國 立 交 通 大 學

電信工程學系

碩 士 論 文

利用適應性濾波器為設計用於多輸入多輸出之

時空柵狀編碼系統之渦輪等化器

Adaptive Filter-Based Turbo Equalization for

MIMO Space-Time Trellis Coded Systems

研究生：曾俊偉

指導教授：紀翔峰

中 華 民 國 九十五 年 九 月

利用適應性濾波器為設計用於多輸入多輸出之

時空柵狀編碼系統之渦輪等化器

Adaptive Filter-Based Turbo Equalization for

MIMO Space-Time Trellis Coded Systems

研究生：曾俊偉　　Student：Jiun-Wei Tzeng

指導教授：紀翔峰博士　　Advisor：Dr. Hsiang-Feng Chi

國 立 交 通 大 學

電信工程學系碩士班

碩士論文

A Thesis

Submitted to Department of Communication Engineering
College of Electrical and Computer Engineering
National Chiao Tung University
in Partial Fulfillment of the Requirements
for the Degree of
Master
in
Communication Engineering

September 2006
Hsinchu, Taiwan, Republic of China

中華民國九十五年九月

利用適應性濾波器為設計用於多輸入多輸

出之時空柵狀編碼系統之渦輪等化器

研究生：曾俊偉　　指導教授：紀翔峰 博士

國立交通大學

電信工程學系碩士班

摘要

時空柵狀編碼是一個使用多根傳送天線的頻寬節省技術，當使用在寬頻的傳送系統時，通道的頻率選擇特性會造成不可忽略的影響。信號之間的互相干擾(ISI)以及同通道中的外來干擾(CCI)都會降低時空柵狀編碼所帶來的的效能增長。在這個論文裡，我們的目的是設計與建立一種接收的技巧，讓時空柵狀編碼系統即使經歷了具有頻率選擇效應的多輸入多輸出通道，仍能正確的接受資訊。我們設計了一個渦輪等化器[3]，使用一根或兩根接受天線來接受多根傳送天線的訊號。等化的動作是用一個簡單的適應性濾波器爲主的等化器。濾波器的係數是用一個名稱爲”最少平均平方” [2]的適應性演算法來獲得。在每次遞迴時，等化器和解碼器產生出外質的訊息，然後這些外質的訊息會在下次遞迴的時候被拿來使

用，就像渦輪碼一樣。我們用兩根、三根、四根的傳送天線來跑模擬，藉此來觀察傳送天線數目對系統效能的增進。關於不同大小的交錯器對我們的渦輪等化器的影響，我們也用模擬來觀察。這些模擬結果顯示出，我們提出的渦輪等化器，配上適當大小的交錯器，能夠有效的對抗多路徑造成的不好影響以及同通道中的干擾。

Adaptive Filter-Based Turbo Equalization for MIMO Space-Time Trellis Coded Systems

Student：Jiun-Wei Tzeng　　Advisor：Dr. Hsiang-Feng Chi

Department of Communication Engineering

National Chiao Tung University

Hsinchu, Taiwan

Abstract

Space-time trellis code (STTC) [1] is a bandwidth-efficient technique utilizing multiple transmit antennas. When it is applied to wideband systems, the channel's frequency-selective characteristics are not negligible. The inter-symbol interference (ISI) and the co-channel interference (CCI) will deteriorate the system performance improvement provided by STTC. In this thesis, our goal is to develop a receiving scheme for space-time trellis coded system over frequency-selective multiple-input-multiple-output (MIMO) channels. We design the turbo-equalization

[3] with one or two receive antennas to recover all symbols from multiple transmit antennas. Equalization is performed using a simple adaptive filter-based equalizer. The filter coefficients of the equalizer are obtained using an adaptive algorithm called "least-mean-squared (LMS)" [2]. At each iteration, extrinsic information is extracted from the detection and decoding units and is then used at the next iteration as in turbo-decoding. Simulations of two, three and four transmit antennas are conducted to see the performance improvement provided by the different number of transmit antennas. Simulations with different interleaver sizes are also conducted to observe the effect of the interleavers on a turbo system. The simulation results show that the proposed turbo-equalizer with an appropriate interleaver size can successfully combat the multipath effects and the co-channel interference.

謝 辭

在探究知識與理論的這條路上，重要的不僅是我得到了什麼，更重要的是我重新回想了我一生所學、所愛與被愛，在掙扎、矛盾與歡笑的交織中，重新發現、了解自己。

這本論文是我學生生涯的一個段落，也象徵著人生新里程碑的開始，不論未來是光明或黑暗，我將帶著這個小小的結晶再出發。

謝謝我的指導教授紀翔峰老師，總是耐心的引領我一步一步完成我的研究，還要感謝其他在學路上給予我教導的師長，讓我有足夠的力量迎向未來的挑戰。

謝謝 913 實驗室所有學長、同學陪我走過的歡笑和淚水，在我遇挫折時給我的鼓勵、幫忙，我永遠不會忘記那些一起奮戰的夜晚！

這篇謝辭更重要的是要特別感謝在我生命中扮演非常重要角色的人，謝謝我的寶貝品客，因爲妳，我才有動力完成這篇論文，妳讓我期待每一天的到來，因爲妳會讓我的每個下一刻都是幸福開心的，也希望我們會有更多美好的未來。

最後，感謝一直以來都很辛苦的媽媽和爸爸，謝謝你們給我很大的空間和疼愛，在我心目中你們是最棒最好的爸媽了！擁有你們真的讓我感到很幸福!

我想我一定還沒有謝完，真的很開心我可以成爲這麼幸福的人，我也會在你們的關愛環繞下，繼續我的下一個旅程。

俊偉 2006 年 9 月 7 日 於交通大學

Content

Figure Content

Table Content

CHAPTER 1

Introduction to MIMO Systems

The demands for high data rate transmission are increasing rapidly recently. But, the traditional approaches to increase data rate, such as increasing bandwidth, or increasing transmission rate, are becoming impractical due to the limited resource. This leads to considerable effort in finding and developing new approaches in addition to the two aforementioned approaches.

A popular approach is to develop a multiple-input-multiple-output (MIMO) system. In this chapter, we will show how the advantages are obtained from the MIMO system by deriving the fundamental capacity of MIMO channels and comparing it to that of the traditional single-input-single-output (SISO) channels.

At the end of this chapter, we will introduce the motivation of this work and the organization of this thesis.

1.1 MIMO Channel

1.1.1 MIMO Channel Model

Assuming a MIMO channel with n_T transmit antennas and n_R receive antennas, and the path between each antenna is frequency-flat, then we can express this system as:

$$\mathbf{r} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{1.1}$$

where $\mathbf{r} = \begin{bmatrix} r_1 & \cdots & r_{n_R} \end{bmatrix}^T$ is the $n_R \times 1$ received vector, $\mathbf{x} = \begin{bmatrix} x_1 & \cdots & x_{n_T} \end{bmatrix}^T$ is the $n_T \times 1$ transmitted vector, $\mathbf{n} = \begin{bmatrix} n_1 & \cdots & n_{n_R} \end{bmatrix}^T$ is the $n_R \times 1$ additive noise vector. $\mathbf{H}$ is the $n_R \times n_T$ fading matrix of the form:

$$\mathbf{H} = \begin{pmatrix} h_{11} & \cdots & h_{1n_T} \\ \vdots & \ddots & \vdots \\ h_{n_R 1} & \cdots & h_{n_R n_T} \end{pmatrix} \tag{1.2}$$

Let the total transmitted power be constrained to P, regardless of the number of transmit antennas n_T. The power can be represented as

$$P = tr(\mathbf{R}_{xx}) \tag{1.3}$$

where $tr(\bullet)$ denotes the trace obtained as the sum of diagonal elements, and $\mathbf{R}_{xx}$ denotes the covariance matrix obtained by:

$$\mathbf{R}_{xx} = E\left\{\mathbf{x}\mathbf{x}^H\right\} \tag{1.4}$$

where $E\{\bullet\}$ denotes expectation operation. It is common to consider the transmitted signals to be zero mean independent identically distributed (i.i.d.) Gaussian variables. Let us consider the case when transmitted power is split into n_T parts evenly and distributed to n_T transmit antennas, then equation (1.4) can be rewritten as

$$\mathbf{R}_{xx} = \frac{P}{n_T}\mathbf{I}_{n_T} \tag{1.5}$$

where $\mathbf{I}_{n_T}$ is the $n_T \times n_T$ identity matrix.

For normalization purposes, we assume that the received power for each of n_R receive antennas is equal to the total transmitted power P. Thus we obtain a normalization constraint for the elements of $\mathbf{H}$ as

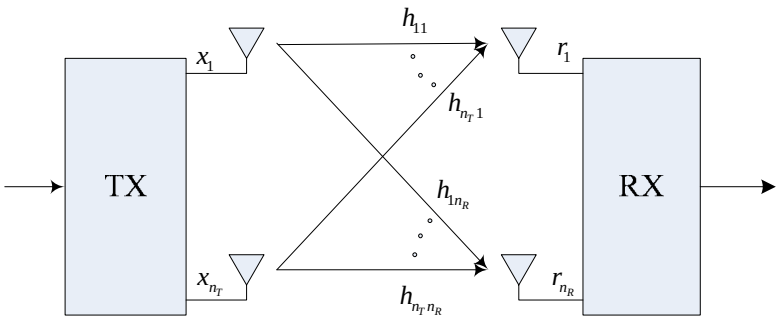

Figure 1.1 MIMO channel with n_T transmit antennas and n_R receive antennas

$$\sum_{i=1}^{n_T} |h_{ij}|^2 = n_T, \quad j = 1, \ldots, n_R \tag{1.6}$$

This system is shown in figure 1.1.

1.1.2 MIMO Channel Capacity

Before we start to derive the channel capacity of MIMO channels, we first review the Shannon's third theorem **[17]**, the information capacity theorem, as a reminder:

The information capacity of a continuous channel of bandwidth B hertz, perturbed by additive white Gaussian noise of power spectral density $N_0/2$ ***and limited in bandwidth to B, is given by***

$$C = B\log_2\left(1+\frac{P_r}{N_0 B}\right) \text{ bits per second} \tag{1.7}$$

where P_r represents the received signal power.

Let $N_0 B = \sigma_n^2$ be the total power of the noise, we rewrite the capacity formula as:

$$C = B\log_2\left(1+\frac{P_r}{\sigma_n^2}\right) \text{ bits per second} \tag{1.8}$$

Now we are ready to derive the channel capacity of MIMO channel. In light of

singular value decomposition (SVD) theorem [4], any $n_R \times n_T$ matrix $\mathbf{H}$ can be decomposed as:

$$\mathbf{H} = \mathbf{U}\mathbf{D}\mathbf{V}^H \tag{1.9}$$

where $\mathbf{D}$ is an $n_R \times n_T$ non-negative and diagonal matrix, $\mathbf{U}$ and $\mathbf{V}$ are $n_R \times n_R$ and $n_T \times n_T$ unitary matrices, respectively. The diagonal elements of $\mathbf{D}$ are the non-negative square roots of the eigenvalues of $\mathbf{H}\mathbf{H}^H$. Furthermore, the column vectors of $\mathbf{U}$ are the eigenvectors of $\mathbf{H}\mathbf{H}^H$, and the column vectors of $\mathbf{V}$ are the eigenvectors of $\mathbf{H}^H\mathbf{H}$. The non-negative square roots of the eigenvalues of $\mathbf{H}\mathbf{H}^H$ are also referred to as the singular values of $\mathbf{H}$. We denote $\sqrt{\lambda_i}$ as the ith element on the diagonal of $\mathbf{D}$, and then we have the following equations:

$$\mathbf{H}\mathbf{H}^H\mathbf{y} = \lambda_i\mathbf{y} \tag{1.10}$$

where $\mathbf{y}$ is an eigenvector corresponding to eigenvalue λ_i By unitary matrices, we imply the following equations exist:

$$\begin{aligned}\mathbf{U}\mathbf{U}^H &= \mathbf{I}\\ \mathbf{V}\mathbf{V}^H &= \mathbf{I}\end{aligned} \tag{1.11}$$

Substituting (1.9) into (1.1), and multiplying both sides by $\mathbf{U}^H$, we can rewrite equation (1.1) as:

$$\begin{aligned}\mathbf{U}^H\mathbf{r} &= \mathbf{U}^H\mathbf{U}\mathbf{D}\mathbf{V}^H\mathbf{x} + \mathbf{U}^H\mathbf{n}\\ &= \mathbf{D}\mathbf{V}^H\mathbf{x} + \mathbf{U}^H\mathbf{n}\end{aligned} \tag{1.12}$$

Introducing the transformations blow

$$\begin{aligned}\mathbf{r}' &= \mathbf{U}^H\mathbf{r}\\ \mathbf{x}' &= \mathbf{V}^H\mathbf{x}\\ \mathbf{n}' &= \mathbf{U}^H\mathbf{n}\end{aligned} \tag{1.13}$$

we obtain an equivalent channel model:

$$\mathbf{r}' = \mathbf{D}\mathbf{x}' + \mathbf{n}' \tag{1.14}$$

We say that equation (1.14) is equivalent to equation (1.1) is because $\mathbf{V}$ and $\mathbf{U}$ are unitary matrices, therefore the transformation that $\mathbf{H}$ represents is identical to

$\mathbf{D}$ with respect to different basis.

Due to the diagonal property of $\mathbf{D}$, each element in $\mathbf{r}'$ can be expressed as:

$$r_i' = \sqrt{\lambda_i} x_i' + n' \quad , i = 1, 2, \cdots n_R \tag{1.15}$$

If $rank(\mathbf{H}) = \gamma$, which also means $rank(\mathbf{H}\mathbf{H}^H) = \gamma$, then $\lambda_i \neq 0$ for $i = 1, 2, \cdots \gamma$ and $\lambda_i = 0$ for $i = \gamma + 1, \cdots n_R$. Rewrite equation (1.13) and apply the property above leads to:

$$\begin{aligned} r_i' &= \sqrt{\lambda_i} x_i' + n_i \quad ; i = 1, \cdots \gamma \\ r_i' &= n_i \qquad\qquad ; i = \gamma + 1, \cdots n_R \end{aligned} \tag{1.16}$$

This new equivalent model is shown in figure 1.2

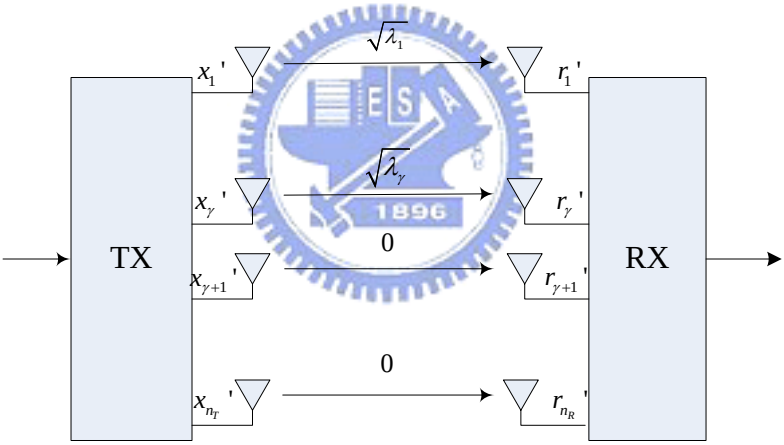

Figure 1.2 Equivalent MIMO channel after SVD

We have changed the system in figure 1.1 into figure 1.2. From figure 1.2, we see that the original MIMO channel can be expressed as γ parallel equivalent SISO channels, and each SISO channel capacity can be directly computed using the Shannon capacity formula reviewed in equation (1.8). From this viewpoint, the capacity of MIMO channel is now easily obtained by summing all the equivalent SISO channel capacities since they are parallel:

$$C = \sum_{i=1}^{\gamma} B \log_2 \left(1 + \frac{P_{ri}}{\sigma_i^{\,2}} \right) \text{ bits per second} \tag{1.17}$$

where P_{ri} denotes the received signal power of r_i', and $\sigma_i^{\,2}$ denotes the noise power of n_i'. P_{ri} can be obtained using following equations under our normalization constraint (1.6)

$$P_{ri} = tr(\mathbf{R}_{r_i' r_i'}) = tr(\mathbf{U}^H \mathbf{R}_{r_i' r_i'} \mathbf{U}) = tr(\mathbf{R}_{r_i' r_i'}) = tr(\mathbf{H}\mathbf{R}_{x_i' x_i'} \mathbf{H}^H) = \frac{\lambda_i P}{n_T} \tag{1.18}$$

Thus, the channel capacity can be written as

$$\begin{aligned} C &= \sum_{i=1}^{\gamma} B \log_2 \left(1 + \frac{\lambda_i P}{\sigma_i^{\,2} n_T} \right) \\ &= B \log_2 \prod_{i=1}^{\gamma} \left(1 + \frac{\lambda_i P}{\sigma_i^{\,2} n_T} \right) \quad \text{bits per second} \end{aligned} \tag{1.19}$$

Since $\lambda_i \ , i = 1, \ldots, \gamma$ are the eigenvalues of $\mathbf{H}\mathbf{H}^H$, to obtain all λ_i we have to find the roots of such equation:

$$\det(\lambda \mathbf{I} - \mathbf{H}\mathbf{H}^H) = 0 \tag{1.20}$$

We can also rewrite equation (1.20) by introducing the roots into the equation:

$$\prod_{i=1}^{\gamma} (\lambda - \lambda_i) = 0 \tag{1.21}$$

Therefore, we can equal these two equations as blow:

$$\det(\lambda \mathbf{I} - \mathbf{H}\mathbf{H}^H) = \prod_{i=1}^{\gamma} (\lambda - \lambda_i) \tag{1.22}$$

Substituting $-\frac{n_T \sigma^2}{P}$ for λ in (1.22), and multiplying both sides a constant, we get

$$\det(\mathbf{I} + \frac{P}{n_T \sigma^2} \mathbf{H}\mathbf{H}^H) = \prod_{i=1}^{\gamma} (1 + \frac{\lambda_i P}{n^T \sigma^2}) \tag{1.23}$$

The purpose of doing this is to write the channel capacity formula in(1.19) as

$$C = B \log_2 \det \left(\mathbf{I} + \frac{P}{n_T \sigma^2} \mathbf{H}\mathbf{H}^H \right) \tag{1.24}$$

1.1.3 Antenna Numbers versus Capacity

From the capacity formula (1.17), it is obvious that γ is an important factor in

determining the capacity. It represents the equivalent parallel SISO channel number. Since $\mathbf{H}$ is a $n_R \times n_T$ matrix, the rank of $\mathbf{H}$, γ, will always be less or equal to $\min(n_R, n_T)$. If $n_R \le n_T$, the equivalent parallel SISO channel number of the MIMO channel is less or equal to the transmit antenna number. If $n_T \le n_R$, the equivalent parallel SISO channel number of the MIMO channel is less or equal to the receive antenna number. It is often a fallacy to think that more antennas there are, more equivalent SISO channels we must have. The fact is: it depends on the rank of the channel. A 3-to-2 MIMO system may give less equivalent SISO channels than a 2-to-2 MIMO system. But for full-rank channels defined as the cases $\gamma = \min(n_R, n_T)$, of course, more antennas give more capacity. Therefore, a full-rank channel is always considered as a good channel for a MIMO system.

