# JOINT CHANNEL ESTIMATION AND EQUALIZATION IN MULTICARRIER MODULATION SYSTEM USING CYCLIC PREFIX

*Xiaowen Wang and K. J. Ray Liu*

Electrical Engineering Department, University of Maryland
College Park, MD 20742, USA
Email: kjrliu@eng.umd.edu

## ABSTRACT

Multicarrier modulation (MCM) is a promising technique for high rate data transmission. A one-tap equalizer is an essential part of the MCM system and the channel estimation is needed to get the coefficient of the equalizer. Lack of correct channel estimation may cause significant performance degradation. We propose to use the cyclic prefix to estimate the channel for the MCM system. We found that the cyclic prefix originally used solely to guarantee the optimality of modulation using discrete Fourier transform (DFT) can be viewed as a source of channel information. Based on this observation, we propose a joint channel estimation and equalization algorithm using the cyclic prefix. Our simulations show that the algorithm can adaptively track the variation of a moderately time varying channel and has about 1-2dB gain over the system using the channel estimation obtained by the conventional training schemes.

## 1. INTRODUCTION

Multicarrier modulation (MCM) is now considered an effective technique for high rate data communications in both wire and wireless environments. The principle of MCM is dividing the transmit data into several parallel low bit rate data streams, and using these data streams to modulate several carriers, which in frequency domain is equivalent to partition the entire channel into several parallel subchannels.

MCM provides an optimal way for channel capacity usage by adjusting the bit rate and transmit power according to the SNR of subchannels. MCM also has a relative longer symbol duration since it is a block oriented technique. The long symbol duration produces greater immunity to impulse noise and fast fading. Because of these advantages, MCM is considered a promising approach in digital subscriber line (xDSL), digital video/audio broadcasting, and wireless communications.

In MCM system, usually a one-tap equalizer is needed for each subchannel to get the estimations of transmitted data. The channel information is essential to the coefficients of the equalizers. Some techniques, such as differential PSK modulation, are used to eliminate the need for channel estimation and equalization. However, differential demodulation causes 3-4dB Signal to Noise Ratio(SNR) loss compared with coherent demodulation if channel information is known. Moreover, channel information is also very important for the bit and power allocation.

In applications such as xDSL, some training processes are performed to estimate the channel before the communication is set up. Then, this channel estimate is used through the entire communication. If the channel changes, retraining is required to track the variation. Recently some research has been done on channel estimation and tracking in wireless communications. A minimum mean square error estimation algorithm is proposed in [5].

In this paper we propose a new channel estimation scheme that can track the change of the channel parameters without retraining. Usually a cyclic prefix or a guarding period is added between two symbols in MCM system in order to reduce the intersymbol interference (ISI). We propose to use the cyclic prefix, which is normally discarded, for channel estimation and equalization. We observed that the prefix actually provides a constantly sent training sequence if accurate transmit signal can be recovered by the conventional MCM systems. A joint channel estimation and equalization algorithm using the cyclic prefix is proposed based on this observation. Simulations were performed under the asymmetric digital subscriber line(ADSL) environment to show the effectiveness of the algorithm.

## 2. MCM SYSTEM USING CYCLIC PREFIX

MCM partitions a spectrally shaped channel into a number of parallel and subchannels by modulating a set of orthonormal basis functions. Most of the MCM systems choose the inverse discrete Fourier transform (IDFT) as the orthonormal basis. Fig. 1 shows a MCM system using IDFT as modulation scheme.

Input data are first buffered into blocks which are used to form the symbols transmitted in channel. Each block of data is then divided into $m/2$ bit streams in a manner determined during system initialization and mapped to some complex subsymbols to form the input of a $m$-point IDFT which is represented as $\mathbf{X}_k = [X_{0,k}\ X_{1,k}\ \cdots\ X_{m-1,k}]^T$, where $X_{i,k}$ is the $i$th input of IDFT. The modulation is then performed by $m$-point IDFT and the result is $\mathbf{x}_k = [x_{0,k}\ x_{1,k}\ \cdots\ x_{m-1,k}]^T$.

