} \mathbf{x} f_{Y|\mathbf{X},\mathbf{H}}(Y^n | \mathbf{x}, \mathbf{H}_k)}{\sum_{\mathbf{x} \in \Omega} f_{Y|\mathbf{X},\mathbf{H}}(Y^n | \mathbf{x}, \mathbf{H}_k)} \tag{18}$$

where the sum is taken over all M^K possible combinations. The complexity of this update equation is dominated by the computation of $\overline{\mathcal{X}^T \mathcal{X}}$, which is of order $\mathcal{O}(LK^2 M^K)$. For each element in the $K \times K$ matrix $\overline{\mathbf{X}^n \mathbf{X}^{nT}}$, the conditional expectation is computed by summing over M^K points. This can be compared to ML, which has complexity of order $\mathcal{O}(LK^2 M^{KL})$. This is due to the computation of (7) for all M^{KL} possible values of $\mathcal{X}$.

B. EM Algorithm for Channel Gain and Noise Variance

The problem is cast in a similar fashion as above with the same complete and incomplete data. The results from the previous section apply except that σ^2 is replaced by σ_k^2, where again k denotes the iteration number. The updates on channel gains and noise variance are done alternatively until convergence is achieved. The channel update equation becomes

$$\hat{\mathbf{H}}_{k+1} = (\overline{\mathcal{X}^T\mathcal{X}})^{-1}\overline{\mathcal{X}}^T\mathbf{Y} \tag{19}$$

where $\overline{\mathcal{X}^T\mathcal{X}} = E[\mathcal{X}^T\mathcal{X}|\mathbf{Y}, \mathbf{H}_k, \sigma_k^2]$ and $\overline{\mathcal{X}} = E[\mathcal{X}|\mathbf{Y}, \mathbf{H}_k, \sigma_k^2]$. To get the noise variance update equation, we need to maximize the log likelihood with respect to σ^2. Taking the derivative of the log likelihood,

$$\nabla_{\sigma^2}[\log f_Z(Y^n, \mathbf{X}^n; \mathbf{H})] = -\frac{1}{2\sigma^2} + \frac{(Y^n - \mathbf{H}^T\mathbf{X}^n)^2}{2\sigma^4}. \tag{20}$$

The update equation for the noise variance estimate is:

$$\sigma_{k+1}^2 = \frac{1}{L}\sum_{n=1}^{L} E[(Y^n - \mathbf{H}^T\mathbf{X}^n)^2 \mid Y^n, \sigma^2 = \sigma_k^2, \hat{\mathbf{H}}_{k+1}]. \tag{21}$$

The expectation is done over $\mathbf{X}^n$ and its complexity is of the same order as the channel gain complexity, since finding $\sum_{n=1}^{L}\overline{\mathbf{X}^n\mathbf{X}^{nT}}$ is also needed.

V. Simulation Results

The EM algorithm is used for the wireless channel with $K = 2$ users. The initial conditions are acquired from (7) and (8). 20 random symbols of training from a block size of 200 symbols are used for 21000 randomly generated channels with $\|\mathbf{H}\| = 1$. Each user transmits ± 1 in each subchannel and the channel gains are chosen from an iid Gaussian random variable. The MSE of both the channel gain estimates and noise variance are compared for the case when the system uses the training data alone (20 symbols) and when it uses the EM algorithm. Figures 3 and 4, where "tr" denotes training, show the normalized MSE of the channel gains and the MSE of the noise variance estimate. The normalized channel gain MSE is defined as $MSE_{\mathbf{H}} = \frac{\|\mathbf{H}-\hat{\mathbf{H}}\|^2}{K}$ and the noise variance MSE as $MSE_{\sigma^2} = |\sigma^2 - \hat{\sigma}^2|^2$. The "tr, L=20" curve is the ML result from using training only over 20 symbols. The "tr, L=200" curve is the ML result from using training over 200 symbols. Thus, the performance of the proposed algorithm for both channel and noise gain is close to optimal (10dB gain) when the noise variance is small (i.e. $\sigma^2 < 0.01$). Note that, on average, EM always improves performance in both the channel gain and the noise gain estimate. As expected, the "tr" channel gain MSE is proportional to σ^2. Also, as σ^2 increases, the MSE_{σ^2} curve for EM decreases while the $MSE_{\mathbf{H}}$ curve stays the same. This is due to the algorithm's ability to detect the high noise variance region and consequently reduce the MSE in the estimate of σ^2. The EM algorithm's stopping criterion is chosen to be $\frac{\|\mathbf{H}_{k+1}-\mathbf{H}_k\|}{\|\mathbf{H}_k\|} < 0.005$ and $\frac{|\sigma_{k+1}^2-\sigma_k^2|}{\sigma_k^2} < 0.005$. The number of iterations needed for convergence is shown in Fig. 5. Few iterations (1-2) are needed in the low noise variance case. Fig. 6 shows the probability of making a symbol error when trying to decode both users using the soft information $E[\mathcal{X}|\mathbf{Y}]$ acquired from the EM algorithm.