1.1.4 MIMO Capacity over Frequency-Selective Channels

In this section, we consider the capacity of an OFDM-based MIMO channel, and then extend it to general frequency-selective channel capacity. For a K-point OFDM-based system, channel can be considered as K parallel subchannels. If K is large enough, then each subchannel can be approximated as frequency-nonselective. Therefore, a frequency-selective MIMO channel in an OFDM-based system can be approximated as K frequency-nonselective MIMO subchannels if K is large enough as illustrated in figure 1.3. The instantaneous channel capacity of such system is given in **[5]** as

$$C \cong \frac{B}{K}\sum_{k=1}^{K}\log_2\left[\det(\mathbf{I} + SNR^k\mathbf{H}^k\cdot(\mathbf{H}^k)^H\right] \tag{1.25}$$

where $\mathbf{H}^k$ represents the kth subchannel which is a $n_R \times n_T$ matrix, and SNR^k represents the Signal-to-Noise Ratio of the kth subchannel at the receive antenna, and B is the bandwidth of the overall channel.

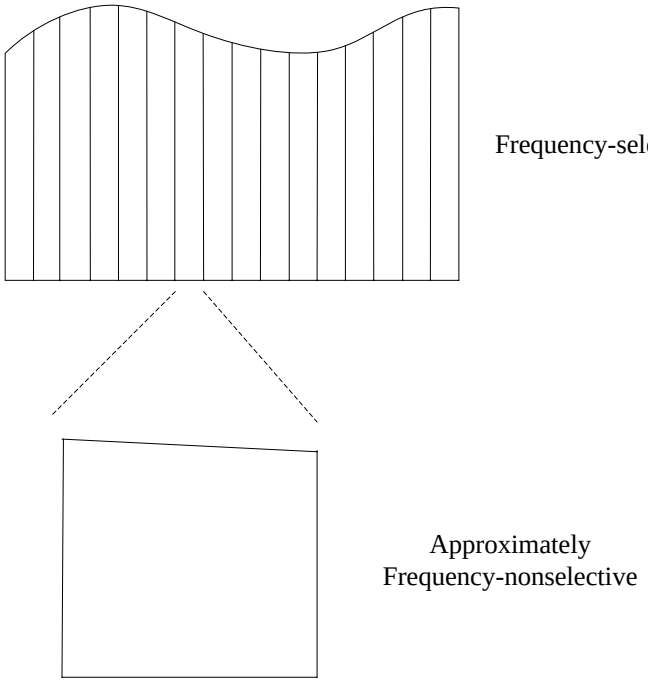

Figure 1.3 Subchannel of OFDM systems

By comparing equations (1.25) and (1.24), it is not difficult to see that this formula is simply to sum the capacities over the K frequency-nonselective subchannels. The derivation seems plausible; however, it is an approximation because we approximate each subchannel to be frequency-nonselective. To derive the exact formula instead of the approximation, Riemann integral theorem can be applied. That is, when K is approaching to infinite, all subchannels (with the bandwidth $\frac{B}{K}$) become frequency-nonselective. The summation in (1.25) becomes integration, and we obtain the exact MIMO frequency-selective channel formula:

$$C = B\int_0^B \log_2\left[\det\left(\mathbf{I}+\mathbf{H}(f)\mathbf{S}_{xx}(f)\mathbf{H}^H(f)\mathbf{S}_{nn}{}^{-1}(f)\right)\right]df \quad \text{bits per second} \qquad (1.26)$$

where $\mathbf{S}_{xx}(f)$ represents power spectrum matrix of x at frequency f obtain by $\mathbf{S}(f) = FT\{\mathbf{R}\}$, and $\mathbf{H}(f)$ is the frequency response of the MIMO channel at frequency f.

1.2 Introduction to Receivers for MIMO Channel

There are many different transmitting schemes for MIMO systems. The Space-time trellis code is a well-known transmitting scheme which introduces diversity and coding gain. However, as the transmission bandwidth increases beyond the coherent bandwidth of the channel, ISI becomes a major performance-limiting impairment. Equalization becomes indispensable. The theory of equalization for single-input single-output channels has been well-developed. Different criteria and algorithms such as MMSE, BCJR, LMS and RLS are applied to obtain a variety of equalizers. In the recently years, the techniques of turbo equalization are proposed and shown to provide much better performance than the tradition receiving schemes in which the equalization and the decoding are conducted separately. The turbo equalization scheme performs equalization and decoding in an iterative manner and obtains the performance near the Shannon limit.

When MIMO channels are considered, these design criteria and algorithms for the SISO systems remain valid and effective. However, due to the nature of MIMO channel, some modifications or extra effort have to be made. Trellis-based equalizers suffer from the heavy computational complexity. This problem can be mitigated by the use of several techniques such as prefiltering **[25]** and in-phase/quadrature-phase detection **[26]**. Filter-based equalizers are low-complexity alternatives. With known channel state information (CSI), MMSE criterion can be derectly applied to the filter-based equalizer as it does in **[23]** for SISO channels. Without the CSI, MMSE solutions can be approached by the adaptive algorithms such as LMS and RLS as it does in **[21]** for SISO channels.

In this thesis, the channel state information is assumed to be unknown at the receiver, and we apply adaptive algorithm directly to the equalizer. This structure

can be deemed as the MIMO version of the one in **[21]**.

1.3 Motivation

Achieving high bit rates over bandlimited wireless channels makes many applications possible. The use of multiple-antennas brings a bandwidth efficient solution while achieving high bit rate at the same time. The fundamental phenomenon which makes reliable wireless transmission difficult is the multipath fading. Therefore, to overcome this phenomenon is especially important. For some applications, the receivers are also required to be small and power efficient. Thus, a simple receiver structure is attractive. However, the equalization for MIMO channels is generally considered impractical due to the heavy computational complexity. In this thesis, our goal is to develop a low-complexity turbo equalizer for broadband MIMO systems.

1.4 Organization

This thesis is organized as follows. Chapter 2 introduces the space-time trellis codes, and its advantages over single-input-single-output trellis codes. The encoding structure and the decoding algorithm are addressed in details. Chapter 3 is the equalizer overview. The filter-based equalizer is the candidate of our design and thus is the main topic of this chapter, too. The ideal of turbo equalization is introduced in chapter 3. In chapter 4, we propose the receiver in combination of the topics in chapter 2 and 3. In this chapter, the details of the receiver are given, including equalization, decoding, interleaving, mapping, and demapping. In chapter 5, we show the simulation results of our proposed system, and some comparison of the systems. Chapter 6 is the conclusion we make on the simulation results and comparisons.

CHAPTER 2

Space-Time Trellis Codes

In chapter 1, we have shown the capacity a MIMO channel, but no system implementation has been mentioned. In this chapter, we will introduce the Space-Time Codes which is one approach to take advantages of MIMO channel and provide high data rate or high link quality. Space-time codes have a variety of different structures, each of them has their one advantages and disadvantages. In this chapter, we will first introduce briefly the three diversities commonly used in communication systems. Then we move on to space-time codes which may combine several of the three diversities. The famous space-time block codes will be briefly introduced, and then we focus on the space-time trellis codes which are the channel codes we used in our proposed system in chapter 4.

2.1 Diversity Techniques

Many channels, especially wireless channels, suffer from attenuation due to destructive addition of multipaths in the propagation media and due to the interference from other users. Severe attenuation makes it impossible for the receiver to determine the transmitted signal unless some less-attenuated replica of the transmitted signal is provided to the receiver. This resource is called diversity and it is an

important contributor to reliable wireless communications. According to the domain where diversity is introduced, there are three categories of diversity: time diversity, frequency diversity, space diversity.

2.1.1 Time Diversity

The replicas of transmitted signals are provided to the receiver in the form of redundancy in time domain. Identical messages are transmitted in different time slots, and the receiver would receive several uncorrelated fading signals. To be uncorrelated, the time separation between identical messages must be at least the coherent time of the channel. The definition of the coherent time is the period over which the channel fading process is correlated. In many systems, redundancy in time domain is introduced by the error control coding (ECC), and an interleaver is placed after error control coding to provide time separation greater than the coherent time. However, in the receiver, deinterleaving process introduces message delay. For slow fading channels, a larger interleaver is required to exceed the coherent time, and therefore, a larger message delay is introduced. This drawback may be vital to some delay-sensitive applications, especially voice applications. Another drawback of this scheme is that there will be a certain bandwidth efficiency loss due to the redundancy in time domain.

2.1.2 Frequency Diversity

The replicas of transmitted signals are provided to the receiver in the form of redundancy in frequency domain. The frequency separation is required to be at least the coherent bandwidth to obtain uncorrelated fading replicas in the receiver. The definition of the coherent bandwidth is similar to the coherent time: the frequency span over which the channel fading property is uncorrelated. Several mature

communication systems introduce the frequency diversity to increase the data rate or improve the link quality. Spread spectrum is one example, this technique includes direct sequence spread spectrum (DSSS), frequency hopping, multicarrier modulation, and CDMA systems. A combination of error-control coding and OFDM can also be considered as frequency diversity, because the time diversity provided by ECC has been transferred into frequency domain by OFDM modulator. These techniques use bandwidths that are far more than enough just to provide frequency diversity, thus, like time diversity, it induces a loss in bandwidth efficiency due to the redundancy introduced in frequency domain.

2.1.3 Space Diversity

The replicas of transmitted signals are provided to the receiver in the form of redundancy in spatial domain. It is typically implement using multiple antennas or antenna arrays arranged in space in a certain manner. Therefore, space diversity is also called antenna diversity. The space separation between antennas is required to be at least the coherent distance. Usually, a few wavelengths are enough for the antennas to experience different fadings. One advantage of this technique is that unlike time and frequency diversity, it doesn't suffer from the loss in bandwidth efficiency. This advantage makes it very attractive to high data rate wireless communications.

2.1.4 Transmit and Receive Diversity

We can further classify space diversity into receive diversity and transmit diversity depending on where the multiple antennas are applied. The receive diversity is adapted in a variety of mobile communication systems with the aim to both suppress co-channel interference and minimize the fading effects. It is reasonable to apply

receive diversity at the base station for uplink (from mobiles to base stations) communication because the power requirement and the dimension requirement are easier to meet compared with mobile devices. For example, in GSM systems, multiple antennas are used at the base station to create uplink receive diversity, compensating for the relatively low transmission power from the mobile. For downlink (from base station to mobiles), it is much more difficult to apply receive diversity at the mobiles. Firstly, placing multiple antennas in a portable mobile device is against the public favor in a smaller device. Secondly, multiple antennas mean more power consumption which is also against the public favor in a power-saving device. Therefore, transmit diversity is more adequate for the downlink communication.

However, in contrast to receive diversity which is widely applied in mobile systems, transmit diversity has gained little attention and is less understood. The reason is that it is more difficult to exploit transmit diversity, and the difficulty is because the transmitted signals are mixed up before they arrive the receiver. Therefore, the receiver requires extra signal processing to separate the transmitted signals before the transmit diversity can be exploited. This signal processing is not always perfect, and may suffer from performance loss. In the next session, we will introduce a transmit diversity technique called “Space-Time Codes”, and in Chapter 4, we will propose a signal processing to separate the transmitted signals.

2.1.5 Combine Different Diversity

Not all forms of diversity can be available at all times. To obtain diversity, the resources providing redundancies must be uncorrelated. Otherwise, you simply obtain the same information twice. For example, in slow fading channels, time diversity is not an option due to large coherent time. When delay spread is small,

frequency diversity is not an option due to large coherent bandwidth. When the platform is a small mobile device, space diversity is not suitable due to limited dimensions. Nevertheless, combining different diversities will improve the data rate or link quality if they are available. Transmit diversity and receive diversity can also be combined to provide more advantages.

2.2 Space-Time Coding

In 1993, Wittneben proposed a delay diversity scheme [6]. This scheme transmits the same information from both antennas but with a delay of one symbol interval as shown in figure 2.1. The effect of this process is to introduce an artificial multipath channel to the receiver which changes a narrowband purely frequency-nonselective fading into a frequency-selective fading channel. With this multipath channel, the receiver can utilize a trellis-based equalizer and gain performance improvement from it. In [7], it was shown that the use of a maximum-likelihood sequence estimator at the receiver is capable of providing dual branch diversity.

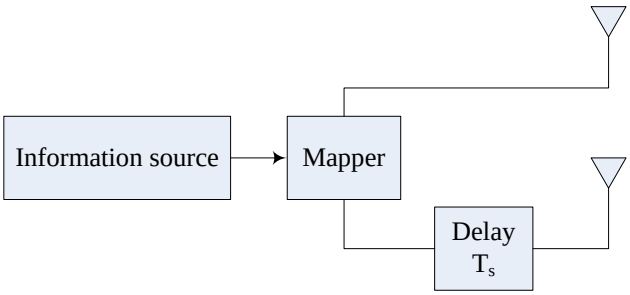

Figure 2.1 Delay diversity transmitter

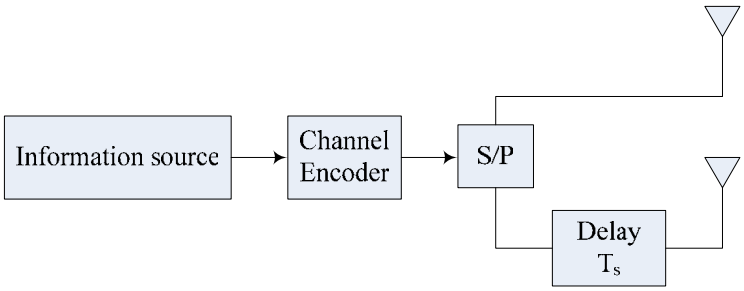

Figure 2.2 Delay diversity transmitter

This framework is identical to figure 2.2 where the channel code is a repetition code with code rate $R=\frac{1}{2}$. In this structure, it is natural to ask if it is possible to design a channel code that is better than repetition code in order to improve performance further. The answer is yes, and we use the term "Space-Time Codes" to refer to these channel codes. In space-time coding, the delay element before antenna unit is removed, and the structure is shown in figure 2.3. Coding is performed in both spatial and temporal domains to introduce correlation between signals transmitted from various antennas at various time periods. Space-time coding can achieve transmit diversity over spatially uncoded systems without sacrificing the bandwidth. There are various approaches in coding structures, including space-time block codes (STBC), and space-time trellis codes (STTC).

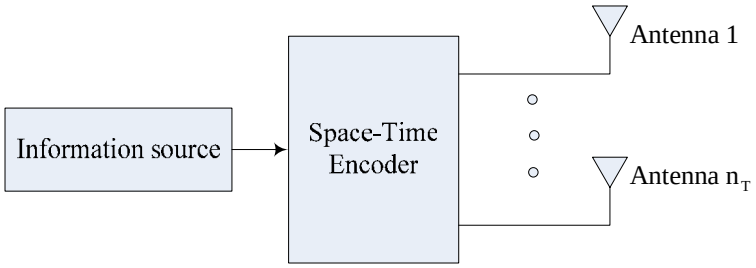

Figure 2.3 Space-time coded transmitter

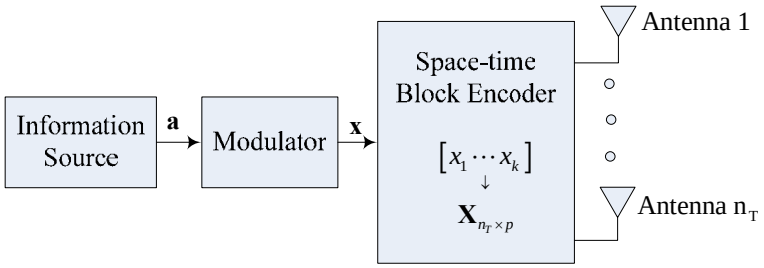

Figure 2.4 Space-time block coded transmitter

2.2.1 Space-Time Block Codes

Space-time block codes operate on a block of input symbols producing a matrix output whose columns represent time and rows represent antennas. Their key feature is the provision of full diversity with extremely low encoder/decoder complexity under frequency-nonselective channels. The space-time block code system is shown in figure 2.4. Assuming n_T antennas are used, and p symbols per antenna are used to convey k uncoded symbols. The code rate is given by

$$R = \frac{k}{p} \tag{2.1}$$

To describe the space-time block codes, we use the transmission matrix $\mathbf{X}$ which is a $n_T \times p$ matrix. The element of $\mathbf{X}$ in the ith row and jth column, $x_{i,j}$, $i = 1,\ldots,n_T$, $j = 1,\ldots,p$ represents the signals transmitted from the antenna i at time j.

The key property of this system is the orthogonality between the sequences generated by the different transmit antennas. This feature was generalized in [8] to an arbitrary number of transmit antennas by applying the theory of orthogonal designs. It is also shown in [8] that to achieve full transmit diversity, the code rate of a space-time block code must be less than or equal to one, $R \le 1$ which requires an bandwidth expansion of $\frac{1}{R}$.

Each element of the transmission matrix $\mathbf{X}$ is a linear combinations of the k

modulated symbols $x_1,\ldots,x_k$ and their conjugates $x_1^*,\ldots,x_k^*$. To achieve the orthogonality, $\mathbf{X}$ must satisfy the following equation:

$$\mathbf{X}\cdot\mathbf{X}^H = c\left(|x_1|^2+\cdots|x_k|^2\right)\mathbf{I}_{n_T} \tag{2.2}$$

where c is a constant, $\mathbf{X}^H$ is the Hermitian of $\mathbf{X}$ and $\mathbf{I}_{n_T}$ is an $n_T\times n_T$ identity matrix. Through the orthogonal designing, the signal sequence from any two transmit antennas are orthogonal. This property enables the receiver to decouple the signals transmitted from different antennas and consequently, a simple maximum likelihood decoding, based only on linear processing of the received signals.

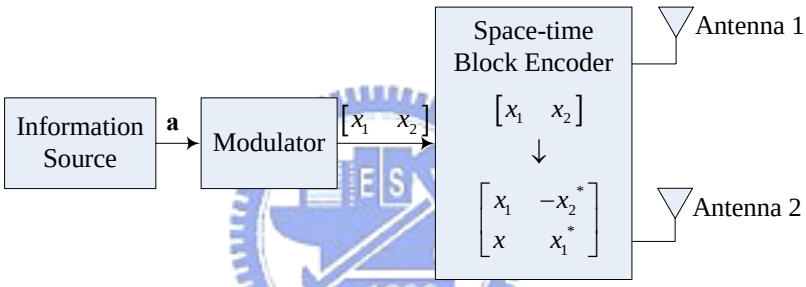

Figure 2.5 Alamouti space-time block encoder

The Alamouti code [9], was the first and the most famous space-time block code. It is achieve a full diversity gain using two transmit antennas and a simple maximum-likelihood decoding algorithm. The transmission matrix is given by

$$\mathbf{X}=\begin{bmatrix} x_1 & -x_2^* \\ x_2 & x_1^* \end{bmatrix} \tag{2.3}$$

Figure 2.5 shows an encoder structure for Alamouti code. It is easy to see that the transmit sequence from antennas one and two are orthogonal,

$$\mathbf{x}^1\cdot\mathbf{x}^2 = x_1x_2^* - x_2^*x_1 = 0 \tag{2.4}$$

The decoding of space-time block codes was based on maximum likelihood algorithm. Assuming one receive antenna and perfect channel state information (CSI) is know at the receiver. Using Alamouti code as an example, the maximum

likelihood decoder choose a pair of signals $(\hat{x}_1, \hat{x}_2)$ from the signal modulation constellation to minimize the distance metric

$$d^2\left(r_1, h_1\hat{x}_1 + h_2\hat{x}_2\right) + d^2\left(r_2, h_1\hat{x}_2^* + h_2\hat{x}_1 *\right) \tag{2.5}$$

With some simple transformation, the decision rule can be derived as

$$\begin{aligned} \hat{x}_1 &= \arg\min_{\hat{x}_1 \in \mathbf{S}} d^2\left(\tilde{x}_1, \hat{x}_1\right) \\ \hat{x}_2 &= \arg\min_{\hat{x}_2 \in \mathbf{S}} d^2\left(\tilde{x}_2, \hat{x}_2\right) \end{aligned} \tag{2.6}$$

where $\mathbf{S}$ is the set of all possible modulated symbol pairs $(\hat{x}_1, \hat{x}_2)$ and

$$\begin{aligned} \tilde{x}_1 &= h_1^* r_1 + h_2 r_2^* \\ \tilde{x}_2 &= h_2^* r_1 + h_1 r_2^* \end{aligned} \tag{2.7}$$

The decision rules derivation can be extended to other cases with other number of receive antennas, and they can be found in **[8]**.