The channel is usually modeled as a FIR filter with length $v+1$. The impulse response of the channel is $\mathbf{h} = [h_0, h_1, \cdots, h_v]^T$. To reduce the ISI caused by the channel memory, a cyclic prefix $\mathbf{x}_k^{(f)} = [x_{-v,k}\ \cdots\ x_{-1,k}]^T$, which consists the last $v$ samples of $\mathbf{x}_k$, i.e., $x_{-i,k} = x_{m-i,k},\ i = 1, \cdots, v$, is appended in front of $\mathbf{x}_k$ before transmission.

At the receiver, the prefix part $\mathbf{y}_k^{(f)} = [y_{-v,k}\ \cdots\ y_{-1,k}]^T$ is discarded, only $\mathbf{y}_k = [y_{0,k}\ y_{1,k}\ \cdots\ y_{m-1,k}]^T$ is used for demodulation. The demodulation is performed by the DFT operation and the result is $\mathbf{Y}_k = [Y_{0,k}\ Y_{1,k}\ \cdots\ Y_{m-1,k}]^T$.

It can be proved that the above modulation scheme is optimal due to the use of cyclic prefix in the sense that the mutual infor-

0-7803-5041-3/99 $10.00 © 1999 IEEE 

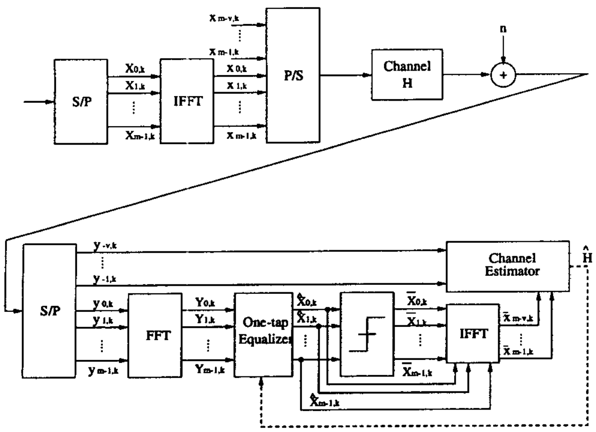



Figure 1: MCM System with Cyclic Prefix and Adaptive Channel Estimation

mation is maximized. As $m$ goes large, the subchannels can be viewed as independent with each other, i.e.,

$$Y_{i,k} = X_{i,k}\mathcal{H}_i + N_{i,k}$$

$$\mathcal{H}_i = \sum_{l=0}^{v} h_l e^{-j2\pi il/m} \tag{1}$$

are samples of $m$ point DFT of $\mathbf{h}$. $N_{i,k}$ are samples of $m$ point DFT of the channel noise.

Assuming $N_{i,k}$ are independent with each other, the best estimation of $X_{i,k}$ from $Y_{i,k}$ is achieved by applying a one-tap equalizer $w_i$ to $Y_{i,k}$, i.e.,

$$\hat{X}_{i,k} = Y_{i,k} \cdot w_i. \tag{2}$$

The optimal coefficient for the one-tap equalizer is:

$$w_i = \frac{\Gamma_i^{\frac{1}{2}} \mathcal{H}_i^*}{\Gamma_i \|\mathcal{H}_i\|^2 + \Delta_i} \tag{3}$$

where $\Gamma_i$ is the transmitted power of $X_{i,k}$ and $\Delta_i = \mathrm{E}[\|N_{i,k}\|^2]$.

Then, $\bar{X}_{i,k}$ is the hard decision result of $\hat{X}_{i,k}$, i.e., $\bar{X}_{i,k} = q(\hat{X}_{i,k})$, where $q(\cdot)$ is some kind of quantization function.

