Convergence Analysis of an Adaptive Channel Estimation Algorithm using Cyclic Prefix in Multicarrier Modulation Systems

Xiaowen Wang and K. J. Ray Liu

Electrical and Computer Engineering Department, University of Maryland

College Park, MD 20742, USA

{xiaowenw, kjrliu}@eng.umd.edu

Abstract—Multicarrier Modulation (MCM) gains growing interests in high data rate communications for both wire and wireless environment. The channel estimation is a crucial aspect in MCM systems. A novel channel estimation algorithm [6] exploiting the channel information contained in the cyclic prefix outperforms the existing adaptive equalization scheme [2]. In this paper theoretical analysis is carried on the convergence of the adaptive channel estimation algorithm considering both channel noise and decision error. We prove that the algorithm is guaranteed to converge with proper loading. We also analyze several factors that affect the convergence rate such as data structure and loading function. Computer simulation shows that our analytical results are quite close to the simulation.

I. INTRODUCTION

Multicarrier modulation (MCM) is now considered an effective technique for both wire and wireless communications [1]. MCM provides an easy way to optimize the channel capacity usage by adjusting the bit rate and transmit power according to the conditions of subchannels. It has a relative longer symbol duration which produces greater immunity to impulse noise and fast fading. Because of these advantages, MCM becomes a promising approach in digital subscriber line (xDSL), digital video/audio broadcasting, and wireless communications [1].

The channel information is essential to bit and power allocations and signal detection in MCM systems. Without perfect knowledge of channel parameters, the MCM system either can not work or may incur significant performance loss. The channel estimation can be done using some training processes. If the channel is time varying, periodical training sequence, such as pilot symbols [5] have to be sent to track the variation. However, people are trying to estimate the channel more efficiently by directly estimating from the transmitted data. In MCM systems, there is a feature makes such an estimation even more practical, which is the cyclic prefix. The cyclic prefix originally designed to reduce intersymbol interference (ISI). However, in [6] and [7], it is observed that the cyclic prefix can be viewed as periodically sending training sequence. An adaptive block recursive least square (RLS) estimation algorithm is proposed based on this observation. The algorithm uses decision directed samples and hence no extra training is needed. The system with this adaptive channel estimation outperforms more robustly than the existing system with adaptive equalization [2]. In this paper, we try to analyze the convergence of the algorithm theoretically.

A lot of research has been done on the convergence of the decision directed blind equalization approaches [8][10][11][12]. The understanding of those blind algorithms is that due to the nonlinearity of the decision directed scheme, the cost function usually has more than one local minima and some kinds of smart initialization schemes must be used to force the system converges to the global minimum. However, the algorithm proposed is quite different from those blind equalization schemes. The goal of the adaptive estimation algorithm is to estimate the channel itself instead of channel inverse. The decision directed samples are treated as the filter input data and appear in both data correlation matrix and cross-correlation vector instead of just the cross-correlation vector of *Wiener-Hopf equation*. Moreover, most of work before is done under the *open eye pattern* assumption which means no decision error exists. The channel noise is also ignored in those analysis. In this paper, we are trying to consider both channel noise and decision error. The problem becomes complicated because the decision error affects the channel estimation while the channel estimation error also affects the decision error. We first separate the analysis problem into two parts and study the effect of decision error on the estimation algorithm and the effect of estimation error on signal detection. We then link the two parts together by constructing an recursive mapping of SER. The conclusions are drawn by studying the convergence of this recursive mapping.