Space-time block codes can achieve a maximum possible diversity advantage with a simple decoding algorithm. It is very attractive because of its simplicity. However, no coding gain can be provided by space-time block codes, and also, non-full rate space-time block codes can introduce bandwidth expansion.

2.2.2 Space-Time Trellis Codes

Space-time trellis codes operate on one input symbol at a time producing a sequence of vector symbols whose length represents antenna number. Like traditional trellis coded modulation (TCM) for the single-antenna channel, space-time trellis codes provide coding gain. Since they also provide full diversity gain, their key advantage over space-time block codes is the provision of coding gain. Space-time trellis codes are nowadays widely discussed as it can simultaneously offer a substantial coding gain, spectral efficiency, and diversity improvement on flat fading channels. The case for frequency-selective channels are studied in **[13][18]**, and the conclusion of **[13]** is that the frequency-selective channel doesn't affect the diversity

provided by space-time trellis codes. However, compared with space-time block codes, this code is far more difficult to design, and also requires a computationally intensive encoder and decoder. The key development was done by Tarokh, Seshadri and Calderbank in 1998 **[1]**, and some other improved development was done in **[10][11][12]**. In next session, we will talk about it in detail.

2.3 Space-Time Trellis Codes

2.3.1 Trellis description

The space-time trellis codes are described by trellis structures. Considering a space-time trellis coded M-PSK modulation with n_T transmit antennas. The encoder takes a group of $m = \log_2 M$ information bits at time t given by

$$\mathbf{a}_t = \left(a_t^{\,1}, \ldots, a_t^{\,m}\right) \tag{2.8}$$

and produces a group of mn_T coded bits

$$\mathbf{c}_t = \left(\left(c_{t,1}^{\;\;1}, \ldots, c_{t,m}^{\;\;1}\right), \ldots, \left(c_{t,1}^{\;\;n_T}, \ldots, c_{t,m}^{\;\;n_T}\right)\right) \tag{2.9}$$

Each $\left(c_{t,1}^{\;\;i}, \ldots, c_{t,n_T}^{\;\;i}\right)$, $i = 1, \ldots, n_T$ are mapped into symbol $x_t^{\,i}$, and thus $\mathbf{c}_t$ is mapped into a group of n_T symbols given by

$$\mathbf{x}_t = \left(x_t^{\,1}, \ldots, x_t^{\,n_T}\right) \tag{2.10}$$

where each $x_t^{\,i}$,$i = 1, \ldots, n_T$ is an M-PSK modulated signal. At each time t, depending on the state of the encoder and the input bits, a transition branch is chosen. On the trellis, each branch transition is labeled of $\left(x_t^{\,1}, \ldots, x_t^{\,n_T}\right)$ which means the transmit antenna i is used to send symbol $x_t^{\,i}$, and all these transmissions are simultaneous.

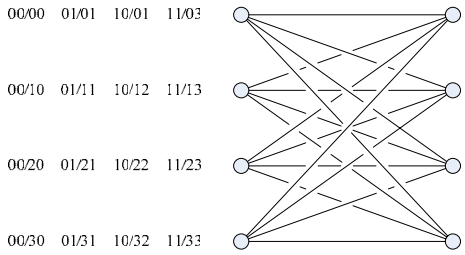

Figure 2.6 Trellis description for a 4-state space-time trellis codes with 2 transmit antennas and 4-PSK signal constellation

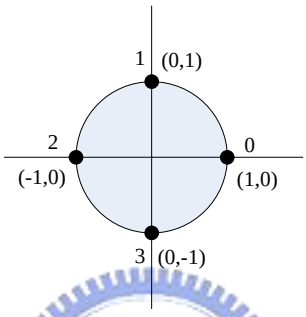

Figure 2.7 4-PSK signal constellation

Figure 2.6 shows an example of trellis description for a 4 states space-time trellis code with 2 transmit antennas and 4-PSK signal constellation. 4-PSK signal constellation is given in figure 2.7. The number pairs $a_t^1\ a_t^2 / x_t^1\ x_t^2$ in front of each state are the labels of each branch transition starting from that state. The left-most number pair is corresponding to the top-most branch transition of the state. As an illustration, if the encoder is in the second state at time t, the input at this time is 11, then the encoder chooses the branch transition from the second state to the fourth state. This branch transition is labeled 11/13 which means antennas (1,2) will transmit the symbol (1,3) respectively.

The trellis is a representation which can fully describe the space-time trellis codes. However, it is common to describe the space-time trellis codes as generator descriptions in the encoders.

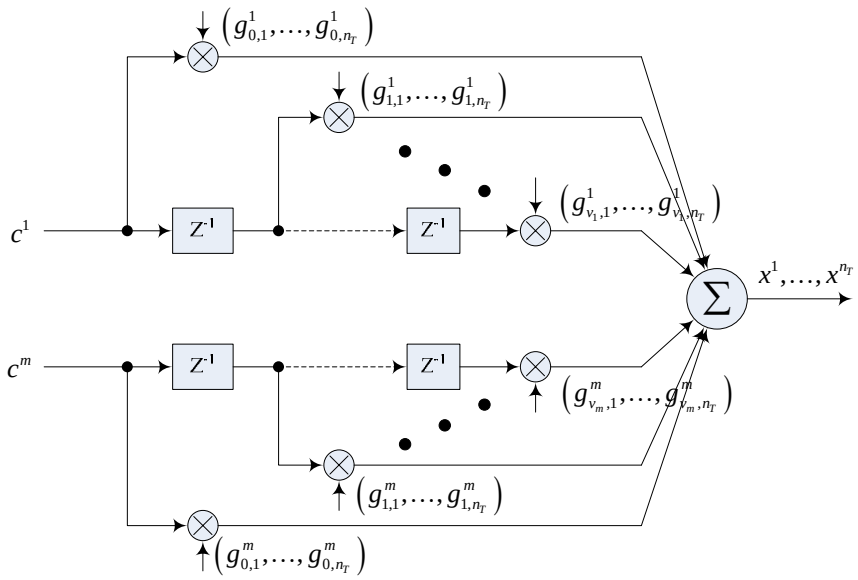

Figure 2.8 Encoder structure of space-time trellis codes

2.3.2 Generator Description

The encoder structure of space-time trellis code is shown in figure 2.8. The k-th input sequence $\mathbf{a}^k=\left(a_0{}^k,\ldots,a_t{}^k,\ldots\right)$, $k=1,\ldots,m$ is fed into the k-th shift register and multiplied by a generator sequence $\mathbf{g}^k$, $k=1,\ldots,m$. The multiplier outputs from all shift registers are added up in modulo-M to give the encoder output $\mathbf{x}=\left(\mathbf{x}_0,\ldots,\mathbf{x}_t,\ldots\right)$. The generator sequence $\mathbf{g}^k$, $k=1,\ldots,m$ is of the form:

$$\mathbf{g}^k=\left[\left(g_{0,1}^{k},\ldots,g_{0,n_T}^{k}\right),\ldots,\left(g_{v_k,1}^{k},\ldots,g_{v_k,n_T}^{k}\right)\right],\ k=1,\ldots m \tag{2.11}$$

The total memory order of the encoder, denoted by v is given by

$$v=\sum_{k=1}^{m}v_k \tag{2.12}$$

where v_k, $k=1,\ldots,m$ is the memory order for the k-th lane of shift register, and is given by

$$v_k=\left\lfloor\frac{v+k-1}{\log_2 M}\right\rfloor \tag{2.13}$$

The encoder output at time t for transmit antenna i is now given as

$$x_t^i = \sum_{k=1}^{m} \sum_{j=0}^{v_k} g_{j,i}^k a_{t-j}^k \mod M, \quad i = 1, \ldots, n_T \tag{2.14}$$

As an example, let us consider a scheme of 4-state space-time trellis coded QPSK system with 2 transmit antennas and the generator sequences are

$$\begin{aligned} \mathbf{g}^1 &= [(02),(20)] \\ \mathbf{g}^2 &= [(01),(10)] \end{aligned} \tag{2.15}$$

The resulting trellis structure is shown in figure 2.6, which is the same as the one we use as an example in the previous session.

2.3.3 Design Criteria

For a given encoder structure, a set of encoder coefficients is determined by minimizing the error probability. It is shown in **[1]** that the error rate for slow fading channels, the upper bound is depend on the value of r which is the rank of codeword distance matrix $\mathbf{A}(\mathbf{X}, \hat{\mathbf{X}})$. It can be obtained by utilizing the codeword difference matrix $\mathbf{B}(\mathbf{X}, \hat{\mathbf{X}})$

$$\mathbf{A}(\mathbf{X}, \hat{\mathbf{X}}) = \mathbf{B}(\mathbf{X}, \hat{\mathbf{X}}) \cdot \mathbf{B}^H(\mathbf{X}, \hat{\mathbf{X}}) \tag{2.16}$$

where $\hat{\mathbf{X}}$ is the erroneous decision mad by decoder when the transmitted sequence was in fact $\mathbf{X}$. Therefore, in order to minimize the error probability, we have two different criteria:

- **Rank & determinant criteria**: If $rn_R < 4$, the minimum rank r over all pairs of distinct codewords should be maximized. Also, the determinant of $\mathbf{A}(\mathbf{X}, \hat{\mathbf{X}})$ along the pairs of distinct codewords with the minimum rank should maximized, too.
- **Trace criteria**: If $rn_R \geq 4$, the minimum trace of $\mathbf{A}(\mathbf{X}, \hat{\mathbf{X}})$ among all pairs of distinct codewords should be maximized.

These criteria were derived and summarized in [1] along with the criteria for fast fading channels. The above criteria was referred to as the Tarokh/Seshadri/Calderbank (TSC) codes. An improved criterion was proposed and referred to as the Baro/Bauch/Hansmann (BBH) codes [14].

2.4 Decoding Algorithm

2.4.1 Maximum A posterior Probability (MAP) Decoder

To recover the information bits at the receiver, it is natural to choose a receiver that achieves the minimum probability of error $P(a_k \neq \hat{a}_k)$ where $\hat{a}_k$ is the decision of k-th information bit. It is well known that this is achieved by setting $\hat{a}_k$ to the value which maximizes the a posteriori probability (APP) $P(a_k = \hat{a}_k | \mathbf{y})$ given the received sequence $\mathbf{y}$, i.e.

$$\hat{a}_k = \arg\max_{\hat{a}_k} P(a_k = \hat{a}_k | \mathbf{y}) \tag{2.17}$$

The algorithms that achieve this task are commonly referred to as maximum a posteriori probability (MAP) algorithms.

When probabilities are concerned, it is often convenient to work with log-likelihood ratios (LLRs) rather than actual probabilities. The LLR for a variable a_k is defined as

$$L(a_k) = \ln \frac{P(a_k = 1)}{P(a_k = 0)} \tag{2.18}$$

The LLRs contains the same information as the probabilities $P(a_k = 1)$ or $P(a_k = 0)$. In fact, the sign of $L(a_k)$ determines whether $P(a_k = 1)$ is larger than $P(a_k = 0)$ or it is on the contrary. Therefore, by using this property of LLRs, the decision rules (2.17) can be further written as

$$\hat{a}_k = \begin{cases} 1, & L(a_k | \mathbf{y}) \geq 0 \\ 0, & L(a_k | \mathbf{y}) < 0 \end{cases} \tag{2.19}$$

The main problem of the MAP approach is that the APPs calculation is computation-intensive. To show this, we use Bayes' rule and the theorem of total probability **[15]** on $P(a_k = \hat{a}_k \mid \mathbf{y})$, and obtain

$$P(a_k = \hat{a}_k \mid \mathbf{y}) = \sum_{\forall \mathbf{a}: a_k = \hat{a}_k} P(\mathbf{a} \mid \mathbf{y}) = \sum_{\forall \mathbf{a}: a_k = \hat{a}_k} \frac{P(\mathbf{y} \mid \mathbf{a}) P(\mathbf{a})}{P(\mathbf{y})} \tag{2.20}$$

The computation loading in this equation is too heavy. Therefore we introduce algorithms which require less computation complexity than (2.20).

2.4.2 BCJR Algorithm

This algorithm was proposed in1974 by Bahl, Cocke, Jelinek, and Raviv **[16]**, and is now known as the name "BCJR algorithm" composed of the initials of four authors. This is a algorithm to efficiently compute the APPs of interest $P(a_k = \hat{a}_k \mid \mathbf{y})$.

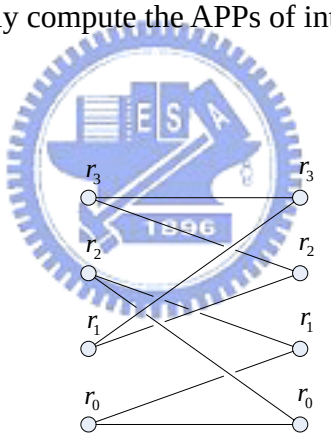

Figure 2.9 A trellis example with 4 states

We denote the pair (i, j) as the branch transition from state r_i to state r_j. Define a set $\boldsymbol{\beta}$ of (i, j), and each pair in the set is a valid branch transition to the chosen trellis. For example, the trellis in figure 2.9 has the set

$$\boldsymbol{\beta} = \{(00),(01),(12),(13),(20),(21),(32),(33)\} \tag{2.21}$$

We also denote $s_k \in S$ as the state of the encoder when k-th set of information bits is the input. We denote by $S = \{r_0, \ldots, r_{2^v - 1}\}$ as the set of all possible states where v

is the memory order of the encoder.

We begin with the computation of the probability that the transmitted sequence path in the trellis contained the branch transition from state r_i to state r_j when the k-th set of information bits is the input, i.e., to compute $P(s_k = r_i, s_{k+1} = r_j \mid \mathbf{y})$. Applying the chain rule for joint probabilities, i.e., $P(a,b) = P(a)P(b \mid a)$, we can obtain

$$P(s_k, s_{k+1}, \mathbf{y}) = P(\mathbf{y}) \cdot \mathrm{P}(\mathrm{s_k}, \mathrm{s_{k+1}} | \mathbf{y}) \tag{2.22}$$

The left hand side of this equation can be written as

$$P(s_k, s_{k+1}, \mathbf{y}) = P\left(s_k, s_{k+1}, (y_1, \ldots y_{k-1}), y_k, (y_{k+1}, \ldots y_N)\right) \tag{2.23}$$

Applying the chain rules again on (2.23), we obtain the key decomposition

$$\begin{aligned} P\left(s_k, s_{k+1}, \mathbf{y}\right) &= P\left(s_k, (y_1, \ldots y_{k-1})\right) P\left(s_{k+1}, y_k, (y_{k+1}, \ldots y_N) \middle| s_k, (y_1, \ldots y_{k-1})\right) \\ &= \underbrace{P\left(s_k, y_1, \ldots y_{k-1}\right)}_{\alpha_k(s_k)} \underbrace{P\left(s_{k+1}, y_k \middle| s_k\right)}_{\gamma_k(s_k, s_{k+1})} \underbrace{P\left(y_{k+1}, \ldots y_N \middle| s_{k+1}\right)}_{\beta_{k+1}(s_{k+1})} \end{aligned} \tag{2.24}$$

It is easy to see that the term $\alpha_k(s_k)$ can be computed from $\alpha_{k-1}(s_{k-1})$ and $\gamma_{k-1}(s_{k-1}, s_k)$, and therefore it can be extended to a recursive computation:

$$\alpha_k(s_k) = \sum_{s_{k-1} \in S} \alpha_{k-1}(s_{k-1}) \gamma_{k-1}(s_{k-1}, s_k) \tag{2.25}$$

with initial states $\alpha_0(r_0) = 1$ and zero for other states. Likewise, the term $\beta_k(s_k)$ has also a recursive computation formula:

$$\beta_k(s_k) = \sum_{s_{k+1} \in S} \beta_{k+1}(s_{k+1}) \gamma_k(s_k, s_{k+1}) \tag{2.26}$$

with end states $\beta_N(r_0) = 1$ and zeros for other states. The assigned values of $\alpha_0(s)$ and $\beta_N(s)$ are because an encoder always starts at the zeros state and end at the zero state. For all k, the probability of $\alpha_k(s_k)$ and $\beta_k(s_k)$ must be normalized to 1 which means

$$\begin{aligned} \sum_{i=0}^{2^v - 1} \alpha_k(r_i) &= 1 \\ \sum_{i=0}^{2^v - 1} \beta_k(r_i) &= 1 \end{aligned} \tag{2.27}$$

From (2.25) and (2.26), we know that if we have the knowledge of all $\gamma_k(s_k, s_{k+1})$, $k = 0, \ldots, N-1$, then we can compute all $P(s_k, s_{k+1}, \mathbf{y})$, $k = 0, \ldots, N-1$. The

term $\gamma_k(s_k, s_{k+1})$ can be further decomposed into

$$\gamma_k(s_k, s_{k+1}) = P(s_{k+1} \mid s_k) \cdot P(y_k \mid s_k, s_{k+1}) \tag{2.28}$$

With specific state numbers, (2.28) can be written as

$$\gamma_k(s_k = r_i, s_{k+1} = r_j) = \begin{cases} P(\mathbf{a}_k = \mathbf{a}_{i,j}) \cdot P(y_k \mid x_k = x_{i,j}), & (i,j) \in \boldsymbol{\beta} \\ 0, & (i,j) \notin \boldsymbol{\beta} \end{cases} \tag{2.29}$$

where $\mathbf{a}_{i,j}$ represent the set of information bits corresponding to branch transition from state r_i to state r_j. Usually, the information bits are assumed to be independent identically distributed (i.i.d.), which makes $P(\mathbf{a}_k) = \left(\frac{1}{2}\right)^m$. As for $P(y_k \mid x_k = x_{i,j})$ term, how to calculate it is dependent on the system. For the simple system with only one decoder and AWGN channel, this term is simply the pdf. of the noise distribution, i.e.

$$P(y_k \mid x_k) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_k - x_k)^2}{2\sigma^2}\right) \tag{2.30}$$

where σ^2 is the power of the noise. In this thesis, we will propose a turbo equalizer system, and this term will be provided by the equalizer.