VI. Conclusions

Through training and the EM algorithm, this paper finds accurate channel estimates as well as soft information for the receiver so that all transmitters' data can be decoded simultaneously. In the low noise variance case, the EM algorithm is also able to significantly refine the channel gain and noise estimates found from training.

VII. Acknowledgment

The authors would like to thank Elisabeth de Carvalho for her comments on this contribution.

References

[1] J. A. C. Bingham, "Multicarrier modulation for data transmission: an idea whose time has come", *IEEE Communications Magazine*, May 1990, pp. 5-14.

[2] V.K. Jones, G. Raleigh, "Channel estimation for wireless OFDM systems", *Globecom 1998*, Vol 2 , pp. 980-985.

[3] H. Cirpan, M. Tsatsanis, "Stochastic maximum likelihood methods for semi-blind channel estimation", *IEEE Signal Processing Letters*, Vol. 5, No. 1, January 1998, pp. 21-24.

[4] C. Cozzo, B. Hughes, "Joint channel estimation and data symbol detection in space-time communications", *ICC, Commun. Theory Mini-Symposium*, June 2000, pp.287-291.

[5] C. Georghiades, J. Han, "Sequence estimation in the presence of random parameters via the EM algorithm", *IEEE Trans. on Communications*, March 1997, pp.300-308.

[6] Y. Zhao, "An EM algorithm for linear distortion channel estimation based on observations from mixture of gaussian sources", *IEEE Trans. on Speech and Audio Processing*, Vol. 7, No.4, July 1999, pp. 400-413.

[7] H. Zamiri-Jafarian, S. Pasupathy, "Recursive channel estimation for wireless communication via the EM algorithm", *IEEE International Conference on Personal Wireless Communications*, 1997, pp. 33-37.

[8] G. Kaleh, and R. Vallet, "Joint parameter estimation and symbol detection for linear or nonlinear unknown channels", *IEEE Trans. on Comm.*, Vol. 42, No. 7, July 1994, pp. 2406-2413.

[9] M. Feder, E. Weinstein, "Parameter estimation of superimposed signals using the EM algorithm", *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Vol. 36, No. 4, April 1988, pp. 477-489.

[10] S. Zhou, G. Giannakis, "Finite-alphabet based channel estimation for OFDM and related multi-carrier systems", *2000 Conference on Information Sciences and Systems*, Princeton University, March 15-17, 2000.

[11] S. Talwar, M. Viberg, and A. Paulraj, "Blind separation of synchronous co-channel digital signals using an antenna array-part1: Algorithms", *IEEE Trans. on Signal Proc.*, Vol. 44, No. 5, May 1996, pp. 1184-1197.

[12] X. Wang, K.J. Liu, "Adaptive channel estimation using cyclic prefix in multicarrier modulation system", *IEEE Communications Letters*, Vol. 3, No. 10, October 1999, pp. 291-293.

[13] J. Manton, Y. Hua, "A frequency domain deterministic approach to channel identification", *IEEE Signal Processing Letters*, Vol. 6, No. 12, December 1999, pp. 323-326.

[14] G.J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*, John Wiley and Sons, 1997.

[15] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm" *J.R. Stat Soc. B*, vol. 39, 1977, pp.1-38.