II. MCM SYSTEM AND ADAPTIVE CHANNEL ESTIMATION ALGORITHM

A. MCM System Using Cyclic Prefix

Fig. 1 shows a MCM system using cyclic prefix with adaptive channel estimation. The system has $m/2$ complex parallel subchannels. Input data are first buffered to blocks and then divided into $m/2$ bit streams and mapped to some complex constellation points $X_{i,k}$, $i = 1, \cdots, m/2$ at time k. The modulation is implemented by m-point inverse discrete Fourier transform (IDFT), $x_{i,k} = \frac{1}{\sqrt{m}} \sum_{l=0}^{m-1} X_{l,k} e^{j\frac{2\pi li}{m}}$, where the last $m/2$ samples of $\mathbf{X}_k = [X_{0,k} \ X_{1,k} \ \cdots \ X_{m-1,k}]^T$ are just the conjugate of the first $m/2$ samples. Therefore the modulated data
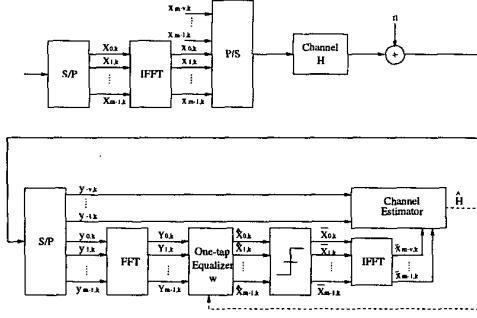
Fig. 1. The MCM System with Cyclic Prefix and Adaptive Channel Estimation

$\mathbf{x}_k = [x_{0,k} \ x_{1,k} \ \cdots \ x_{m-1,k}]^T$ are real samples. A cyclic prefix $\mathbf{x}_k^{(f)} = [x_{-v,k} \ \cdots \ x_{-1,k}]^T$ is constructed by $x_{-i,k} = x_{m-i,k}$, $i = 1, \cdots, v$, and transmitted before $\mathbf{x}_k$. At the receiver, the prefix part $\mathbf{y}_k^{(f)} = [y_{-v,k} \ \cdots \ y_{-1,k}]^T$ is discarded. The demodulation is performed only on $\mathbf{y}_k = [y_{0,k} \ y_{1,k} \ \cdots \ y_{m-1,k}]^T$ by the DFT operation. The demodulated data is $\mathbf{Y}_k = [Y_{0,k} \ Y_{1,k} \ \cdots \ Y_{m-1,k}]^T$ with $Y_{i,k} = \frac{1}{\sqrt{m}} \sum_{l=0}^{m-1} y_{l,k} e^{-j\frac{2\pi li}{m}}$.

The channel is usually modeled as a FIR filter with length $v + 1$. The impulse response of the channel is $\mathbf{h} = [h_0, h_1, \cdots, h_v]^T$. The channel noise $n_{i,k}$'s are assumed to be i.i.d. Gaussian distribution with zero mean and variance σ^2. If the cyclic prefix is longer enough, there is no ISI between $\mathbf{X}_k$'s and the subchannels can be viewed as independent with each other, i.e.,

$$Y_{i,k} = X_{i,k}H_i + N_{i,k}, \quad (1)$$

where $H_i = \sum_{l=0}^{v} h_l e^{-j\frac{2\pi li}{m}}$. and $N_{i,k}$ is the noise of the ith subchannel also with variance σ^2. $N_{i,k}$'s are also independent with each other.

For the independent subchannel of (1) only a one-tap equalizer $W_{i,k}$ is needed to get the estimation of $X_{i,k}$ from $Y_{i,k}$, i.e., $\hat{X}_{i,k} = Y_{i,k} \cdot W_{i,k}$. For the zero-forcing equalizer, $W_{i,k} = \frac{1}{H_i}$. Then the decision is made upon $\hat{X}_{i,k}$ resulting in $\bar{X}_{i,k} = q(\hat{X}_{i,k})$, where $q(\cdot)$ is some type of quantization function.

B. Adaptive Channel Estimation Algorithm

In the MCM system, usually the received cyclic prefix part $\mathbf{y}_k^{(f)}$ is discarded. However, it is found in [6] that if all the prefix parts concatenate together, it forms a training sequence. Based on this observation, a block RLS algorithm is adopted to adaptively estimate the channel. The is summarized as follows.