Our goal is to compute the APPs $P(a_k = \hat{a}_k \mid \mathbf{y})$. Since we have derived the formula for $P(s_k, s_{k+1} \mid \mathbf{y})$, we can sum the APPs $P(s_k, s_{k+1} \mid \mathbf{y})$ over all branches that correspond to $a_k = \hat{a}_k$. This idea is actually the theorem of total probability **[15]**.

$$\begin{aligned} P(a_k = \hat{a}_k \mid \mathbf{y}) &= \sum_{\forall (i,j) \in \boldsymbol{\beta}: a_{i,j} = a_k} P(s_k = r_i, s_{k+1} = r_j \mid \mathbf{y}) \\ &= \sum_{\forall (i,j) \in \boldsymbol{\beta}: a_{i,j} = a_k} \frac{P(s_k = r_i, s_{k+1} = r_j, \mathbf{y})}{P(\mathbf{y})} \\ &= \sum_{\forall (i,j) \in \boldsymbol{\beta}: a_{i,j} = a_k} \frac{\alpha_k(r_i)\gamma_k(r_i, r_j)\beta_{k+1}(r_j)}{P(\mathbf{y})} \end{aligned} \tag{2.31}$$

Finally, the LLRs of the APPs are

$$L(a_k = \hat{a}_k \mid \mathbf{y}) = \frac{\sum_{\forall (i,j)\in\boldsymbol{\beta}: a_{i,j}=1} \frac{\alpha_k(r_i)\gamma_k(r_i,r_j)\beta_{k+1}(r_j)}{P(\mathbf{y})}}{\sum_{\forall (i,j)\in\boldsymbol{\beta}: a_{i,j}=0} \frac{\alpha_k(r_i)\gamma_k(r_i,r_j)\beta_{k+1}(r_j)}{P(\mathbf{y})}} = \frac{\sum_{\forall (i,j)\in\boldsymbol{\beta}: a_{i,j}=1} \alpha_k(r_i)\gamma_k(r_i,r_j)\beta_{k+1}(r_j)}{\sum_{\forall (i,j)\in\boldsymbol{\beta}: a_{i,j}=0} \alpha_k(r_i)\gamma_k(r_i,r_j)\beta_{k+1}(r_j)} \tag{2.32}$$

The turbo equalizer system will also require the computation of $L(c_k = \hat{c}_k \mid \mathbf{y})$, and this is done similarly as follows,

$$L(c_k = \hat{c}_k \mid \mathbf{y}) = \frac{\sum_{\forall (i,j)\in\boldsymbol{\beta}: c_{i,j}=1} \frac{\alpha_k(r_i)\gamma_k(r_i,r_j)\beta_{k+1}(r_j)}{P(\mathbf{y})}}{\sum_{\forall (i,j)\in\boldsymbol{\beta}: c_{i,j}=0} \frac{\alpha_k(r_i)\gamma_k(r_i,r_j)\beta_{k+1}(r_j)}{P(\mathbf{y})}} = \frac{\sum_{\forall (i,j)\in\boldsymbol{\beta}: c_{i,j}=1} \alpha_k(r_i)\gamma_k(r_i,r_j)\beta_{k+1}(r_j)}{\sum_{\forall (i,j)\in\boldsymbol{\beta}: c_{i,j}=0} \alpha_k(r_i)\gamma_k(r_i,r_j)\beta_{k+1}(r_j)} \tag{2.33}$$

These LLRs will be used in the turbo equalizer system described in Chapter 3.

CHAPTER 3

Turbo Equalization

For wideband systems, because the bandwidth of the transmitted signals is large and often exceeds the coherent bandwidth of the channel, the channels usually have frequency-selective characteristics. The selectivity in frequency is aroused by the multipath fading channel in which transmission paths have different delays and fadings. This channel can be charactered by a tapped-delay line filter. The objective of an equalizer is to eliminate the distortion induced by the frequency-selective channel. Structures of a variety of equalizers are surveyed in this chapter. The algorithms to derive the optimal filter coefficients are also briefly introduced here. At the end of this chapter, a promising equalization scheme incorporated with the turbo principle is introduced.

3.1 Frequency-Selective Channel

Figure 3.1 shows an equivalent discrete baseband system including transmitter and receiver. Here we focus on channel and equalizer only. The system does not include any coding layer which is commonly found in practical communication systems. To add coding layers into this system is straightforward, that is to replace

the information source with the coded bits. Information bits $\mathbf{a}$ in figure 3.1 is mapped into modulated symbols $\mathbf{x}$ according to a chosen modulation. A transmit filter and a receiver filter are used to meet the spectrum requirement and to mitigate the effect of frequency-selective channel. The objective of the equalizer at the receiver is to recover the symbols $\mathbf{x}$ from the output of the receive filter $\mathbf{y}$. The relation between $\mathbf{x}$ and $\mathbf{y}$ is

$$\mathbf{y} = \left(\left(\mathbf{x} * \mathbf{g}_T\right) * \mathbf{h}\right) * \mathbf{g}_R + \mathbf{n} * \mathbf{v} \tag{3.1}$$

where $*$ is the operation of convolution. We define a new channel $\mathbf{h}_T$ to be

$$\mathbf{h}_T = \mathbf{g}_T * \mathbf{h} * \mathbf{g}_R \tag{3.2}$$

and rewrite (3.1) to be

$$\mathbf{y} = \mathbf{h}_T * \mathbf{x} + \mathbf{n}_T \tag{3.3}$$

where $\mathbf{n}_T = \mathbf{n} * \mathbf{v}$. The $\mathbf{h}_T$ represents the total effect of the transmit filter, channel, and the receive filter. It is the channel that equalizers have to cope with, and can be charactered by the tapped-delay line model as shown in figure 3.2. Thus, channels can be described by the impulse response $\left(h_T[0],\ldots,h_T[L]\right)$ where L is the channel order.

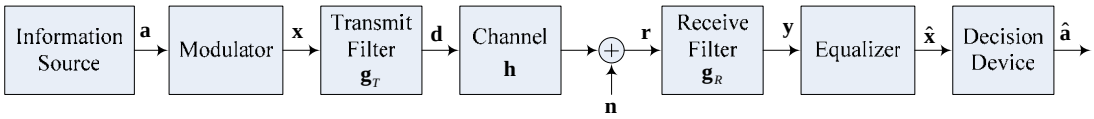

Figure 3.1 Equivalent discrete baseband system

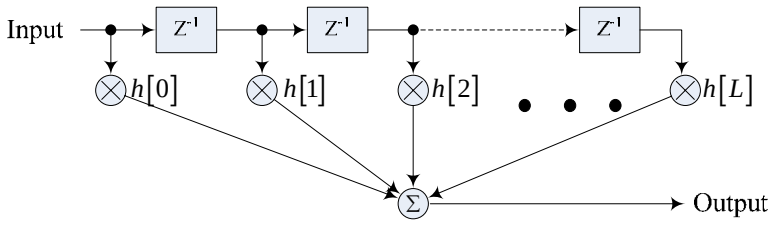

Figure 3.2 Tapped delay line model of the channel

3.2 Equalizer Overview

3.2.1 Trellis-Based

Due to the tapped delay line model of the channel, we can derive a trellis description for the channel. With this trellis description, the BCJR algorithm we presented in 2.4.2 can also be applied readily to detect the symbols. To apply BCJR algorithm in the equalizer, the initial values of $\alpha_0(s_0)$ and $\beta_N(s_N)$ need to be modified according to the property of channels. Usually, the channel starts at the zero state and ends at arbitrary state, which leads to

$$\begin{aligned} \alpha_0(r_i) &= \begin{cases} 1 & ,i=0 \\ 0 & ,i\neq 0 \end{cases} \\ \beta_N(r_i) &= 1 \quad \text{for all } i \end{aligned} \tag{3.4}$$

Or, if the channel was preoccupied by the previous transmission, the value $\alpha_0(s_0)$ will be $\alpha_0(r_i)=1$ for all i.

This approach of equalization requires the knowledge of channel state information (CSI). This requires another effort in channel estimating, and the accuracy of the estimate will affect the detection performance in a certain level.

However, the problem of an trellis-based equalizer is the complexity. The state number of the equivalent trellis is dependent on the channel order L, the signal constellation M-PSK, and transmit antenna n_T in a MIMO system. It can be calculated by

$$\text{state number} = \left(\mathrm{M}^{\mathrm{n_T}}\right)^L \tag{3.5}$$

For example, if the transmitter uses 2 transmit antennas and 4-PSK modulation, and the channel order is 5, then we will have a equivalent trellis of state number $\left(4^2\right)^5 = 1048576$. This is a huge number for a decoder to build, not to mention the one with more antennas or with higher constellation. Therefore, this approach is

only practical when M, L, and n_T are very small.

3.2.2 Filter-Based

Due to the high complexity required by trellis-based equalizer, we turn to a much simpler approach, filter-based equalization. The trellis-based approach does not recover the signals, but it calculates the APPs and then chooses the one with max APPs to be the estimate signals. The filter-based does not calculate the APPs, but instead, it tries to recover the signals of interest and makes decisions on them whether in soft or hard decisions. The structures of filters can be categorized into linear filters or decision-feedback filters.

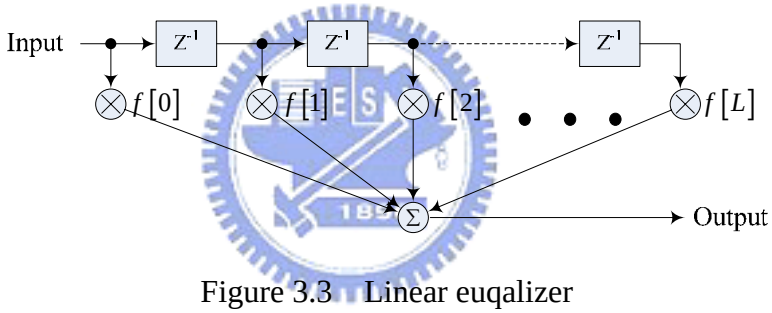

Figure 3.3 Linear euqalizer

3.2.2.1 Linear Equalizer (LE)

To compensate for the channel distortion, we may employ a linear filter with adjustable coefficients $\mathbf{f}$. The filter coefficients are adjusted on the basis of measurements of the channel characteristics. There are different criteria to derive the filter coefficients and will be addressed in 3.2.3. A linear equalizer is shown in figure 3.3. Assuming the filter order is L_f, the linear operation of the equalization can be expressed as

$$\hat{x}_k = \mathbf{f}^H \cdot \mathbf{y}_k \tag{3.6}$$

where $\mathbf{y}_k$ is the received vector $\begin{bmatrix} y_k & \cdots & y_{k-L_f} \end{bmatrix}^T$

The time delay τ between adjacent taps may be selected as large as T_s, the symbol interval, in which case the FIR equalizer is called a symbol-space equalizer. In this case the input to the equalizer is the sampled received sequence at a sampling rate equal to $\frac{1}{T_s}$. On the other hand, when the time delay τ between adjacent taps is selected such that $\frac{1}{\tau} > \frac{1}{T_s}$, the channel equalizer is said to have fractionally spaced taps and it is called a fractionally spaced equalizer. The advantages of a fractionally spaced equalizer are that it provides the function of match filtering and it is less sensitive to symbol sampling timing. The disadvantage is the computation load increase.

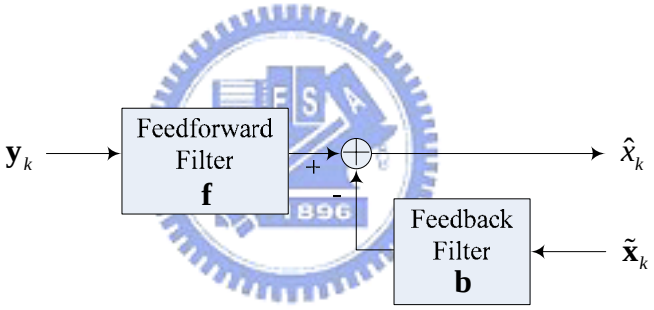

Figure 3.4 Decision-feedback equalizer

3.2.2.2 Decision-Feedback Equalizer (DFE)

The linear filter equalizers described above are very effective on channels where the ISI is not severe. A decision-feedback equalizer is a nonlinear equalizer that employs previous decisions to eliminate the ISI caused by previously detected symbols on the current symbol to be detected, i.e., to eliminate the post-cursor part of ISI. The DFE consists of two filters as shown in figure 3.4. The first filter is called a feedforward filter and is generally a fractionally spaced FIR filter. This filter is identical in structure to the LE described above. The second filter is a feedback filter.

It is also implemented as an FIR filter with symbol-spaced taps $\mathbf{b}$. Its input is the set of previously detected symbols. In a simple system where equalizer process one symbol only once, the input of the feedback filter is usually the hard decision of previous equalized symbols. In the turbo system we proposed in chapter 4, the input to the feedback filter are the estimate symbols produced from the decoder in the previous iteration.

Assuming the feedback filter order is L_b, and the feedforward filter order is L_f, the output of a DFE can be expressed as

$$\hat{x}_k = \mathbf{f}^H \cdot \mathbf{y}_k - \mathbf{b}^H \cdot \bar{\mathbf{x}}_k \tag{3.7}$$

where $\mathbf{y}_k$ is the received vector $\begin{bmatrix} y_k & \cdots & y_{k-L_f} \end{bmatrix}^T$ and $\bar{\mathbf{x}}_k$ is the feedback filter input of the form $\begin{bmatrix} \bar{x}_k & \cdots & \bar{x}_{k-L_b} \end{bmatrix}^T$. Since DFE was introduced by Austin in 1967, it has received considerable attention from many researchers due to its improved performance over the linear equalizer and reduced implementation complexity as compared to the optimal trellis-based equalizer we mentioned in 3.2.1. However, due to the feedback of previously detected symbols, a DFE suffers from error propagation. Especially when SNR is low, the previously detected symbols are erroneous, and thus the DFE under this condition may not outperform the LE.

3.2.3 Filter Design Algorithm

Filtering theorem is well-established and has a rich history. Therefore, there are lots of algorithms and criteria to determine the filter coefficients. The zero-forcing criterion and the minimum mean square error criterion are the most famous two.

3.2.3.1 Zero-Forcing (ZF)

The ideal of zero-forcing equalizer is simple and straightforward. It is to compensate for the channel distortion and ignore all other interference including noise. The optimal coefficients for an infinite length LE are the samples of the inverse filter of channel. The output of a linear equalizer can be expressed as

$$\begin{aligned}\hat{x}[n] &= f[n] * y[n] \\ &= f[n] * h_T[n] * x[n]\end{aligned} \tag{3.8}$$

which, to be equal to $x[n]$, requires

$$f[n] * h_T[n] = \delta[n] \tag{3.9}$$

where $\delta[n]$ is the unit impulse function. Applying the z-transform to (3.9) gives the z-transform of the filter coefficient:

$$F(z) = \frac{1}{H_T(z)} \tag{3.10}$$

For most cases, $f[n]$ are insignificant for large n. Thus we can set the coefficients, $f[n]$, by the long division of $\frac{1}{H_T(z)}$ which leads to a finite impulse response (FIR) filter structure. The long division formula shown below requires little computation for small n:

$$f[n] = \frac{\sum_{i=0}^{n-1}\left(-f[i] \times h_T[n-i]\right)}{h_T[0]} \tag{3.11}$$

The problem of zero-forcing criterion is the noise enhancement. From equation (3.10), if the frequency response of the channel is small or even null at some frequency, then the zero-forcing equalizer compensates it by placing a large gain at that frequency. Consequently, the noise at that frequency is greatly enhanced, too.

3.2.3.2 Minimum Mean Square Error (MMSE)

The MMSE criterion is to minimize the mean-square-error (MSE) between the

actual equalizer output $\hat{x}_k$ and the desired value x_k, i.e.,

$$MSE = E\left[\left|e_k\right|^2\right] \tag{3.12}$$

where $e_k = x_k - \hat{x}_k$. To minimize the mean-square-error, we take the gradient of MSE and find its root. Assuming a linear filter is applied, which means $\hat{x}_k = \mathbf{f}^H \cdot \mathbf{y}_k = \sum_{j=0}^{L_f} f_j^* y_{k-j}$, the gradient of (3.12) with respect to f_i is

$$\begin{aligned} \nabla_i J &= -2E\left[y_{k-i} e_k^*\right] \\ &= -2E\left[y_{k-i}\left(x_k^* - \sum_{j=0}^{L_f} f_j y_{k-j}^*\right)\right] \end{aligned} \tag{3.13}$$

To find the roots, $\nabla_i J$ must be zeros for all $i = 0, \ldots, L_f$ which leads to the Winer-Hopf equations

$$\begin{aligned} E\left[y_{k-i}\left(x_k^* - \sum_{j=0}^{L_f} f_j y_{k-j}^*\right)\right] &= 0 \\ \sum_{j=0}^{L_f} f_j E\left[y_{k-i} y_{k-j}^*\right] &= E\left[y_{k-i} x_k^*\right] \end{aligned} \tag{3.14}$$

Or in matrix form:

$$\begin{aligned} \underbrace{E\left[\mathbf{y}_k \mathbf{y}_k^H\right]}_{\searrow} \cdot \underbrace{\mathbf{f}}_{\downarrow} &= \underbrace{E\left[\mathbf{y}_k \cdot x_k^*\right]}_{\swarrow} \\ \mathbf{R} \cdot \mathbf{f} &= \mathbf{p} \end{aligned} \tag{3.15}$$

The MMSE solution is the solution to the Winer-Hopf equation (3.15), and the filter coefficients are

$$\mathbf{f} = \mathbf{R}^{-1}\mathbf{p} \tag{3.16}$$

$\mathbf{R}$ and $\mathbf{p}$ are called the autocorrelation matrix of $\mathbf{y}_k$ and the crosscorrelation matrix between $\mathbf{y}_k$ and $\overline{\mathbf{x}}_k$ respectively. For decision-feedback filters where the output of the equalizer can be expressed as

$$\hat{x}_k = \mathbf{f}^H \cdot \mathbf{y}_k - \mathbf{b}^H \cdot \overline{\mathbf{x}}_k = \begin{bmatrix} \mathbf{f} \\ -\mathbf{b} \end{bmatrix}^H \cdot \begin{bmatrix} \mathbf{y}_k \\ \overline{\mathbf{x}}_k \end{bmatrix} \tag{3.17}$$

Equation (3.16) can be applied readily with some modifications

$$\begin{bmatrix} \mathbf{f} \\ -\mathbf{b} \end{bmatrix} = \begin{pmatrix} E\left[\mathbf{y}_k \mathbf{y}_k{}^H\right] & E\left[\mathbf{y}_k \overline{\mathbf{x}}_k{}^H\right] \\ E\left[\overline{\mathbf{x}}_k \mathbf{y}_k{}^H\right] & E\left[\overline{\mathbf{x}}_k \overline{\mathbf{x}}_k{}^H\right] \end{pmatrix}^{-1} E\begin{bmatrix} \mathbf{y}_k \cdot x_k^* \\ \overline{\mathbf{x}}_k \cdot x_k^* \end{bmatrix} \tag{3.18}$$

The minimum-mean-square-error criterion is considered to be superior to the zero-forcing criterion. However, it requires the knowledge of the statistics $\mathbf{R}$ and $\mathbf{p}$ which are normally unknown. Also, even the statistics are know, the direct computation of (3.16) is heavy, too.