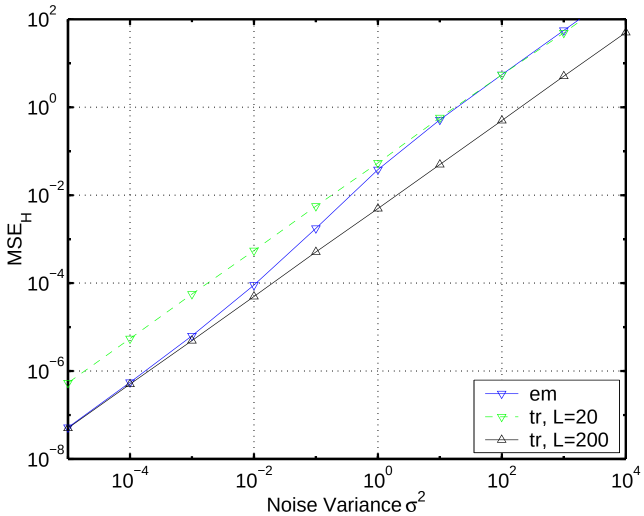

Fig. 3. Gain MSE for 2 users

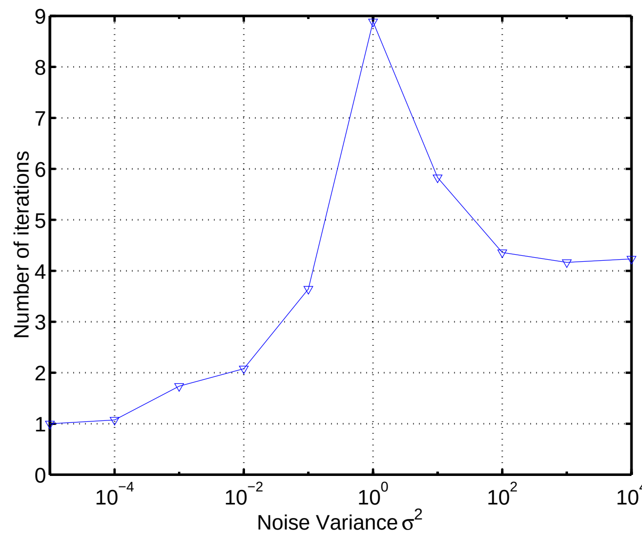

Fig. 5. Number of Iterations

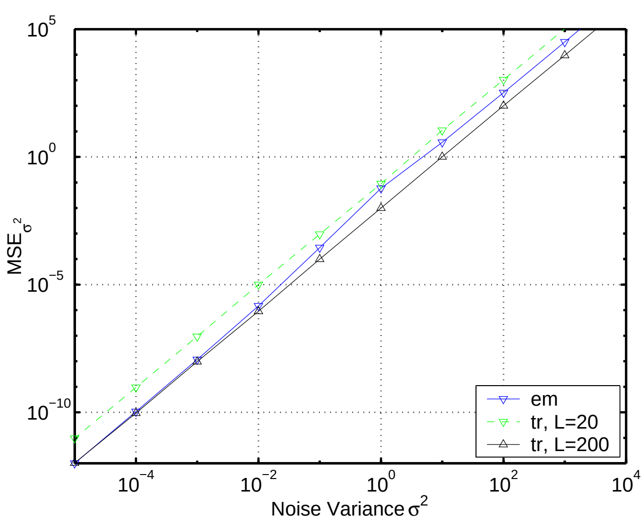

Fig. 4. Noise Variance MSE for 2 users

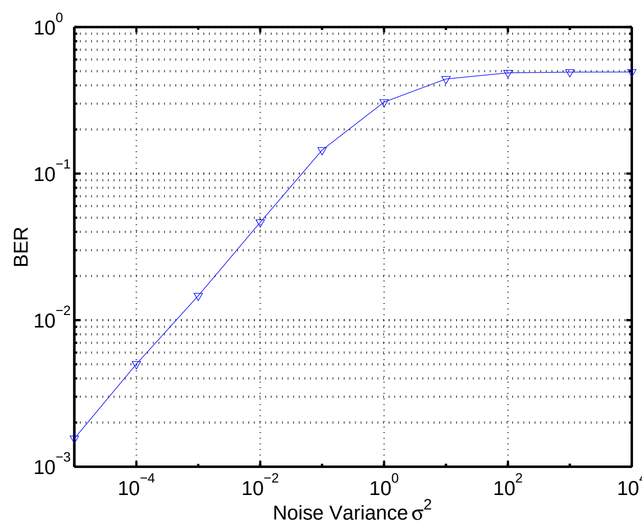

Fig. 6. Probability of bit error for 2 users