Input: $\mathbf{y}_k^{(f)}$ and $\mathbf{Y}_k$.

Selecting parameters: μ_1 and μ_2.

Initialization: $k = 0$, an initial training process is used to initialize $\hat{\mathbf{h}}(0)$, $\Gamma_i \triangleq \mathrm{E}[\|X_{i,k}\|^2]$ and $\bar{\Phi}(0)$.

Computation: $k = 1, 2, 3, \cdots$

1. $\hat{H}_{i,k-1} = \sum_{l=0}^{v} \hat{h}_{l,k-1} e^{-j2\pi il/m}$,
   $\hat{W}_{i,k-1} = \frac{1}{\hat{H}_{i,k-1}}$.

2. $\hat{X}_{i,k} = Y_{i,k}\hat{W}_{i,k-1}$, $i = 0, \cdots m-1$

3. $\bar{x}_{i,k} = \frac{1}{\sqrt{m}} \sum_{l=0}^{m-1} q(\hat{X}_{l,k}) e^{j\frac{2\pi il}{m}}$, $i = m - v, \cdots, m-1$

4. $\bar{\Phi}(k) = \mu_1 \bar{\Phi}(k-1) + \sum_{l=1}^{v} \mu_2 \bar{\mathbf{u}}_l(k)\bar{\mathbf{u}}_l^H(k)$,
   $\bar{\mathbf{z}}(k) = \mu_1 \bar{\mathbf{z}}(k-1) + \sum_{l=1}^{v} \mu_2 \bar{\mathbf{u}}_l(k)y_{-l,k}^*$,

5. $\hat{\mathbf{h}}(k) = \bar{\Phi}^{-1}(k)\bar{\mathbf{z}}(k)$

In above algorithm, a loading algorithm is performed to allocate bit and energy at the initialization based on the the information got by initial training. The loading is not only try to optimize the system performance, but also try to limit the decision error to such a level that the system behaves near the equilibrium.

The loading is done to meet some requirements on SER with energy constraint $E_x \triangleq \mathrm{E}[\|x_{i,k}^2\|^2] = \frac{1}{m}\sum_{i\in U}\Gamma_i \leq C$, where U is the set of all the used subchannels. For the QAM constellation used in our systems, the SER requirement is $P_{e,i}(0) = P_e(0) = 4Q\left(\frac{d_i\|H_{i,0}\|}{\sqrt{2}\sigma}\right)$, where $P_{e,i}(0)$ is the initial SER and d_i is the minimum distance between constellation points of the ith subchannel. $P_e(0)$ is some preset required value of SER which will guarantee the system performance and $Q(x) = \int_x^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$.

The transmitted energy of ith subchannel is $\Gamma_i = \frac{1}{m}\sum_{i=0}^{m-1} \frac{M_i-1}{6} d_i^2$, with M_i as the number of constellation points. After calculating d_i form the SER requirement, using the energy constraint we can find M_i for each subchannel. If $M_i = 0$, then the subchannel is unused.

Fig. 2 compares above adaptive channel estimation algorithm with the adaptive equalization scheme in [2]. In such a scheme, the one tap equalizer updated independently for each subchannels. The equalizer coefficient is updated by $W_{i,k} = W_{i,k-1} + \mu(\hat{X}_{i,k} - \bar{X}_{i,k})Y_{i,k}^*$.

The MCM system used in the simulation has 256 complex subchannels. The average transmit energy is 1. initially the channel transfer function used is $H_0(D) = \frac{0.9923}{1-0.12D}$ and the loading is done according to it. After some time, the channel changes to $H(D) = \frac{0.9527}{1-0.3D}$. The figure shows the *averaged* SER which is defined as

$$\tilde{p}(k) \triangleq \frac{1}{|U|} \sum_{i\in U} P_{e,i}(k), \quad (2)$$

where $P_{e,i}(k)$ is SER of ith subchannel of kth block.