3.3 Adaptive Equalizer

The Wiener-Hopf solution can be found by a recursive method known as the method of steepest descent. Under the appropriate conditions, the solution obtained by the method of steepest descent will converges to the Wiener solution without the need to invert the correlation matrix of the input vector. However, it still requires the knowledge of the statistics. Using instantaneous estimates of these statistics, we obtain a simple but effective algorithm to approach the Wiener solution. This algorithm is called “least mean square algorithm”.

3.3.1 Least Mean Square Algorithm

Define a cost function $J(\mathbf{f})$ to be the mean square error $E\left[|e_k|^2\right]$ as a function of filter coefficients $\mathbf{f}$. The ideal of the method of steep decent is to adjust the coefficient $\mathbf{f}$ in the direction of steepest decent, that is, in a direction opposite to the gradient vector of the cost function $J(\mathbf{f})$, which is denoted by $\nabla J(\mathbf{f})$. Accordingly, the steepest decent algorithm is formally described by

$$\mathbf{f}(n+1) = \mathbf{f}(n) - \frac{1}{2}\mu \cdot \nabla J\left(\mathbf{f}(n)\right) \tag{3.19}$$

Where n denotes the iteration, μ is a positive constant called the step-size, and the

factor $\frac{1}{2}$ is introduced for mathematical convenience. The gradient $\nabla J\left(\mathbf{f}\left(n\right)\right)$ can be derived as

$$\nabla J\left(\mathbf{f}\left(n\right)\right) = -2\mathbf{p} + 2\mathbf{R}\mathbf{f}\left(n\right) \tag{3.20}$$

Accordingly, (3.19) is now calculated as

$$\mathbf{f}\left(n+1\right) = \mathbf{f}\left(n\right) + \mu \cdot \left(\mathbf{p} - \mathbf{R}\mathbf{f}\left(n\right)\right) \tag{3.21}$$

To compute $\mathbf{f}$ by using (3.21) still requires the knowledge of $\mathbf{p}$ and $\mathbf{R}$. If we discard the actual statistics and use instantaneous estimates of $\mathbf{p}$ and $\mathbf{R}$ by

$$\begin{aligned} \hat{\mathbf{R}}\left(n\right) &= \mathbf{y}\left(n\right)\mathbf{y}^{\mathrm{H}}\left(n\right) \\ \hat{\mathbf{p}}\left(n\right) &= \mathbf{y}\left(n\right)x^{*}\left(n\right) \end{aligned} \tag{3.22}$$

The gradient $\nabla J\left(\mathbf{f}\left(n\right)\right)$ becomes

$$\hat{\nabla} J\left(\mathbf{f}\left(n\right)\right) = -2\mathbf{y}\left(n\right)\left(x^{*}\left(n\right) - \mathbf{y}^{H}\left(n\right)\hat{\mathbf{f}}\left(n\right)\right) \tag{3.23}$$

where $x^{*}\left(n\right) - \mathbf{y}^{H}\left(n\right)\hat{\mathbf{f}}\left(n\right)$ is recognized as $e^{*}\left(n\right)$. Here we obtain another recursive algorithm known as least mean square (LMS) algorithm

$$\hat{\mathbf{f}}\left(n+1\right) = \hat{\mathbf{f}}\left(n\right) + \mu\mathbf{y}\left(n\right)e^{*}\left(n\right) \tag{3.24}$$

Figure 3.5 shows the signal-flow graph representation of the LMS algorithm.

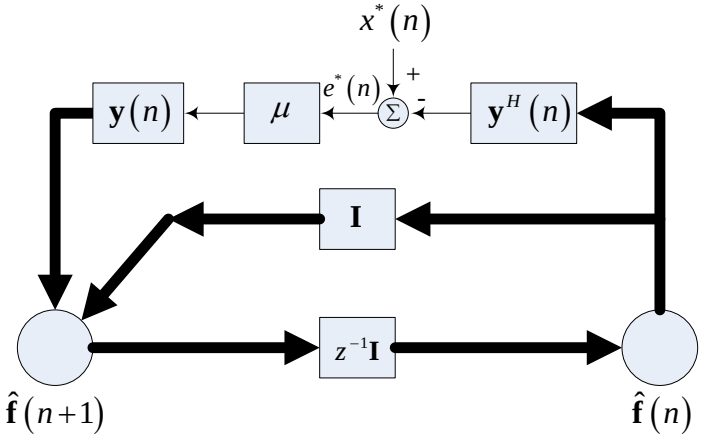

Figure 3.5 Signal-flow graph representation of the LMS algorithm

The stability analysis [3] shows that a necessary condition for the LMS algorithm to converge to Wiener-Hopf solution is

$$0 < \mu < \frac{2}{\lambda_{\max}} \tag{3.25}$$

where $\lambda_{\max}$ is the largest eigenvalue of the correlation matrix $\mathbf{R}$. The convergence of LMS algorithm is in the sense of convergence in the mean square, i.e., there is always a misadjustment exists. The convergent rate and the value of misadjustment are highly related to the step size. In the rage of (3.25), a larger step size results in a faster convergent rate but a large misadjustment. On the other hand, a smaller step size results in a slower convergent rate but a smaller misadjustment.

3.3.2 Adaptive Decision Feedback Equalization

The adaptation of an equalizer with feedforward and feedback filters can be obtained similarly. The equalizer output can be expressed in matrix form

$$\hat{x}_k = \mathbf{f}^H \cdot \mathbf{y}_k - \mathbf{b}^H \cdot \overline{\mathbf{x}}_k = \begin{bmatrix} \mathbf{f} \\ -\mathbf{b} \end{bmatrix}^H \cdot \begin{bmatrix} \mathbf{y}_k \\ \overline{\mathbf{x}}_k \end{bmatrix} \tag{3.26}$$

Thus the LMS algorithm for such system is

$$\begin{bmatrix} \hat{\mathbf{f}}(n+1) \\ -\hat{\mathbf{b}}(n+1) \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{f}}(n) \\ -\hat{\mathbf{b}}(n) \end{bmatrix} + \mu \begin{bmatrix} \mathbf{y}(n) \\ \overline{\mathbf{x}}(n) \end{bmatrix} e^*(n) \tag{3.27}$$

3.4 Equalization and Decoding

For digital communication, the information bits are protected by channel codes, and the coded bits are distorted by channel and noise. The receiver's ultimate goal is to recover the information bits. Due to the multi-layer structure, there are different structures for the receiver to achieve the ultimate goal. Figure 3.6 shows the transmitter for such system. The interleaver is placed to boost the coding gain.

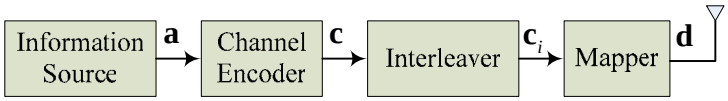

Figure 3.6 Transmitter with encoder and interleaver

3.4.1 Optimal Joint Equalization and Decoding

An optimal detector must solve the channel equalization and decoding problems at the same time in order to decide a sequence of information bits that is the most probable sequence. If we concatenate the tapped delay line model of the channel with the encoder shift register, we obtain a new trellis description whose is formed by combining two convolution expressions and of the enormous size.. The optimal detector can be obtained by sequence decoding the trellis. The maximum likelihood sequence detection of this large-sized trellis is impractical due to high computation complexity. The state number of the trellis is exponentially increasing with the sum of channel order and encoder order.

3.4.2 Separate Equalization and Decoding

The traditional way of equalization and decoding is to split them into two blocks as shown in figure 3.7. The equalizer takes care of only the channel compensation, and the channel decoder handles only the channel code. Performance of such structure is strongly sensitive to the error of the equalization but the equalizer gets no help from the decoder.

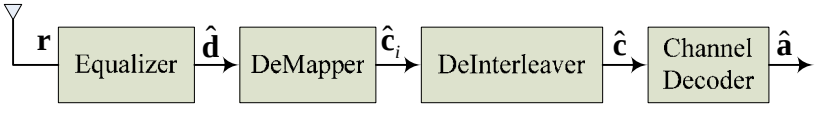

Figure 3.7 Separate equalization and decoder

3.4.3 Iteratively Equalization and Decoding

The rich research of turbo codes has provided substantial insights into the tremendous performance improvement achieved in an iterative manner. With this success in turbo codes, the iterative process has been applied in the equalization/decoding and been referred to as “turbo equalizer”.

The turbo equalizer was first proposed by Douillard et al. in 1995 **[21]**. The ideal of turbo equalization is to exchange information about the subjects between the two blocks: equalizer and decoder so that each block can benefit from the other block. Through the iterative process, each block will receive more and more reliable a prior information as long as the feedback information is not too erroneous. Figure 3.8 shows the block diagram in turbo equalization. The a priori information we mentioned is usually the log-likelihood ratio (LLR) of the subject. A block uses this a priori information and/or local observations to calculate its output which is also log-likelihood ratio (LLR) but referred to as a posteriori probabilities as shown in figure 3.9 which uses the equalizer block as an example.

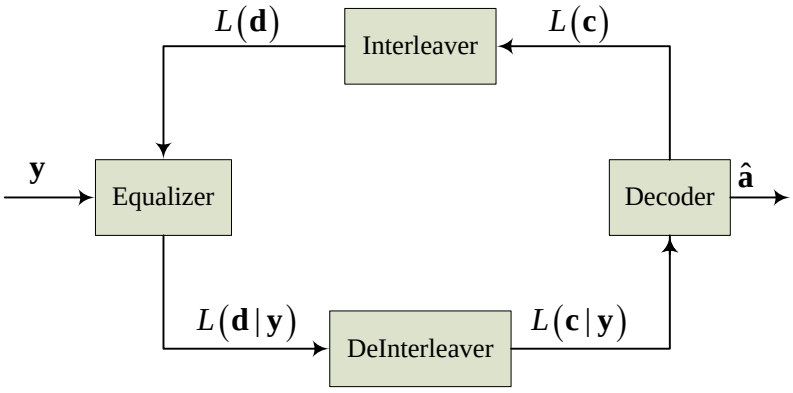

Figure 3.8 Block diagram of turbo equalizer

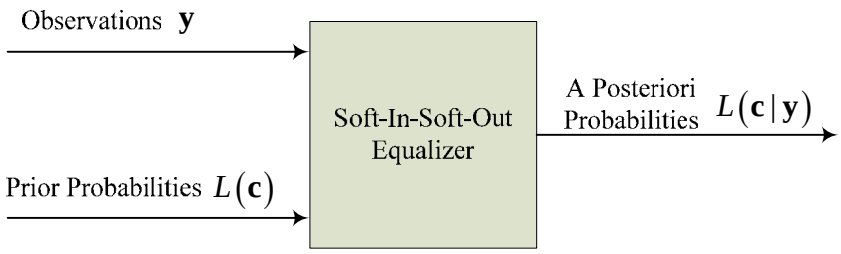

Figure 3.9 Equalizer block in turbo equalizer

To avoid creating direct feedback that is too strong, only extrinsic information is fed from one block to the other block. Using the notation in figure 3.9, the equalizer calculates the a posteriori probabilities $L(c_k|\mathbf{y})$ as

$$\begin{aligned} L(c_k|\mathbf{y}) &= \ln \frac{\sum_{\forall \mathbf{c}:c_k=1} P(\mathbf{y}|\mathbf{c})\cdot P(\mathbf{c})}{\sum_{\forall \mathbf{c}:c_k=0} P(\mathbf{y}|\mathbf{c})\cdot P(\mathbf{c})} \\ &= \ln \frac{\sum_{\forall \mathbf{c}:c_k=1} P(\mathbf{y}|\mathbf{c})\cdot \prod_{i=1}^{K} P(c_i)}{\sum_{\forall \mathbf{c}:c_k=0} P(\mathbf{y}|\mathbf{c})\cdot \prod_{i=1}^{K} P(c_i)} \\ &= \ln \frac{\sum_{\forall \mathbf{a}:a_k=1} P(\mathbf{y}|\mathbf{a})\cdot \prod_{i=1:i\neq k}^{K} P(a_i)}{\sum_{\forall \mathbf{a}:a_k=0} P(\mathbf{y}|\mathbf{a})\cdot \prod_{i=1:i\neq k}^{K} P(a_i)} + L(a_k) \\ &= L_{ext}(a_k|\mathbf{y}) + L(a_k) \end{aligned} \tag{3.28}$$

The $L_{ext}(c_k|\mathbf{y})$ is referred to as the extrinsic information that is to be fed into the decoder block. Therefore, before any a priori information is fed into the decoder block, the extrinsic information must be carried out by

$$L_{ext}(c_k|\mathbf{y}) = L(c_k|\mathbf{y}) - L(c_k) \tag{3.29}$$

while $L(c_k)$ is the input to the equalizer. The interleaver and the deinterleaver are necessities in a turbo system because they help to further disperse the direct feedback

effect. In general, the forward/backward algorithm creates outputs that are locally highly correlated. The correlations between neighboring symbols can be largely suppressed by the use of an interleaver.

There are different ways to take advantages of the a priori information. The straightforward approach to turbo equalization uses a soft-input soft-output equalizer based on the BCJR algorithm, but the computational complexity of this equalizer is too large. This motivated the development of reduce-complexity alternatives to the BCJR equalizer, such as soft interference cancellers **[21]**-**[23]**. In **[23]**, Tüchler develop a turbo-equalization based on linear filtering. The filter coefficients are derived to minimize the mean-squared error (MMSE) assuming the perfect channel state information is known. The a priori information fed to the equalizer is used to calculate the statistical values required in the MMSE solution. Tüchler also proposed a decision-feedback equalizer in **[23]** but it feeds back hard decisions on the equalizer output, without combining them with the a priori information. The coefficients of the MMSE equalizer have to be computed anew for every symbol, even when the channel is static.

It has been shown **[19][20]** that under a SISO channel, the decision-feedback equalizer with anti-causal part in the feedback filter can eliminate the ISI completely if perfect channel state information and ideal data feedback are available. However, this assumption is impractical to traditional systems due to the anti-causal part. Nevertheless, for turbo equalizer systems, anti-causal part is no longer impractical after the first iteration. Lopes developed a turbo-equalization based on decision-feedback filtering. Instead of feeding back the hard decisions on the equalizer output, it feeds back the mean values derived from the a priori information which are more accurate. In this structure, the feedback filters are allowed to have anti-causal part because of the iterative process. The coefficients of the equalizer are

also computed according to MMSE criterion but do not have to be recomputed every symbol. By adopting a simple statistical model for the equalizer outputs and a priori information, a time-invariant MMSE equalizer is thus obtained. However, the channel states information is still required.

Glavieux, Loat and Labat proposed an adaptive decision-feedback turbo equalizer **[21]** which does not require the knowledge of channel state information. It applied the adaptive algorithm directly to the filter coefficients. The computational complexity of this adaptive equalization is relatively low compared to the ones where exact MMSE solutions are computed. The feedback filter also has the anti-causal part and the inputs to the feedback filter are the mean values derived from the a priori information. This equalizer structure is shown by simulations to be effective in combating multi-path effect.

In this thesis, we extend the equalizer proposed by Glavieux, Loat and Labat to be a multiple-input multiple-output version in order to combat the multi-path effect in a MIMO channel. The advantage of this equalizer structure is that it does not require the knowledge of channel state information nor the matrix inversion, which is computation-demanding. The computational complexity of equalization for MIMO channel is multiples of that for SISO channel. Therefore, the demand of a low-complexity turbo equalization algorithm is important especially to a practical MIMO receiver.

CHAPTER 4

Adaptive Filter-Based Turbo Equalizer with Space-Time Decoder

In this chapter, we combine the techniques in the previous chapters and propose a turbo equalizer which utilizes the adaptive filtering technique and the space-time trellis coded system in wideband MIMO systems. We will describe how we transmit information bits to the antennas and how we recover the information bits from severely distorted received signals. The proposed receiver is composed of a low-complexity equalizer, but the good performance is still preserved due to the turbo process. The performance evaluation of the proposed receiver is carried out with different parameters, and the perfect-feedback case is used as reference in the figures. These comparisons provide trade-off between complexity and performance. In addition, they reveal some interesting properties of a space-time coded system with wideband transmission.

4.1 Transmitter

A space-time trellis coded transmitter with n_T transmit antennas is shown in figure 4.1. Assume the information bits in the sequence **a** are independent

identically distributed (i.i.d.). The information bit sequence is encoded into n_T coded bit sequences $\mathbf{c}_i$,$i=1,\ldots,n_T$, and the coded bit sequences are then mapped into coded symbol sequences $\mathbf{s}_i$,$i=1,\ldots,n_T$. Interleavers π_i ,$i=1,\ldots,n_T$ are used to shuffle coded symbol sequences $\mathbf{s}_i$,$i=1,\ldots,n_T$ respectively to obtain new sequences $\mathbf{d}_i$,$i=1,\ldots,n_T$. Before symbols are transmitted by antennas, pulse-shaping filters are applied to mitigate the ISI effect. A common choice of the pulse-shaping filter is the squared root raised cosine filter, which will be addressed later.