## 3. THE EXISTING TRAINING METHOD

The channel is modeled as the FIR filter stated before. When a training sequence $x_k$ is sent to a channel, the output of the channel is:

$$y_k = x_k * h_k + n_k.$$

where $n_k$ is uncorrelated random noise. Suppose $\hat{h}_k$ is the estimation of $h_k$, then the estimated output $\hat{y}_k = x_k * \hat{h}_k$. The best $\hat{h}_k$ is chosen to minimize the power of the error $\epsilon_k = y_k - \hat{y}_k$. This is the familiar quadratic form of minimizing the mean square error problem. There are many well-known methods to solve this problem, such as least squares (LS) method which is also used in our algorithm.

For MCM systems, we need to estimate the channel parameters in frequency domain. Instead of the time domain algorithm, an equivalent frequency domain deterministic least squares (DSL) channel identification algorithm can be used. In this algorithm, a training block with length $m$ is sent periodically, and then the channel outputs are collected and averaged to reduce the influence of channel noise. The DFT of channel response is obtained by performing element by element division between the DFT of the averaged channel output and the input training sequence. Several different training blocks with guarding band can be used in order to further average out any non-linear effects. The final estimation is obtained by averaging the results of all these training blocks.

## 4. THE PROPOSED JOINT CHANNEL ESTIMATION AND EQUALIZATION ALGORITHM

### 4.1. Observation on Cyclic Prefix

The training algorithms above is designed for the time-invariant system, which means new training process must be performed if the channel varies. However, in the MCM system using cyclic prefix, we can view the cyclic prefix as a training sequence and use it to track the variation of the channel.

Let's consider the prefix part $\mathbf{y}_k^{(f)}$ which is originally discarded. The relationship between $\mathbf{y}_k^{(f)}$ and the transmit signal is

$$\mathbf{y}_k^{(f)} = \mathbf{A}_k \mathbf{h} + \mathbf{n}_k^{(f)} \tag{4}$$

where

$$\mathbf{A}_k = \begin{bmatrix} x_{-v,k} & x_{m-1,k-1} & \cdots & x_{m-v,k-1} \\ \ddots & \ddots & \ddots & \vdots \\ & \ddots & \ddots & \\ x_{-1,k} & \cdots & x_{-v,k} & x_{m-1,k-1} \end{bmatrix}$$

and $\mathbf{n}_k^{(f)} = [n_{-v,k} \cdots n_{-1,k}]^T$.

The lower triangle part of matrix $\mathbf{A}_k$ is composed by $\mathbf{x}_k^{(f)}$, while the upper triangle part is composed by the last $v-1$ samples of $\mathbf{x}_k$. However, this last $v-1$ samples are also the elements of the prefix $\mathbf{x}_{k-1}^{(f)}$. So if all the prefix parts concatenate together as a pair of sequences $x^{(f)} = \{\cdots x_{-v,k-1} \cdots x_{-1,k-1} x_{-v,k} \cdots x_{-1,k} \cdots\}$ and $y^{(f)} = \{\cdots y_{-v,k-1} \cdots y_{-1,k-1}\ y_{-v,k} \cdots y_{-1,k} \cdots\}$, the relationship between these two satisfies

$$y_k^{(f)} = x_k^{(f)} * h_k + n_k. \tag{5}$$

If we can get accurate estimations of transmitted prefix by the conventional MCM method, i.e., we know $x^{(f)}$, then (5) shows that $y^{(f)}$ and $x^{(f)}$ form a pair of training sequences that can be used to estimate the channel.

One problem here is that we can only get the estimate of this training sequence. This estimate forms a feedback loop for the channel estimation. The error incurred by the inaccurate estimation may propagate. In order to reduce the probability of error propagation the samples after hard decision, $\bar{\mathbf{X}}_k$, are used to estimate the transmit prefix. Since the prefix is a time domain signal while $\bar{\mathbf{X}}_k$ are in frequency domain, an IDFT is performed to get the time domain estimation of $\bar{\mathbf{x}}_k$.

### 4.2. Least Square Method to Estimate $\hat{\mathbf{h}}$

Several methods have been tried to solve $\hat{\mathbf{h}}$ from the training sequence formed by the cyclic prefix. The following method has the best performance in simulation.