In [8][10], the convergence of such an adaptive equalization scheme is proved under the condition that the channel noise is small and no decision error exists. However,
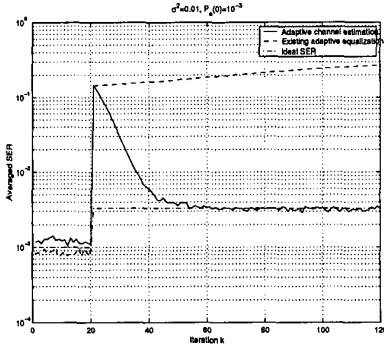
Fig. 2. Average SER ($\mu_1 = 0.9$, $\mu_2 = 1$)

in Fig. 2, our simulation shows that such a system fails to follow the channel variation while our adaptive channel estimation works.

III. CONVERGENCE ANALYSIS

In this adaptive algorithm, if ideal samples of the cyclic prefix is available , then we know from many literatures [13] that the algorithm converges to an unbiased estimation linearly and the convergence rate is determined by the eigenvalue spread of the data correlation matrix $\mathbf{R} = E[\mathbf{u}_l\mathbf{u}_l^H]$. Unfortunately, we can only use the decision directed data in practice. The decision error will affect the channel estimation. However, if the channel variation is within some ranges, then the decision error can be controlled to some small level by doing proper loading. Under this condition, we can begin our analysis.

First define *frequency domain decision error* as $\eta_{i,k} \overset{\Delta}{=} \bar{X}_{i,k} - X_{i,k}$ where $\eta_{i,k}$'s are independent with different i's and k's with the energy $\sigma^2_{\eta_i}(k)$. Define *time domain decision error* as $e_{i,k} \overset{\Delta}{=} \bar{x}_{i,k} - x_{i,k} = \frac{1}{\sqrt{m}}\sum_{l\in U}\eta_{l,k}e^{-j\frac{2\pi li}{m}}, i = -v,\cdots,-1$. The estimated data vector can be written as $\bar{\mathbf{u}}_l(k) = \mathbf{u}_l(k) + \mathbf{e}_l(k)$, where $\mathbf{u}_l(k) = [x_{-l,k}\cdots x_{-1,k}\ x_{-v,k-1}\cdots x_{-l,k-1}]^T$ and $\mathbf{e}_l(k) = [e_{-l,k}\cdots e_{-1,k}\ e_{-v,k-1}\cdots e_{-l,k-1}]^T$. The decision error can be determined by SER with given constellation.

The *time domain estimation error* is defined as $\varepsilon_{i,k} = \hat{h}_{i,k} - h_i$ and $\varepsilon(k) = [\varepsilon_{0,k}\cdots\varepsilon_{v,k}]^T$. And the *frequency domain estimation error* is defined as $\Delta H_{i,k} = \hat{H}_{i,k} - H_i$. The energy of the frequency domain estimation error is $\sigma^2_{H_i}(k) \overset{\Delta}{=} E[\|\Delta H_{i,k}\|^2]$.

In order to analyze the behavior of the channel estimation, we make following assumptions about the input data and decision error.

1. The input data $x_{i,k}$ is independent with the chan-

nel noise $n_{i,k}$.and the decision error $e_{i,k}$, which is true if the constellation used is large enough. We also observed that the decision error $e_{i,k}$ in the estimated cyclic prefix is independent with $n_{i,k}$, $i = -v, \cdots, -1$.

2. The input data vector $\mathbf{u}_l(k)$ is generated from the $(v+1)$-dimensional Gaussian distribution $\mathcal{N}(\mathbf{0}, \mathbf{R})$ independently. Then it is proved that $E[\Phi^{-1}(k)] = \frac{1}{kv-v-2}\mathbf{R}^{-1}$.