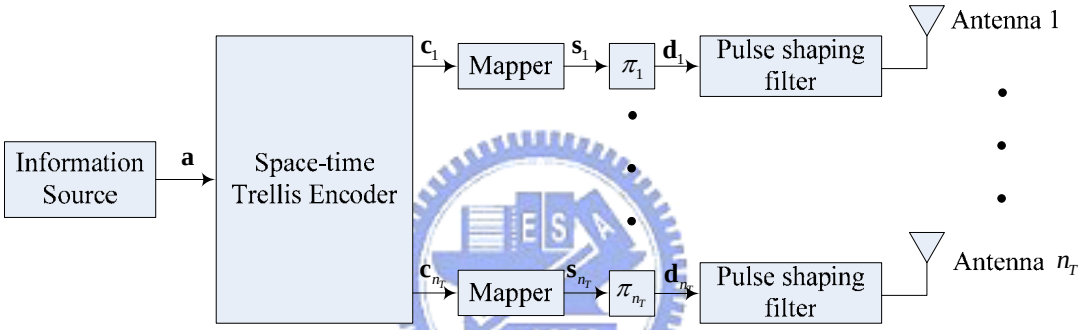

Figure 4.1 Transmitter of a space-time trellis coded system

4.1.1 Space-Time Trellis Encoder

We have introduced the space-time trellis codes in 2.3. Here we select some of these codes in our systems. In this thesis, the space-time trellis codes we choose are all 4-PSK modulated whose signal constellation is shown in figure 2.7. In 4.4.3, we will compare the performance between 2, 3, and 4 transmit antennas. In 4.4.5, we will compare the performance between 4, 8, 16, 32 and 64 states in the encoder. Therefore, to carry out these simulations and comparisons, different codes are adopted for different simulations. Applying the notation in figure 2.8, we list the required codes described by generator sequences in table 4.1 and table 4.2. When the effect of transmit antenna number is concerned, table 4.1 provides codes of 2, 3 and 4

antennas with same state number. When the effect of state number is concerned, table 4.2 provides codes with 4, 8, 16, 32 and 64 states with 2 transmit antennas. In the coding theorem, a code with more states in the trellis results in better performance at the expense of extra computation complexity in the decoder. Moreover, more transmit antennas provide more diversity to the receiver, but the signal detection is more difficult. How to choose the number of transmit antennas and state number depends on the system designer. Simulations with respect to these parameters provide a powerful tool to determine these parameters.

n_T	v	generator sequences
2	5	$\mathbf{g}^1=[(0,2),(2,3),(1,2)]$ $\mathbf{g}^2=[(2,2),(1,2),(2,3),(2,0)]$
3	5	$\mathbf{g}^1=[(0,2,2),(2,3,3),(1,2,2)]$ $\mathbf{g}^2=[(2,2,0),(1,2,2),(2,3,1),(2,0,0)]$
4	5	$\mathbf{g}^1=[(0,2,2,2),(2,3,3,2),(1,2,2,1)]$ $\mathbf{g}^2=[(2,2,0,1),(1,2,2,0),(2,3,1,0),(2,0,0,2)]$

Table 4.1 Generator sequences of 32 states and 2, 3 and 4 transmit antennas

n_T	v	generator sequences
2	2	$\mathbf{g}^1=[(0,2),(1,2)]$ $\mathbf{g}^2=[(2,3),(2,0)]$
2	3	$\mathbf{g}^1=[(2,2),(2,1)]$ $\mathbf{g}^2=[(2,0),(1,2),(0,2)]$
2	4	$\mathbf{g}^1=[(1,2),(1,3),(3,2)]$ $\mathbf{g}^2=[(2,0),(2,2),(2,0)]$

2	5	$\mathbf{g}^1 = \left[(0,2),(2,3),(1,2)\right]$ $\mathbf{g}^2 = \left[(2,2),(1,2),(2,3),(2,0)\right]$
2	6	$\mathbf{g}^1 = \left[(0,2),(3,1),(3,3),(3,2)\right]$ $\mathbf{g}^2 = \left[(2,2),(2,2),(0,0),(2,0)\right]$

Table 4.2 Generator sequences of different state numbers for 2 transmit antennas

Signal power normalization is applied after space-time trellis encoder. We constrain the total transmitted power of all transmit antennas to be 1. To meet the power constraint, each signal must be divided by $sqrt\left(n_T \times 2\right)$.

4.1.2 Interleavers

All trellis codes are sensitive to burst errors. These burst errors are possibly caused by correlated noises and the sequential operations of an equalizer or a decoder. Generally, the noise is assumed to be uncorrelated, i.e. white noise, at the receive antenna. However, the operation of the equalizer will produce a correlated noise, i.e. the colored noise at the equalizer output. In addition, the residual ISI at the equalizer output is a form of correlated interference. These correlated interferences including the colored noise and the residual ISI have an impact on the coding gain and cause performance loss. To relieve the effect of burst errors, an interleaver is usually used. In the receiver, the inverse operation, i.e. deinterleaver, is placed between equalizer and decoder. The deinterleaver will decorrelate the signals at the decoder input and thus avoid the performance loss cause by burst errors.

In a space-time coded system, the correlated interference not only appears in time domain but also in space domain. A space time equalizer outputs several sequences at the same time representing the estimated symbols for different antennas. The

cross filters in the equalizer introduce correlation between different output sequences and thus result in correlated interference in space domain. To handle this kind of correlation, different interleavers are applied for different antennas.

As mentioned in 3.4.3, interleavers plays a crucial role in a turbo system. Because of the iterative process, it is important to remove the correlation at the output of each block. Otherwise, with the iteration increases, the correlation becomes bigger and results in larger performance loss. In the turbo equalizer receiver, the deinterleaver removes the correlation at the output of the equalizer while the interleaver removes the correlation at the output of the decoder.

In this thesis, we set the interleaver length to be 4096 and assign different permutation tables for different antennas. The i-th element in the permutation table is j, and we denote this as a pair (i, j) ,$i = 1,\ldots,4096,\ \mathrm{j} = 1,\ldots,4096$. The i-th input of the interleaver will be the j-th output of the interleaver. On the other hand, the function of a deinterleaver according to a permutation table (i, j) ,$i = 1,\ldots,4096,\ \mathrm{j} = 1,\ldots,4096$ is to put the j-th input in the i-th output.

4.1.3 Pulse Shaping Filter

The raised cosine filter is a simple spectrum shaping filter. The frequency response consists of a flat portion and a rolloff portion that has a sinusoidal form. Define a parameter called the rolloff factor to be

$$\alpha = 1 - \frac{f_r}{W} \tag{4.1}$$

where W is the Nyquist bandwidth of the signal and f_r is the rolloff frequency. The rolloff factor also indicates the excess bandwidth over the Nyquist bandwidth W. Specifically, the transmission bandwidth B_T is defined by

$$\begin{aligned} B_T &= 2W - f_r \\ &= W(1+\alpha) \end{aligned} \tag{4.2}$$

Instead of using one raised cosine filter to meet the Nyquist criterion, we can split it into two filters each of which is a squared root raised cosine (SRRC) filter. By placing one in the transmitter before antenna, and one in the receiver after antenna, we can obtain a raised cosine response.

Define the symbol rate at the output of the interleavers to be R_s. To design the SRRC filters, we first upsample the input by 2 to be $2R_s$. Assigning the rolloff factor to be 0.5, and truncating the filter coefficients to be of length 13, we obtain the normalized SRRC filter of the impulse response shown in figure 4.2. The coefficients are normalized so that the signal power will not be amplified by this filter.

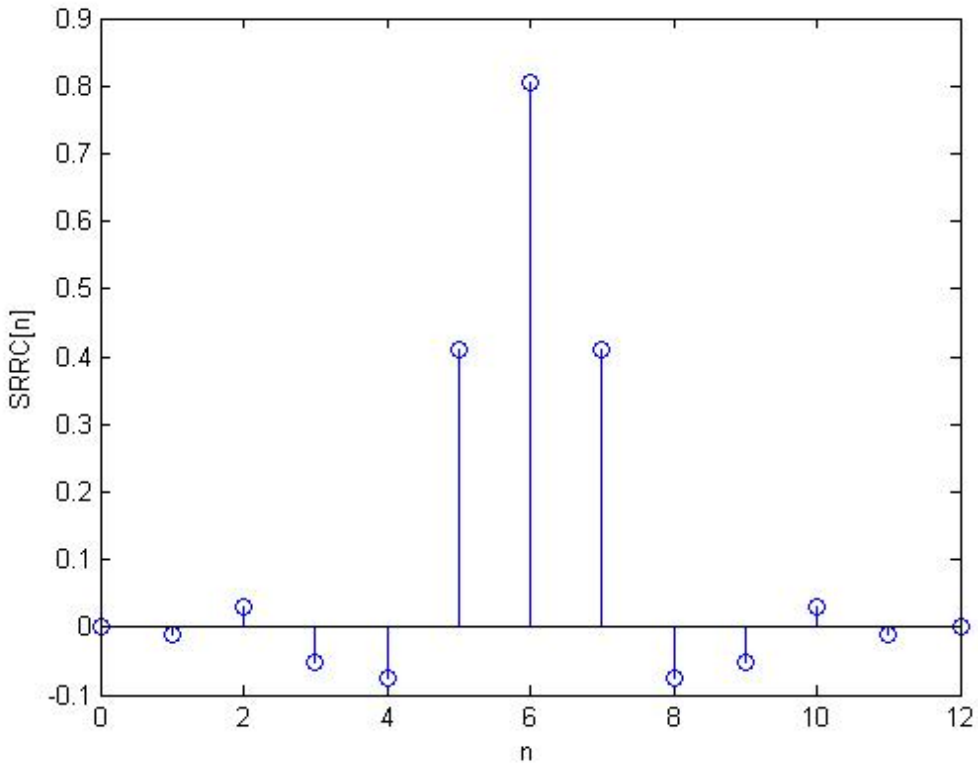

Figure 4.2 Squared rooted raised cosine filter with rolloff factor 0.5 and truncated to be length 13

4.1.4 Training Symbols

We insert training symbols before information data is transmitted. The purpose of these training symbols is to train the equalizer coefficients. The packet format is shown in figure 4.3. The training mode will be described latter along with the equalization. We set the training length to be 4096 symbols in the following simulations.

Figure 4.3 Packet format

4.2 Channel and Noise

In the following simulations, we assume there are 10 multipaths between every transmit antenna and receive antenna. Multipaths are separated by 0.5 symbol period, which results in channel length of five symbols duration. The fading gain of each multipath is a complex random variable of normal distribution with the mean 0 and the variance 1. As mentioned in 1.1, channels must be normalized so that the signal power at each receive antenna is equal to the total signal power of all transmit antennas. To do so, we modified the channel normalization formula to be

$$\sum_{i=1}^{n_T}\sum_{n=0}^{L}\left|h_{ij}[n]\right|^2 = n_T, \quad j = 1,\ldots,n_R \tag{4.3}$$

In most cases, we plot the error rate versus signal-to-noise ratio (SNR) figure to compare the performance. The SNR is calculated as the signal power over noise power. Because of the signal normalization and the channel normalization, the signal power at each receive antenna is 1. The SNR is simply

$$\text{SNR} = \frac{1}{\sigma_n^2} \tag{4.4}$$

where σ_n^2 is the noise power. Noise is assumed to be additive Gaussian white noise at receive antennas.

4.3 Turbo Receiver

The turbo receiver is shown in Figure 4.4. Signals after antennas are sampled at a rate $2R_s$ and then fed into SRRC filters. The SRRC filters are the same square rooted raised cosine filters we placed at the transmitter. The data rates at the SRRC filter output $\mathbf{y}_j$,$j=1,\ldots,n_R$ are $2R_s$. The turbo system is composed of an equalizer block and a decoder block. The equalizer produces the estimated transmitted sequences $\hat{\mathbf{d}}_i$,$i=1,\ldots,n_T$ based on observed sequences from receive antennas $\mathbf{y}_j$,$j=1,\ldots,n_R$ and the mean values $\bar{\mathbf{d}}_i$,$i=1,\ldots,n_T$ from the previous iteration. These estimated sequences $\hat{\mathbf{d}}_i$,$i=1,\ldots,n_T$ are deinterleaved into $\hat{\mathbf{s}}_i$,$i=1,\ldots,n_T$ as the reverse operation of interleavers π_i ,$i=1,\ldots,n_T$ in the transmitter. After the deinterleavers, these sequences are transformed in to log-likelihood ratios of coded bits $L(\mathbf{c}_i \mid \mathbf{y})$,$i=1,\ldots,n_T$ by the demappers. The space-time trellis decoder takes the estimated log-likelihood ratios of coded bits and produces the new log-likelihood ratios of these coded bits $L(\mathbf{c}_i)$,$i=1,\ldots,n_T$ along with decisions of information bits $\hat{\mathbf{a}}$. Then in the next iteration, these LLRs $L(\mathbf{c}_i)$,$i=1,\ldots,n_T$ are passed through mappers and interleavers to provide the mean values $\bar{\mathbf{d}}_i$,$i=1,\ldots,n_T$ to the equalizer. The detail of these blocks will be address in the followings.

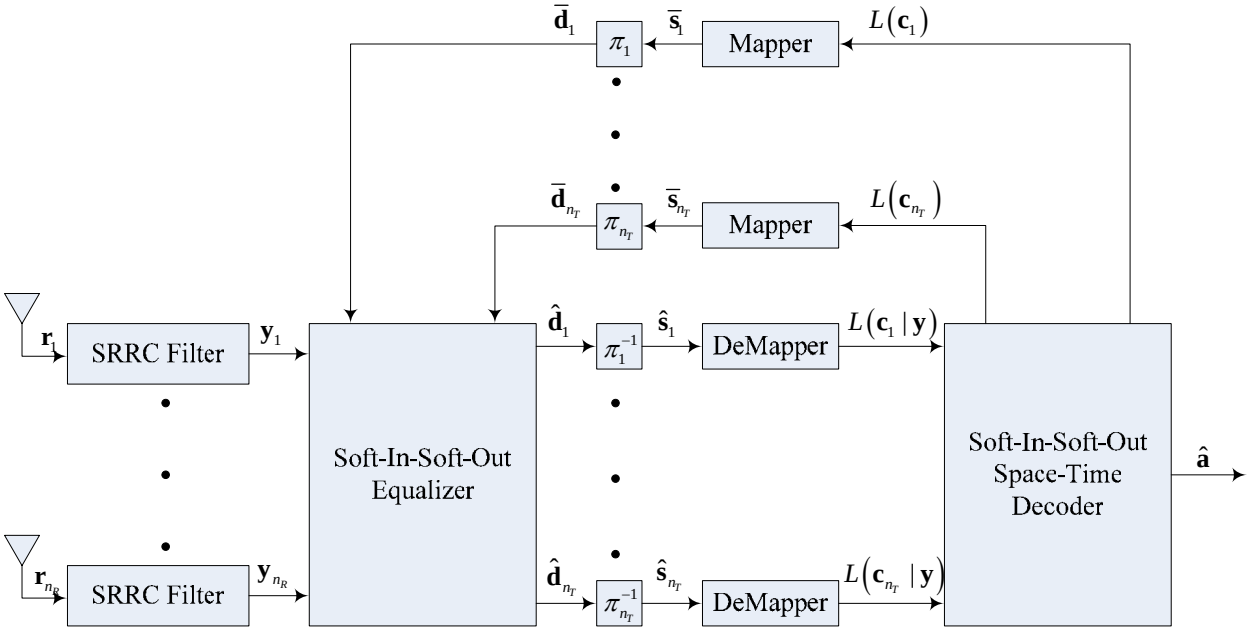

Figure 4.4 Transmitter of a space-time trellis coded system

4.4 Equalization

The aim of the equalizer in this system is to recover the n_T sequences transmitted from the n_T transmit antennas. Figure 4.5 shows the block diagram of the equalizer. The feedforward filter $\mathbf{f}$ is a fractionally-spaced linear equalizer whose input data rate is $2R_s$ and output data rate is R_s. The feedback filter $\mathbf{b}$ is a decision-feedback equalizer whose input is the mean values $\bar{\mathbf{x}}_i$,$i=1,\ldots,n_T$ evaluated from the log-likelihood ratios $L(\mathbf{d}_i)$,$i=1,\ldots,n_T$.

For the first iteration, because there is no available a priori information, only feedforward filter will be applied. In this case, it performs a linear equalization operation. For the other iterations, mean values $\bar{\mathbf{x}}_i$,$i=1,\ldots,n_T$ severed as a priori information are available; therefore it performs a feedback-cancelling equalization operation. The coefficients of $\mathbf{f}$ and $\mathbf{b}$ are obtained by an adaptive algorithm

called least-mean-square (LMS) we introduced in 3.3.

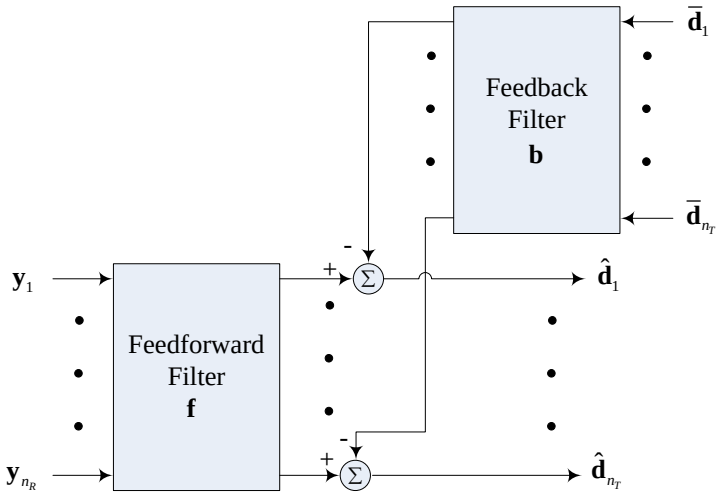

Figure 4.5 Equalizer block diagram of a space-time trellis coded system

We define the length of each sub filter in feedforward part to be $L_f = 2L_1 + 1$ and the length of the each sub filter in feedback part to be $L_b = 2L_2 + 1$. In the following simulations, we assign these parameters as $L_1 = 20,\ L_2 = 20$.

4.4.1 Single Receive Antenna

Let us consider the case where only a single receive antenna is applied. The equalizer is shown in figure 4.6. The equalizer output is given by

$$\hat{d}_i[k] = \mathbf{f}_i^H \mathbf{y}'_{1,k} - \sum_{j=1}^{n_T} \mathbf{b}_{i,j}^H \bar{\mathbf{d}}'_{j,k} \quad ,i = 1,\ldots,n_T \tag{4.5}$$

where $\mathbf{y}'_{1,k} = \left[y_1[2k + D + L_1] \quad \cdots \quad y_1[2k + D] \quad \cdots \quad y_1[2k + D - L_1] \right]^T$ and

$\bar{\mathbf{d}}'_{j,k} = \left[\bar{d}_j[k + L_2] \cdots \quad \bar{d}_j[k] \quad \cdots \quad \bar{d}_j[k - L_2] \right]^T$ and D is an appropriate delay.

When trying to recover symbols from a specific transmit antenna, the signals from other antennas become co-channel interference (CCI) to the symbols of interest.

Therefore, to cancel these co-channel interferences, the feedback filters are also applied. The feedback filters $\mathbf{b}_{i,j}$ $,i \neq j$ are used to cancel CCI. The feedback filter $\mathbf{b}_{i,j}$ $,i = j$ is used to cancel ISI of the symbols themselves. Note that $b_{i,j}[0]$ $,i = j$, must be zero to avoid strong positive feedback.

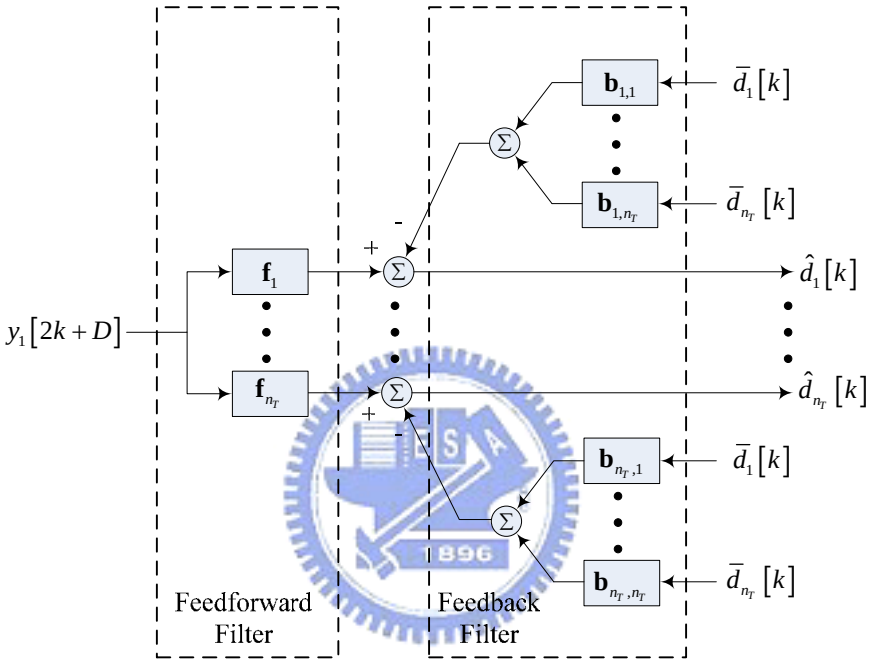

Figure 4.6 Equalizer structure for one receive antenna

The adaptation for these filter coefficients is based on the LMS algorithm. At the beginning, the equalizer is in the training mode in which the desired symbols $d_i[k]$ $,i=1,\ldots,n_T$ are known. Using these known training symbols to calculate the error in the LMS algorithm will result in correct adaptation. Therefore, it is expected that the solution will reach steady state by the time training mode is over. After the training mode is over, the filters are supposed to be close to the Wiener solution which minimizes the mean squared error $E\left\{\left|d_i[k]-\hat{d}_i[k]\right|^2\right\}$. Therefore, the data symbols

can be equalized by using simply the LMS algorithm.