The idea of this method is trying to use LS method directly to solve (4). However, it is observed that $\mathbf{A}_k$ is an under-determined matrix. In order to reduce the effect of random noise, we expand (4) to form the following equation:

$$\begin{bmatrix} \mathbf{y}_{k-N}^{(f)} \\ \vdots \\ \mathbf{y}_{k}^{(f)} \\ \vdots \\ \mathbf{y}_{k+L}^{(f)} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{k-N} \\ \vdots \\ \mathbf{A}_{k} \\ \vdots \\ \mathbf{A}_{k+L} \end{bmatrix} \mathbf{h} + \begin{bmatrix} \mathbf{n}_{k-N}^{(f)} \\ \vdots \\ \mathbf{n}_{k}^{(f)} \\ \vdots \\ \mathbf{n}_{k+L}^{(f)} \end{bmatrix}. \tag{6}$$

After arranging data to the above form, the LS solution is given by

$$\hat{\mathbf{h}} = \mathbf{A}_{N,L}^{\dagger}(k)\mathbf{y}_{N,L}(k)$$

where $\mathbf{A}_{N,L}(k) = [\mathbf{A}_{k-N} \cdots \mathbf{A}_k \cdots \mathbf{A}_{k+L}]^T$ and $\mathbf{y}_{N,L}(k) = \left[\mathbf{y}_{k-N}^{(f)} \cdots \mathbf{y}_{k}^{(f)} \cdots \mathbf{y}_{k+L}^{(f)}\right]^T$. $\mathbf{A}_{N,L}^{\dagger}(k)$ is the pseudo inverse of $\mathbf{A}_{N,L}(k)$ that can be obtained by performing singular value decomposition (SVD) on $\mathbf{A}_{N,L}(k)$, which is an $(N+L+1)v \times (v+1)$ matrix. Usually $N, L > 0$, so the rank of $\mathbf{A}_{N,L}(k)$ is constrained by the length of the prefix, i.e., $\text{Rank}(\mathbf{A}_{N,L}(k)) \leq v+1$.

The roles of $N$ is some kind of similar to forgetting factor. $L$ is used to guarantee the amount of data is enough to get an accurate estimation.

### 4.3. Joint Channel Estimation and Equalization Algorithm

Based on the discussion in section 4.1 and 4.2, we summarize the channel estimation and equalization algorithm as the following:

*Input:* received prefix part $\mathbf{y}_k^{(f)}$ and demodulated signal $\mathbf{Y}_k$.

*Known parameters:* transmitted power $\Gamma_i$ and noise power $\Delta_i$.

*Selecting parameters:* $N$ and $L$.

*Initialization:* $k = 0$, an initial training is used to get the estimation of $\hat{\mathbf{h}}(0)$.

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

$$\begin{aligned} \mathcal{H}_i(k-1) &= \sum_{l=0}^{v} \hat{h}_l(k-1)e^{-j2\pi il/m} \\ w_i(k-1) &= \frac{\Gamma_i^{\frac{1}{2}}\mathcal{H}_i(k-1)^*}{\Gamma_i\|\mathcal{H}_i(k-1)\|^2 + \Delta_i} \\ \hat{X}_{i,k} &= Y_{i,k}w_i(k-1) \\ \bar{X}_{i,k} &= q(\hat{X}_{i,k}),\ i = 0, 1, \cdots, m-1 \\ \bar{x}_{i,k} &= \sum_{l=0}^{m-1} \bar{X}_{l,k}e^{j2\pi il/m},\ i = m-v, \cdots, m-1 \end{aligned}$$

If $k = nL$, where $n$ is an integer, use $\bar{x}_{i,k}$ calculated above to form the matrix $\mathbf{A}_{N,L}(k-L)$, then,

$$\hat{\mathbf{h}}(k) = \mathbf{A}_{N,L}^{\dagger}(k-L)\mathbf{y}_{N,L}(k-L);$$

otherwise, $\hat{\mathbf{h}}(k) = \hat{\mathbf{h}}(k-1)$.