3. The statistics of the decision error varies slowly so that $\mathbf{R}_e(k) = E[\mathbf{e}_l(k)\mathbf{e}_l^H(k)]$.

A. Effect of decision error

In the following analysis, we only consider $\mu_1 = \mu_2 = 1$. We assume that $\mathbf{e}_l(k)$ is stationary, i.e., $\mathbf{R}_e(k) = \mathbf{R}_e$.

Theorem 1: If the previous three assumptions are valid around the equilibrium of the system, the optimal channel estimation in term of minimizing the cost function of RLS algorithm is a biased estimation with stationary decision error. The channel estimation in this case is

$$\hat{\mathbf{h}} = [\mathbf{I} - \mathbf{R}^{-1}\mathbf{R}_e^{\frac{1}{2}}(\mathbf{I} + \mathbf{R}_e^{\frac{1}{2}}\mathbf{R}^{-1}\mathbf{R}_e^{\frac{1}{2}})^{-1}\mathbf{R}_e^{\frac{1}{2}}]\mathbf{h}. \quad (3)$$

If there is no decision error in the data, it is well known that the solution of RLS is unbiased. However, the theorem states that with a stationary decision error, the channel estimation becomes biased. If we further assume that all eigenvalues of $\mathbf{R}^{-1}\mathbf{R}_e$ are much less than 1, this bias can be simplified to

$$\hat{\mathbf{h}} - \mathbf{h} = \mathbf{R}^{-1}\mathbf{R}_e\mathbf{h}. \quad (4)$$

The dynamic of such a convergence can be analyzed when the decision error is near stationary, i.e., $\lim_{k\to\infty}\mathbf{R}_e(k) = \mathbf{R}_e$ and all eigenvalues of $\Phi(k)^{-1}\Delta\Phi(k)$ are all much less than 1, then, the estimation error can be approximated as $\varepsilon(k) = \hat{\mathbf{h}}(k) - \mathbf{h} \approx \Phi^{-1}(k)\epsilon(k)$, where $\epsilon(k) = \sum_{n=1}^{k}\sum_{l=1}^{v}\mathbf{u}_l(n)n^*_{-l,n} + \mathbf{e}_l(n)n^*_{-l,n} - \mathbf{u}_l(n)\mathbf{e}_l^H(n)\mathbf{h} - \mathbf{e}_l(n)\mathbf{e}_l^H(n)\mathbf{h}$. We then can show that

$$E[\varepsilon(k)] = \frac{kv}{kv-v-2}\mathbf{R}^{-1}\mathbf{R}_e(k)\mathbf{h}, \quad (5)$$

As $k \to \infty$, this mean value goes to the stationary point which is given in (4).

B. Effects of the estimation error

Define the *residual noise at the decision point* as $\xi_{i,k} \overset{\Delta}{=} \hat{X}_{i,k} - X_{i,k}$. From section II-A, $\hat{X}_{i,k}$ is the output of the equalizer. The coefficient of the equalizer is $\hat{W}_{i,k} = \frac{1}{\hat{H}_{i,k}}$. Using the first order approximation, it comes $\hat{W}_{i,k} = \frac{1}{\hat{H}_i} \approx W_{i,k}(1+\frac{\Delta H_{i,k}}{H_{i,k}})$. Then we can derive the following theorem.

718

Theorem 2: If we assume that the decision error $e_{i,k}$ are Gaussian distributed, with the first order approximation, the estimation error propagates to the decision point as an extra additive Gaussian noise independent with the transmitted signal conditioned on the knowledge of the ideal data, i.e, $E[\xi_{i,k}X_{i,k}|X_{i,k}] = 0$. and $E[||\xi_{i,k}||^2|X_{i,k}] = \frac{\sigma^2}{||H_i||^2} + \frac{||X_{i,k}||^2 \mathbf{w}_i^H \mathbf{R}^{-1}\mathbf{w}_i[\sigma^2+\mathbf{h}^H\mathbf{R}_e(k-1)\mathbf{h}]}{(kv-2v-2)||H_i||^2}$, where $\mathbf{w}_i = [1 \; e^{-j\frac{2i\pi}{m}} \; \cdots \; e^{-j\frac{2i\pi}{m}v}]^T$.