During the data symbols, the desired symbols $d_i[k]$ are no longer available so that we must search for other signals for computation. There are two options available: the tentative decisions $\tilde{d}_i[k]$ at the equalizer output, and the mean values $\bar{d}_i[k]$ obtained from the pervious iteration. Generally, $\bar{d}_i[k]$ are more correct than the tentative decisions $\tilde{d}_i[k]$. Therefore we choose $\bar{d}_i[k]$ to calculate the error in the LMS algorithm during data symbols. However, for the first iteration, the mean values $\bar{d}_i[k]$ are not available. We can only choose the tentative decisions $\tilde{d}_i[k]$.

The step size is highly related to the convergence rate and the mismatch at steady state. After several experiments, we assign the step size to be 0.001 in the training mode to obtain fast convergence and use the step size 0.00001 for the data symbols to maintain and slightly adjust the coefficients.

With the training symbols length being 4096, the adaptation of these filters is now summarized below for the first iteration and other iterations.

For the first iteration:

$$\begin{aligned}
&\hat{d}_i[k] = \mathbf{f}_{i,k}^H \mathbf{y}_{1,k}^{'} \quad ,i = 1,\ldots,n_T \\
&e_i[k] = \begin{cases} d_i[k] - \hat{d}_i[k] & k \le TrLen \\ \tilde{d}_i[k] - \hat{d}_i[k] & k > TrLen \end{cases} \quad ,i = 1,\ldots,n_T \\
&\mu = \begin{cases} \mu_{Tr} & k \le TrLen \\ \mu_{Data} & k > TrLen \end{cases} \\
&\mathbf{f}_{i,k+1} = \mathbf{f}_{i,k} + \mu \cdot \mathbf{y}_{1,k}^{'} \cdot e_i^*[k] \quad ,i = 1,\ldots,n_T
\end{aligned} \tag{4.6}$$

For the other iterations:

$$\hat{d}_i[k] = \mathbf{f}_{i,k}^H \mathbf{y}'_{1,k} - \sum_{j=1}^{n_T} \mathbf{b}_{i,j,k}^H \bar{\mathbf{d}}'_{j,k} \quad , i = 1, \ldots, n_T$$

$$e_i[k] = \begin{cases} d_i[k] - \hat{d}_i[k] & k \le TrLen \\ \bar{d}_i[k] - \hat{d}_i[k] & k > TrLen \end{cases}$$

$$\mu = \begin{cases} \mu_{Tr} & k \le TrLen \\ \mu_{Data} & k > TrLen \end{cases}$$

$$\mathbf{f}_{i,k+1} = \mathbf{f}_{i,k} + \mu \cdot \mathbf{y}'_{1,k} \cdot e_i^*[k] \quad , i = 1, \ldots, n_T$$

$$\mathbf{b}_{i,j,k+1} = \begin{cases} \mathbf{b}_{i,j,k} - \mu \cdot \mathbf{d}'_{j,k} \cdot e_i^*[k] & k \le TrLen \\ \mathbf{b}_{i,j,k} - \mu \cdot \bar{\mathbf{d}}'_{j,k} \cdot e_i^*[k] & k > TrLen \end{cases} \quad , i = 1, \ldots, n_T \quad , j = 1, \ldots, n_R$$

$$b_{i,j,k+1}[0] = 0 \quad , i = j = 1, \ldots, n_T$$

(4.7)

4.4.2 Two Receive Antennas

In the equalization of multiple transmit antennas system, the co-channel interferences are the main interference and need to be eliminated. Especially when the transmit antenna number is large, the co-channel interferences are far more vital to the system performance than the noise. In the feedforward part, using only single filter to isolate certain symbols from other symbols is difficult. The feedforward part is the only available part at the first iteration. Fortunately, the second receive antenna may ease this problem greatly.

It is well-known that two receive antennas have the ability to remove the interference more completely. The way we take advantages of the extra receive antenna is to add cross filters in the feedforward filter part. The addition of cross filters makes the isolation more successful and results in better performance at the first iteration. The structure of such equalizer is shown in figure 4.7. The feedback filter part is exactly the same as the one in the single receive antenna case.

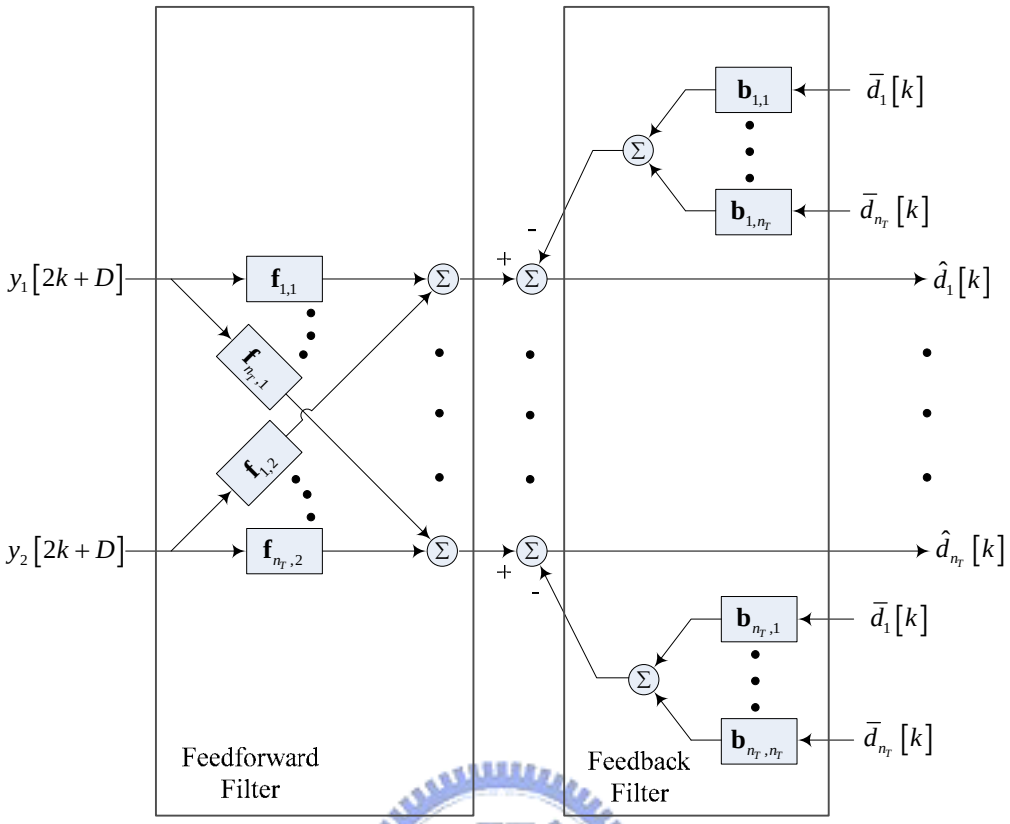

Figure 4.7 Equalizer structure for two receive antennas

The adaptation of such equalizer is similar with additional cross-filter coefficients. We summarize the adaptation blow.

For the first iteration:

$$\hat{d}_i[k] = \sum_{j=1}^{n_R} \mathbf{f}_{i,j,k}^H \mathbf{y}'_{j,k} \quad , i = 1,\ldots,n_T$$

$$e_i[k] = \begin{cases} d_i[k] - \hat{d}_i[k] & k \le TrLen \\ \tilde{d}_i[k] - \hat{d}_i[k] & k > TrLen \end{cases} \quad , i = 1,\ldots,n_T \qquad (4.8)$$

$$\mu = \begin{cases} \mu_{Tr} & k \le TrLen \\ \mu_{Data} & k > TrLen \end{cases}$$

$$\mathbf{f}_{i,j,k+1} = \mathbf{f}_{i,j,k} + \mu \cdot \mathbf{y}'_{j,k} \cdot e_i^*[k] \quad , i = 1,\ldots,n_T \quad , j = 1,\ldots,n_R$$

For the other iterations:

$$\hat{d}_i[k] = \sum_{j=1}^{n_R} \mathbf{f}_{i,j,k}^H \mathbf{y}'_{j,k} - \sum_{j=1}^{n_T} \mathbf{b}_{i,j,k}^H \bar{\mathbf{d}}'_{j,k} \quad , i = 1,\ldots,n_T$$

$$e_i[k] = \begin{cases} d_i[k] - \hat{d}_i[k] & k \le TrLen \\ \bar{d}_i[k] - \hat{d}_i[k] & k > TrLen \end{cases}$$

$$\mu = \begin{cases} \mu_{Tr} & k \le TrLen \\ \mu_{Data} & k > TrLen \end{cases}$$

$$\mathbf{f}_{i,j,k+1} = \mathbf{f}_{i,j,k} + \mu \cdot \mathbf{y}'_{j,k} \cdot e_i^*[k] \quad , i = 1,\ldots,n_T \quad , j = 1,\ldots,n_R$$

$$\mathbf{b}_{i,j,k+1} = \begin{cases} \mathbf{b}_{i,j,k} - \mu \cdot \mathbf{d}'_{j,k} \cdot e_i^*[k] & k \le TrLen \\ \mathbf{b}_{i,j,k} - \mu \cdot \bar{\mathbf{d}}'_{j,k} \cdot e_i^*[k] & k > TrLen \end{cases} \quad , i = 1,\ldots,n_T \quad , j = 1,\ldots,n_R$$

$$b_{i,j,k+1}[0] = 0 \quad , i = j = 1,\ldots,n_T$$

(4.9)

4.5 Mapper and DeMapper

The demapper function transforms the estimated coded symbols into the log-likelihood ratios of the coded bits $L(c_k \mid \mathbf{y})$. The estimated coded symbol $\hat{s}_k$ can be decomposed into two part: desired symbol s_k and the interference w_k.

$$\hat{s}_k = s_k + w_k \tag{4.10}$$

The interference is the combination of filtered noise and residual ISI. Generally, it follows the Gaussian distribution with zero mean and a specific variance. The demapper recovers the signal constellation by using the relation between the coded symbols s_k and the corresponding coded bits. For example, the modulation we used throughout this thesis is 4-PSK whose relation is

$$\begin{aligned} \text{Real}(s_k) &= \begin{cases} \alpha & , c_{2k} = 0 \\ -\alpha & , c_{2k} = 1 \end{cases} \\ \text{Imag}(s_k) &= \begin{cases} \alpha & , c_{2k+1} = 0 \\ -\alpha & , c_{2k+1} = 1 \end{cases} \end{aligned} \tag{4.11}$$

where α is a positive value determined by the signal power normalization. The log-likelihood ratios of c_{2k} can be calculated by

$$
\begin{aligned}
L\left(c_{2k}\,|\,\mathbf{y}\right) &= \ln\frac{P\left(c_{2k}=1|\,\mathbf{y}\right)}{P\left(c_{2k}=0|\,\mathbf{y}\right)} \\
&= \ln\frac{\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2\sigma^2}\left(\mathrm{Real}\left(\hat{s}_k\right)+\alpha\right)^2}}{\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2\sigma^2}\left(\mathrm{Real}\left(\hat{s}_k\right)-\alpha\right)^2}} \\
&= \ln e^{-\frac{1}{2\sigma^2}\left(\left(\mathrm{Real}\left(\hat{s}_k\right)+\alpha\right)^2-\left(\mathrm{Real}\left(\hat{s}_k\right)-\alpha\right)^2\right)} \\
&= \frac{2\mathrm{Real}\left(\hat{s}_k\right)\alpha}{\sigma^2}
\end{aligned}
\tag{4.12}
$$

Likewise, c_{2k+1} can be calculated by

$$
\begin{aligned}
L\left(c_{2k+1}\,|\,\mathbf{y}\right) &= \ln\frac{P\left(c_{2k+1}=1|\,\mathbf{y}\right)}{P\left(c_{2k+1}=0|\,\mathbf{y}\right)} \\
&= \ln\frac{\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2\sigma^2}\left(\mathrm{Imag}\left(\hat{s}_k\right)+\alpha\right)^2}}{\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2\sigma^2}\left(\mathrm{Imag}\left(\hat{s}_k\right)-\alpha\right)^2}} \\
&= \ln e^{-\frac{1}{2\sigma^2}\left(\left(\mathrm{Real}\left(\hat{s}_k\right)+\alpha\right)^2-\left(\mathrm{Imag}\left(\hat{s}_k\right)-\alpha\right)^2\right)} \\
&= \frac{2\mathrm{Imag}\left(\hat{s}_k\right)\alpha}{\sigma^2}
\end{aligned}
\tag{4.13}
$$

The variance σ^2 is related to the filter coefficients, the channel and the noise power at the receive antenna. If we know all these parameter, we can calculate an exact value of σ^2 as in **[22]** for SISO channels. However, it is assumed to be unknown in our simulations, and we assign a constant value to σ^2. The assigned value of σ^2 will affect the values of LLRs calculated by demapper, but the signs of the LLRs is not affected. In other words, it does not affect the correctness of decoding. After several experiment, we assign the value of σ^2 to be 0.5.

The function of mapper in this turbo receiver is to transform the LLRs of coded bits produced by decoder into mean values of coded symbols. Performing the average operation on real part of coded symbols, we obtain

$$\begin{aligned}\mathrm{Real}\left(\bar{s}_k\right) &= \alpha \cdot P\left(c_{2k}=0\right)+\left(-\alpha\right)\cdot P\left(c_{2k}=1\right)\\ &= \alpha \cdot \frac{e^{L(c_{2k})}}{1+e^{L(c_{2k})}}+\left(-\alpha\right)\cdot\frac{1}{1+e^{L(c_{2k})}}\\ &= \alpha \cdot \frac{e^{L(c_{2k})}-1}{1+e^{L(c_{2k})}}\\ &= \alpha \cdot \tanh\left(\frac{L\left(c_{2k}\right)}{2}\right)\end{aligned} \tag{4.14}$$

Likewise, the imaginary part can be calculated as

$$\begin{aligned}\mathrm{Imag}\left(\bar{s}_k\right) &= \alpha \cdot P\left(c_{2k+1}=0\right)+\left(-\alpha\right)\cdot P\left(c_{2k+1}=1\right)\\ &= \alpha \cdot \frac{e^{L(c_{2k+1})}}{1+e^{L(c_{2k+1})}}+\left(-\alpha\right)\cdot\frac{1}{1+e^{L(c_{2k+1})}}\\ &= \alpha \cdot \frac{e^{L(c_{2k+1})}-1}{1+e^{L(c_{2k+1})}}\\ &= \alpha \cdot \tanh\left(\frac{L\left(c_{2k+1}\right)}{2}\right)\end{aligned} \tag{4.15}$$

4.6 Decoding

In our simulations, we choose the BCJR algorithm to implement our decoder. The decoding algorithm has already been described in 2.4.2. The codewords are correctly assigned so that the state starts and ends at the zero state. The codeword length affects the maximally allowed coding gain. After several experiments, we choose the codeword length to be 256. The codes with the length more than 256 do not benefit from the extra code length but suffer from the extra decoding complexity.

CHAPTER 5

Simulation Results and Comparisons

In this chapter, we show the simulation results under a variety of conditions. The simulations are based on the MIMO turbo equalization system which has already been described in detail in the previous chapter. The parameters that appear in this chapter are all well defined in the previous chapters. We discuss the performance difference (improvements or degradations) of these systems when applying different parameters. The performances of ideal case are shown in the figures as references.

5.1 Perfect Feedback

If we feed the feedback filters in the equalizer with original transmitted symbols, the adaptation will lead to an approach to the ideal Wiener solution as guaranteed by the theorem of LMS algorithm. The performance of this receiver is referred to as the perfect feedback case. It is the best performance that the turbo equalizers are able to achieve. The performance of a turbo system can be evaluated by the iterations required to reach an acceptable gap to the perfect feedback case. The amount of such gap can be considered as index of turbo system performance, too. Usually, many iterations are required to obtain an extreme small gap. Therefore, this is a trade-off between the receiver complexity and the performance.

5.2 Improvement by Iterations

To see the improvement made by increasing iterations, simulations are conducted on the transceiver defined below. The transmitter utilizes a space-time trellis code with 2 transmit antennas and 32 states. The receiver use a single receiver antenna and 10 iterations to recover the information bits. Figure 5.1 shows the symbol error rate (SER) of such system. Symbol error rate (SER) is the error rate measured at the equalizer output. We use this error rate as the performance index of an equalizer. From figure 5.1, we see that the equalizer benefits dramatically from the increase of iterations. Within a few iterations, the SER can approach the perfect feedback case closely. However, for low SNR, the SER does not benefit much from the increase of iterations. In fact, it may even become worse. This is reasonable because decisions made on erroneous information cannot be good for sequent iterations. The threshold when the performance starts to benefit from the increase of iterations is dependent on many factors, such as the noise and the channels. Generally, the more coding gain the decoder can provide, the lower this threshold is.

Figure 5.2 shows the bit error rate of the aforementioned system. The bit error rate (BER) calculation is based on the decisions of information bits from the decoder. The BER approaches the perfect feedback case. Because of the error-correcting ability provided by the trellis code, there is an improvement in the error rate between SER and BER.