Here, what we present is actually a block recursive algorithm and the channel estimation is refreshed every $L$ symbols. The symbol by symbol recursion is just the special case as $L = 1$. The reason for such a scheme is that this algorithm is a feedback scheme which combines channel estimation and equalization together. It requires more most recently data to keep on with the channel variation. Our simulation shows that both $N$ and $L$ should be chosen carefully to get the best performance, usually $L > 1$.

## 5. SIMULATION RESULTS

In our simulation, the transmit power of all the used subchannels is set to equal and fixed to 1. QAM signal is used in each subchannel. At first, some target error probability $P_e$ is preset. Then the bit is allocated by the following error probability constraint

$$P_e \leq 4Q(\frac{d_i\|\mathcal{H}_i\|}{\sqrt{\Delta_i}})$$

where $d_i$ is the minimum distance between the signal points in QAM constellation of the $i$th subchannel.

Initially the channel transfer function is

$$H0(D) = \frac{0.1 + 0.8D^2}{1 - 1.5D + 0.54D^2}$$

The bit allocation is done according to this transfer function and will keep unchanged during the simulation. After some time, the channel transfer function will change to some $H(D)$. Two different transfer function are used for $H(D)$.

$$H1(D) = \frac{0.1 + 0.6D^2}{1 - 1.5D + 0.54D^2}$$

$$H2(D) = \frac{0.1 + 0.8D^2}{1 - 1.4D + 0.5D^2}.$$

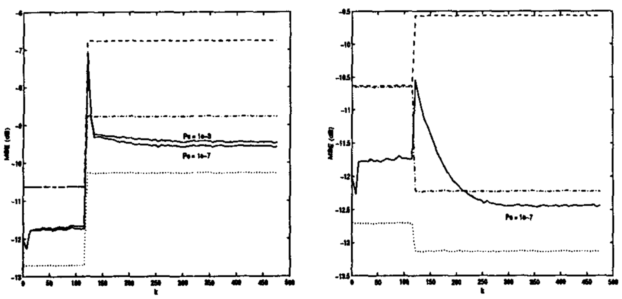
Length of FFT is chosen as $m = 512$. White noise is used in order to simplify simulation, i.e., $\Delta_i = \Delta$.

The averaged mean square error (MSE) per subchannel is defined as

$$err = \frac{\sum_{i \in U} err_i}{|U|}$$

where $err_i = \|X_i - \hat{X}_i\|^2$ is the MSE of the $i$th subchannel and $U$ is the set of all the used subchannels. $|U|$ is the number of all the used subchannels.

In our algorithm, first 2 frames of data are sent as pure training sequence to get the initial channel estimation. After that, the real data are sent and the joint channel estimation and equalization algorithm is used to track the variation of the channel. Fig. 2-4 show the results of the simulation. The solid lines in Fig. 2 and 4 show the result of the joint channel estimation and equalization algorithm. The DSL channel identification algorithm is also performed for comparison. The dashed lines show the result of using the channel estimation obtained by DSL for $H0(D)$ without retraining when channel changes, while the dash-dot lines show that of using the channel estimation obtained by DSL for $H(D)$ when channel changes. However, the training processes are not represented in the following simulation results, since the block lengths of the training sequence and data transmission are different. It should be noted that only the channel estimations obtained by DSL are used in the following simulation and extra training sequences are needed in order to get those estimations.