We then can use those familiar formula derived for AWGN channel to calculate the SER.

C. Recursive mapping of SER and convergence

Approximates the frequency domain decision error as $\sigma_{\eta_l}^2(k) = P_{e,l}(k)d_l^2$. Then the time domain estimation error can be calculated as $\mathbf{h}^H\mathbf{R}_e(k)\mathbf{h} = \frac{1}{m}\sum_{l\in U}\sigma_{\eta_l}^2(k)||H_l||^2$.

Using the formula in [14], the SER of ith subchannel of k block is

$$P_{e,i}(k) = E[4q_{i,k}(1-q_{i,k})], \qquad (6)$$

where $q_{i,k} = \left(1-\frac{1}{\sqrt{M_i}}\right)Q(\lambda)$ with $\lambda =$

$$\sqrt{\frac{d_iH_i}{\frac{\mathbf{w}_i^H\mathbf{R}^{-1}\mathbf{w}_i||X_{i,k}||^2(\sigma^2+\frac{2\sigma^2Q-2}{m}\left(\frac{P_e(0)}{4}\right)\sum_{l\in U}\frac{||H_l||^2}{||H_{l,0}||^2}P_{e,l}(k-1))}{kv-2v-2}+\sigma^2}}$$

Then define $\bar{p}(k) \overset{\triangle}{=} \frac{1}{m}\sum_{i\in U}\frac{||H_i||^2}{||H_{i,0}||^2}P_{e,i}(k)$. The iteration of $\bar{p}(k)$ is

$$\bar{p}(k) = f(\bar{p}(k-1)) = \frac{4}{m}\sum_{i\in U}\frac{||H_i||^2}{||H_{i,0}||^2}$$
$$E[q_{i,k}(\bar{p}(k-1))(1-q_{i,k}(\bar{p}(k-1)))]. \quad (7)$$

The condition for this iteration to converge is $|f'(\bar{p})| < 1$. By some mathematics, we find out $|f'(\bar{p})|$ is bounded by

$$|f'(\bar{p})| \leq \frac{1}{kv-2v-2}\frac{36\sqrt{3}e^{\frac{3}{2}}}{m\sqrt{2\pi}}\sum_{i\in U}\mathbf{w}_i^H\mathbf{R}^{-1}\mathbf{w}_i\Gamma_i. \quad (8)$$

The boundness implies that $\lim_{k\to\infty}|f'(\bar{p})| < 1$, which leads to the following theorem,

Theorem 3: The adaptive channel estimation algorithm is guaranteed to converge linearly just as the usual RLS algorithm.

Further analysis of the bound of $|f'(\bar{p})|$, we found that the algorithm would converge faster if $\mathbf{w}_i^H\mathbf{R}^{-1}\mathbf{w}_i$ is smaller. $\mathbf{w}_i^H\mathbf{R}^{-1}\mathbf{w}_i = \text{tr}[\mathbf{R}^{-1}]$ if $\mathbf{R}$ is diagonal matrix. Then it comes to the familiar conclusion that the convergence rate is decided by the eigenvalue spread of the auto-correlation matrix of data. The well-conditioned data ensures a fast convergence rate. If all the subchannels are used and have same energy, then the condition number of the data auto-correlation matrix is 1. However, loading algorithm may make some channel with small channel gain unused, resulting in a bad-conditioned data auto-correlation matrix and slow convergence.

It is easy to show that $\bar{p}(k)$ should be smaller to accelerate the convergence. Then, from (7), we have $\bar{p}(k) \geq g(H_i, H_{i,0}, q_{i,k})$ where g is some function. The equality is valid if and only if $||H_i|| = A||H_{i,0}||$, where A is a constant. This means that the initial channel response which we use for loading should have the same shape as the ideal channel response to guarantee a fast convergence. Furthermore, A is the larger the better, i.e., the ideal channel response would be better have higher gain than the initial channel response for loading, or it would be better to leave more gap in loading to ensure the convergence of the channel estimation algorithm.

Unlike the existing equalization scheme which treated each subchannel independently, the SER iteration for any individual subchannel depends on the performance of the whole system $\bar{p}(k-1)$ as shown in (6). Our algorithm can recover even though the performances of some subchannels if the overall system still performs well.

IV. Computer Simulation

Since we use some approximation in above analysis, computer simulation is done to verify the analysis results. The same system used in section II-B is also used in this section.

We first show how a stationary decision error affects the channel estimation. Ideal channel information is used for equalization, i.e., no estimation error propagates to decision point. The transfer function used is $H(D) = \frac{0.02+0.16D^2}{1-1.5D+0.54D^2}$. Fig. 3 shows the learning curves with and without decision error. The learning curve evaluates the averaged residual noise which is $\frac{1}{m}\sum_{i\in U}\sigma_{\xi_i}^2$. It can be seen that there is a constant difference between the two curves which corresponds to the bias of the channel estimation caused by the decision error. However, the bias is quite small only about 1dB with the SER requirement of 10^{-2}.

Fig 4 verifies the convergence of the recursive mapping with a flat channel whose H_i is constant. Since the channel is flat, the transmit rate and energy for all the subchannels are the same. Fig. 4 (a) shows the analytical and the simulated results of the iteration of the *averaged SER* defined in (2). The average channel estimation error $\frac{1}{m}\sum_l\sigma_{\eta_l}^2$ and its standard deviation is shown in Fig. 4 (b). The analytical results and the simulation results are very close in both figures. There are two scenarios with different initial SER requirement 10^{-2} and 10^{-3} are shown in Fig. 4 (a). In both scenarios the simulation results are very close to the analytical results. When 10^{-3}, results for different noise levels are plotted. We found that the two iteration curves are very close to each other. This is because the loading does not change and so does correlation of the time domain data due to the flat channel response in both cases.
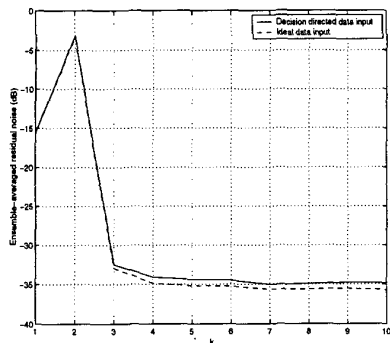
Fig. 3. Learning curves ($v = 64$, $\sigma^2 = 0.1$ and $P_e(0) = 10^{-2}$)

Fig. 5 shows the effect of initial loading. The transfer function used here is the one used in Fig. 2. Two cases are simulated. In one case, the loading is done according to the ideal channel response. In the other case, we assume no knowledge of the channel available, so flat loading is used. Once again, the simulation results verify the analytical result. It also shows that the system converges faster when loading is done according to the ideal channel information. It converges in about 10 iterations when ideal loading is done while in about 20 iterations with flat loading.

V. CONCLUSION

In this paper, we investigates the convergence analysis problem of the adaptive channel estimation algorithm proposed in [6]. Our studies show that existence of decision error results in a biased channel estimation and the channel estimation error appears at decision point as an additional noise. Based on these conclusions, we derive a recursive expression of SER. Considering the convergence of this iteration, we prove that the system is guaranteed to converge as the iteration goes on. The convergence rate is determined by the eigenvalues of the data correlation matrix which is affected by the channel noise and loading algorithm.

REFERENCES

[1] J. A. C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come", *IEEE Communications Magazine*, pp5-14, May, 1990.

[2] J. S. Chow, J. C. Tu and J. M. Cioffi, "A discrete multitone transceiver system for HDSL application", *IEEE Journal on Selected Areas in Communications*, v. 9, N. 6, pp895-908, August, 1991.