(a) (b)

Figure 2: Average MSE per Subchannel ($\Delta = 0.01, v = 128$)

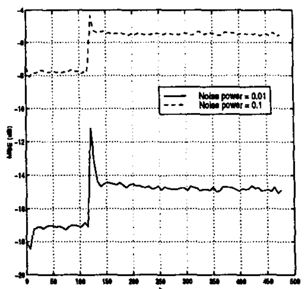



Figure 3: MSE of the 88th Subchannel ($Pe = 10^{-7}, v = 128$)

In Fig. 2, the length of the prefix is 128. $N$ and $L$ are 2 and 5 respectively. The average MSEs per subchannel are plotted. The channel changes from $H0(D)$ to $H1(D)$ in Fig. 2(a) while it changes to $H2(D)$ in Fig. 2(b). We can see that the algorithm converges when the channel varies. However, it converges faster in (a), in about 10 symbols, than in (b), in about 100 symbols. If we consider that the channel change is more dramatic in (b) than in (a), the result is satisfiable. Moreover, the algorithm can not only track the channel variation but also achieve about 1dB gain over the DSL method. In Fig. 2(a) the MSEs are plotted for both the target $P_e = 10^{-7}$ and $P_e = 10^{-3}$. In both cases, all 256 subchannels are used and the SNRs of each subchannel are identical. The only difference between these two cases is the minimum distance between the signal points, which means the errors of the estimation for the transmit prefix are different. The results shows that the adaptive algorithm is robust enough to such an estimation error. The MSE of $Pe = 10^{-3}$ is only a slightly larger than that of $Pe = 10^{-7}$.

In Fig. 3, the MSE of the 88th subchannel are plotted for noise power $\Delta = 0.01$ and 0.1 respectively, which means the SNR of this subchannel as $\Delta = 0.01$ is 10dB higher than that as $\Delta = 0.1$. This difference is compatible with the MSE difference in Fig. 3. As SNR goes down, the MSE is mainly caused by noise and the degradation brought by the inaccurate channel estimation becomes smaller.

Fig. 4 shows the result with much shorter length of prefix, $v = 64$. $N$ and $L$ are chosen as 4 and 7. The other conditions of Fig. 4 are the same as those in Fig. 2 and the result is also similar. The performance gain over the DSL reaches 2dB which is even larger compared to the case with longer prefix length. However, the algorithm converges slower, in more than 100 symbols. This is because we need to collect more symbols to get an accurate estimation, i.e, since $v$ is small, $N$ and $L$ must be large to maintain the dimension of $\mathbf{A}_{N,L}(k)$ to some level. If $N$ and $L$ are too small, the channel estimation error becomes large enough to cause the error propagation through the equalization and the algorithm can not converge. As the prefix length goes even smaller, the error brought by the channel memory becomes too large that will also make the adaptive algorithm collapses. The smallest number we tried in our simulation that can still make the algorithm converge is 32.

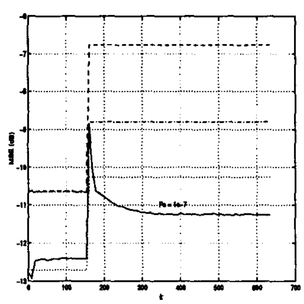


Figure 4: Average MSE per Subchannel ($\Delta = 0.01, v = 64$)

## 6. CONCLUSION

We have presented a joint channel estimation and equalization algorithm using the cyclic prefix in MCM system. This algorithm can adaptively track variation of a moderately time varying channel without additional training. Moreover, it also can give about 1-2dB improvement over the conventional training schemes.

## 7. REFERENCES


[1] John A. C. Bingham, "Multicarrier Modulation for Data Transmission: An Idea Whose Time Has Come", *IEEE Communications Magazine*, May, 1990, pp5-14.

[2] John M. Cioffi, *A Multicarrier Primer.*

[3] Sanjay Kasturia, James T. Aslanis and John M. Cioffi, "Vector Coding for Partial Response Channel", *IEEE Transaction on Information Theory*, v. 36, N. 4, July 1990, pp741-762.

[4] Jerome C. Tu, *Theory, Design and Application of Multi-Channel Modulation for Digital Communications*, Ph.D dissertation, Stanford University, June, 1991.

[5] Ye Li, Leonard J. Cimini, Jr, and Nelson R. Sollengerger, "Robust Channel estimation for OFDM Systems with Rapid Dispersive Fading Channels", *IEEE Transaction on Communications*, v. 46, N. 7, July 1998, pp902-915.