[3] P. S. Chow, J. M. Cioffi and J. A. C. Bingham, "A practical discrete multiton transceiver loading algorithm for data transmission over spectrally shaped channels", *IEEE Transaction on Communications*, pp773-775, Feb./March/April, 1995.

[4] J. K. Cavers, "An analysis of pilot symbol assisted modulation for rayleigh fading channels", *IEEE Transactions on Vehicular Technology*, v. 40, N. 4, pp686-693, Nov. 1991.
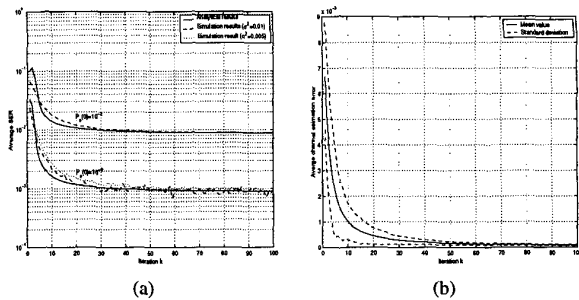
(a)   (b)

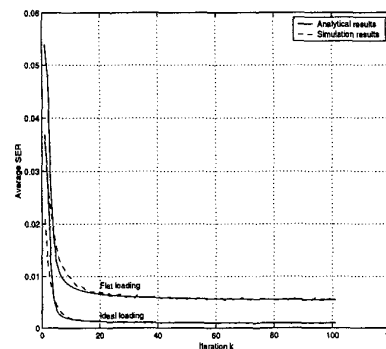Fig. 4.   (a) SER iteration (b) Channel estimation error ($v = 32$ and $P_e(0) = 10^{-3}$)

Fig. 5.   SER iteration ($v = 32$, $\sigma^2 = 0.01$ and $P_e(0) = 10^{-3}$)

[5] P. Hoeher, S. Kaiser and P. Robertson, "Two-dimensional pilot-symbol-aided channel estimation by Wiener filtering", *Proc. of 1997 IEEE International Conf. on Acoustics, Speech and Signal Processing*, Munich, Germany, pp1845-1848, April, 1997.

[6] X. Wang and K. J. R. Liu, "Adaptive channel estimation in multicarrier modulation system using cyclic prefix", *IEEE Communication Letter*, v. , N. pp, Oct., 1999.

[7] X. Wang and K. J. R. Liu, "Joint channel estimation and equalization in multicarrier modulation system using cyclic prefix", *Proc. of ICASSP'99*, v. 5, pp2733-2736, Phoenix, AZ.

[8] O. Macchi and E. Eweda, "Convergence analysis of self-adaptive equalizers", *IEEE Transaction on Information Theory*, v. 30, N. 2, pp161-176, March, 1984.

[9] Z. Ding, C. R. Johnson, Jr. and R. A. Kennedy, "On the (non)existence of undesirable equilibria of Godard blind equalizer", *IEEE Transaction on Signal Processing*, v. 40, N. 10, pp2425-2432, Oct., 1992.

[10] Y. Li and Z. Ding, "Convergence analysis of finite length blind adaptive equalizers", *IEEE Transaction on Signal Processing*, v. 43, N. 9, pp2120-2129, Sep., 1995.

[11] Y. Li and K. J. R. Liu, "Static and dynamic convergence behavior of adaptive blind equalizers", *IEEE Transaction on Signal Processing*, v. 44, N. 11, pp2736-2744, Nov., 1996.

[12] D. Kundur and D. Hatzinakos, "On the use of Lyapunov criteria to analyze the convergence of blind deconvolution algorithm" *IEEE Transaction on Signal Processing*, v. 46, N. 11, pp2918-2925, Nov., 1998.

[13] S. Haykin, *Adaptive Filter Theory*, NJ: Prentice Hall, 1996.

[14] J. G. Proakis, *Digital Communications*, 2nd ed., New York: McGraw-Hill, 1989.