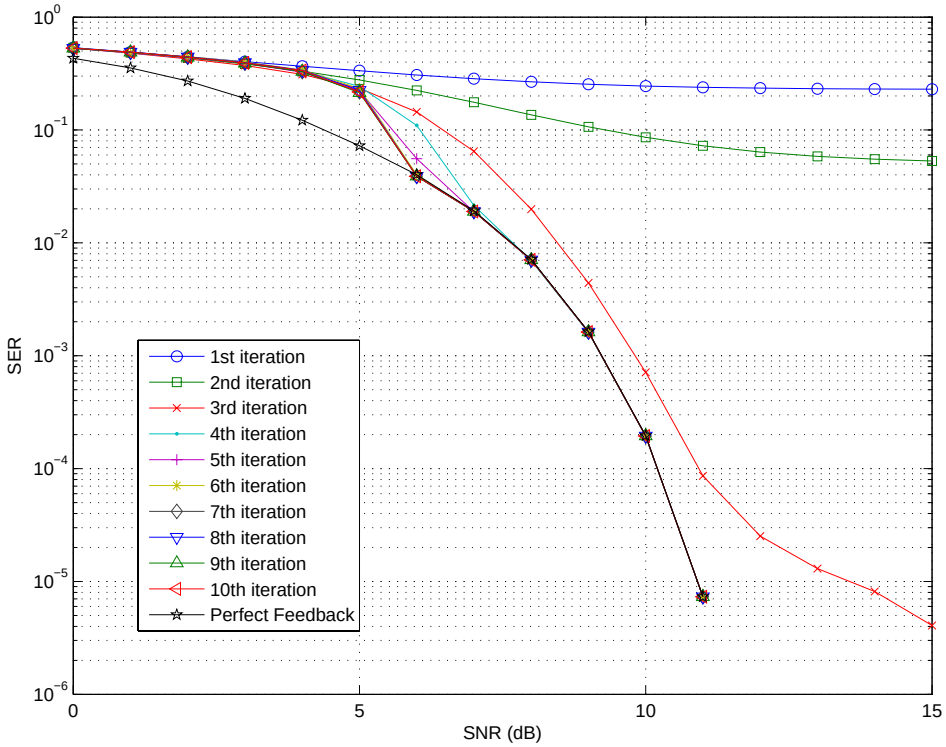

Figure 5.1 SER for 2 transmit antennas and 1 receive antenna and STTC with 32 states

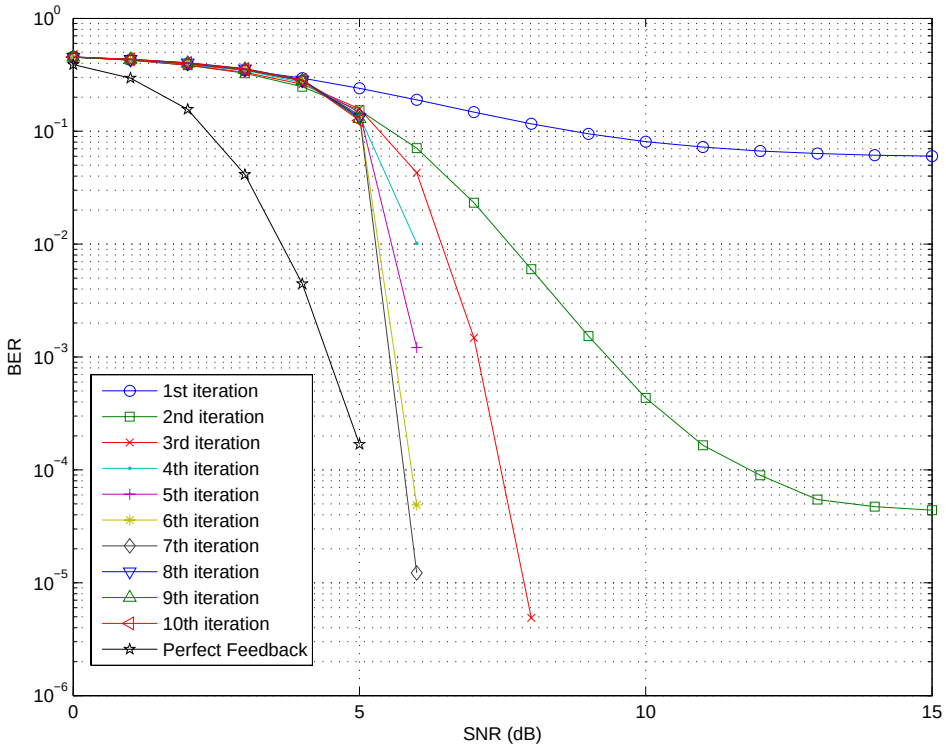

Figure 5.2 BER for 2 transmit antennas and 1 receive antenna and STTC with 32 states

Figure 5.3 and figure 5.4 shows the SER and the BER, respectively, of the same system except that 3 transmit antennas are used. For a system with 3 transmit antennas, it should be more difficult to recover all the transmitted symbols than the system with 2 transmit antennas because the co-channel interferences increase with the number of antennas. However, the theorem of space-time trellis codes suggests that more transmit antennas lead to larger diversity if the channel is frequency-flat. Nevertheless, the comparison between figure 5.3 and figure 5.1 is meaningless because different channel structures are applied in the simulations. Hence, we should consider these simulations independently. The results in figure 5.3 and figure 5.4 show that this turbo equalizer receiver works well for a 3-transmit antennas system,.

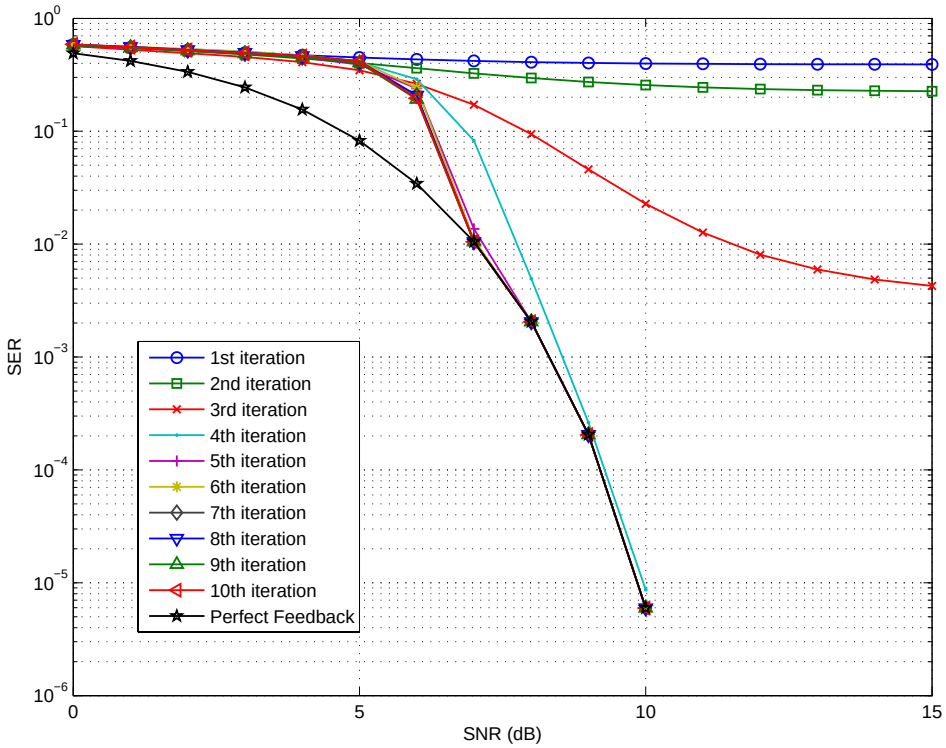

Figure 5.3 SER for 3 transmit antennas and 1 receive antenna and STTC with 32 states

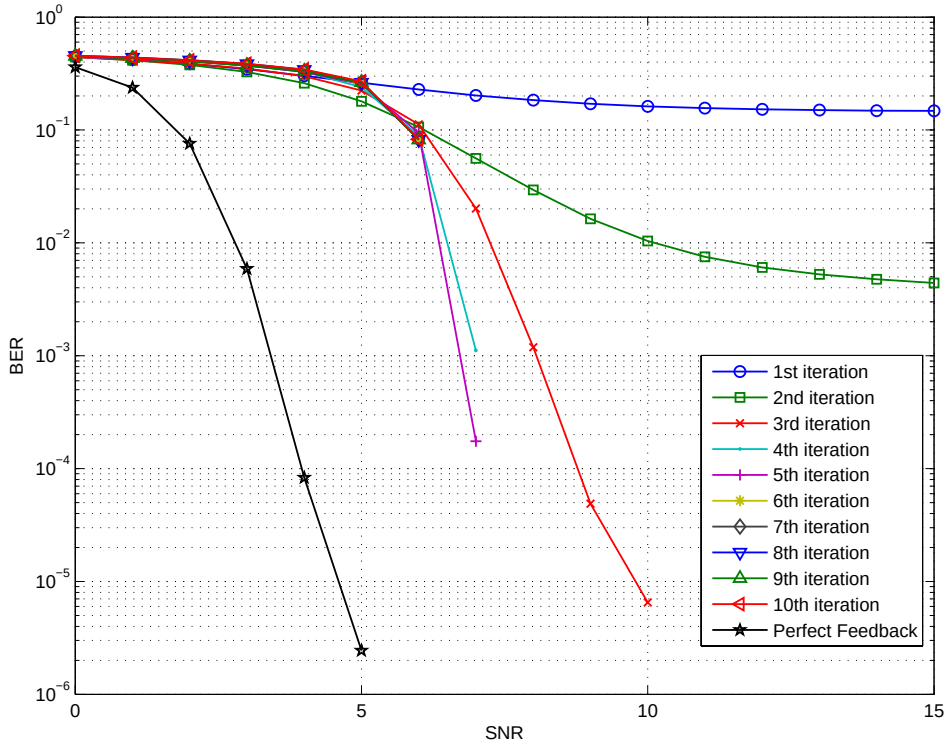

Figure 5.4 BER for 3 transmit antennas and 1 receive antenna and STTC with 32 states

Figure 5.5 and figure 5.6 show the SER and BER for 4 transmit antennas case. For a system with 4 transmit antennas, it should be more difficult to recover all the transmitted symbols than system using 2 or 3 transmit antennas, because the co-channel interferences become larger. Likewise, the comparison between figure 5.5 and figure 5.1 or figure 5.3 is meaningless because different channel structures are applied. We should consider these simulations independently. The results in figure 5.3 and figure 5.4 show that for a system with 4 transmit antennas, this turbo equalization works well, too.

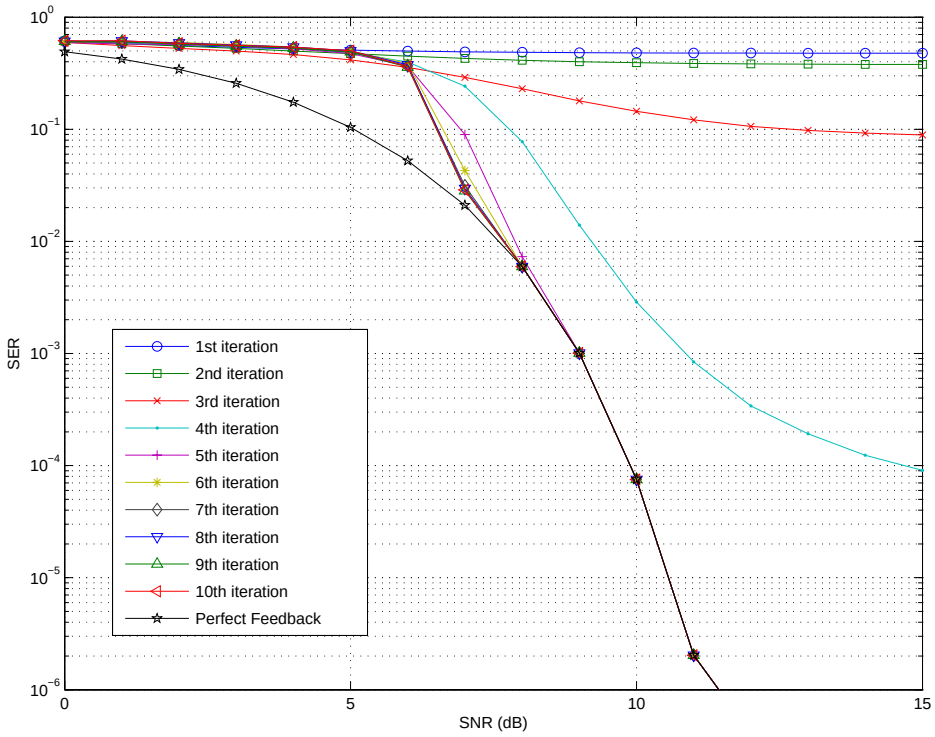

Figure 5.5 SER for 4 transmit antennas and 1 receive antenna and STTC with 32 states

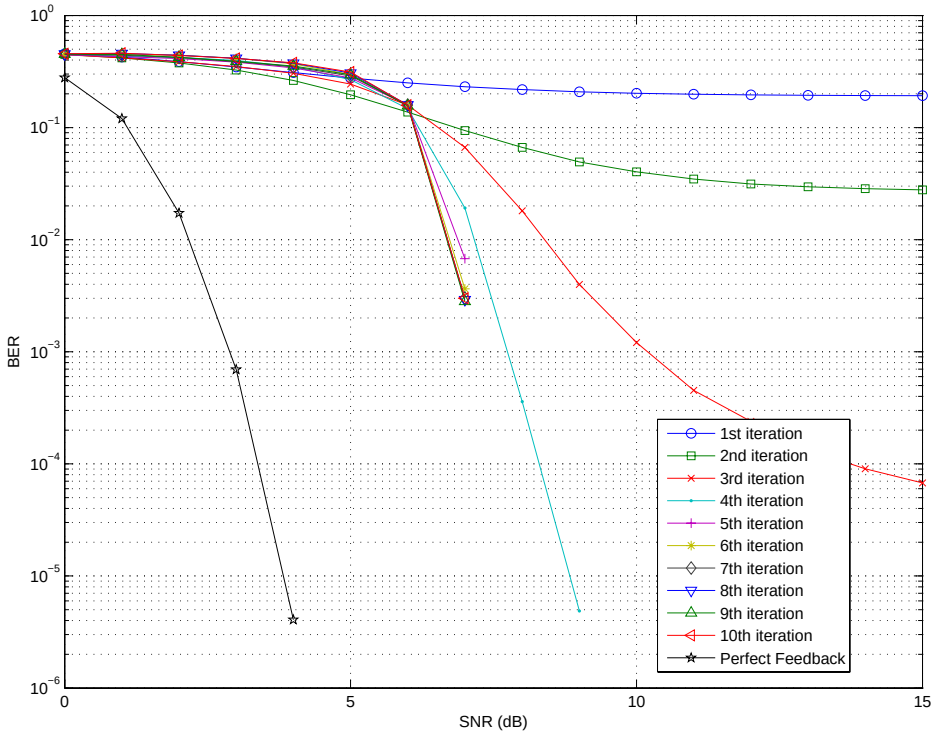

Figure 5.6 BER for 4 transmit antennas and 1 receive antenna and STTC with 32 states

5.3 Improvement by Two Receive Antennas

The advantages of adding extra antenna are described in 4.4.2. In this section, we conduct simulations with two receive antennas. Except for the receive antenna number, the other system parameters in the following simulations are the same as the ones in the single receive antenna case. Figure 5.7 and figure 5.8 show the SER and BER for 2 transmit antennas. Comparing the trends in figure 5.7 and figure 5.1, we can see that the equalizer of two receive antennas case makes great progress in the first iteration so that it approaches the perfect case more rapidly than the single antenna case. The number of iterations required to obtain the SER closed to the perfect case is 5 for single receive antenna case while only 2 iterations are needed for the two receiver antenna case. Similar result also holds for the BER. When we compare performance between one and two receive antennas, performance trends are our concern, not the values of SER or BER.

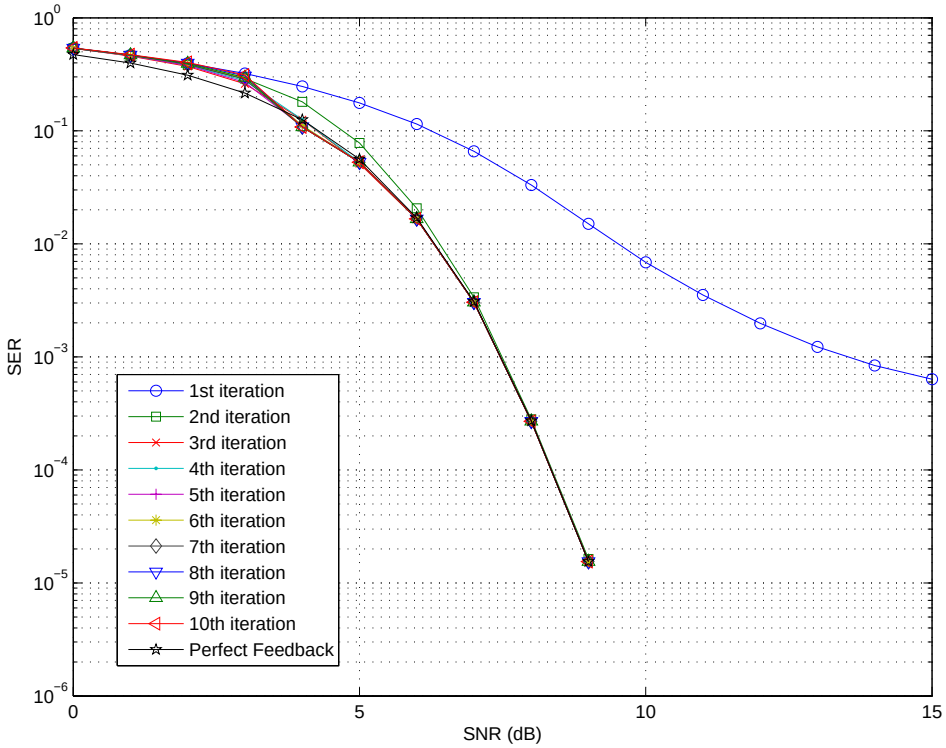

Figure 5.7 SER for 2 transmit antennas and 2 receive antenna and STTC with 32 states

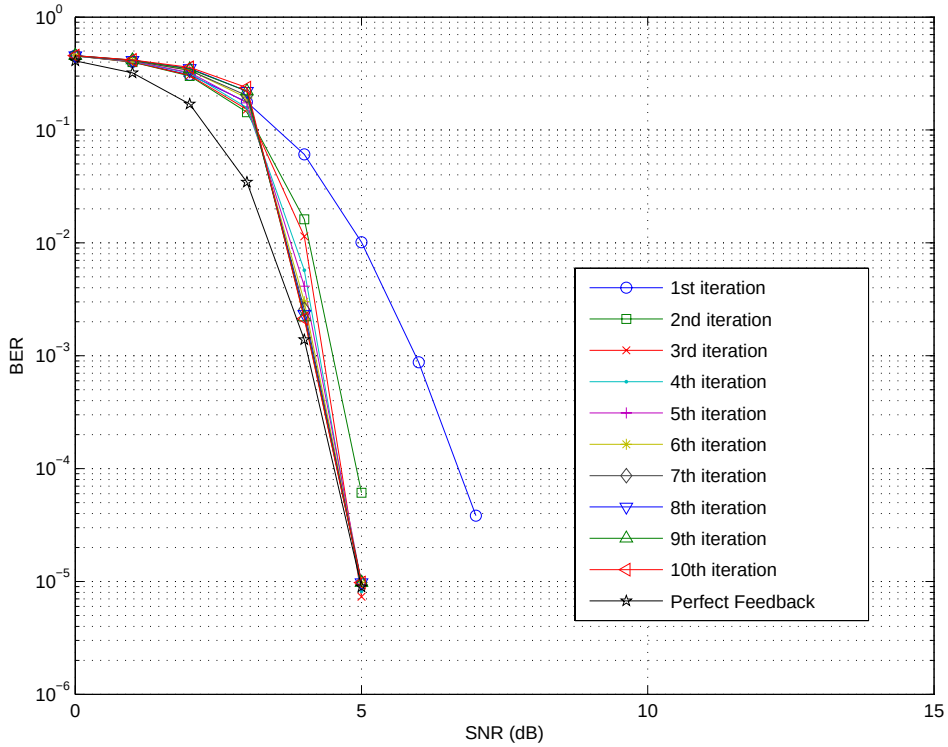

Figure 5.8 BER for 2 transmit antennas and 2 receive antenna and STTC with 32 states

Figure 5.9 to figure 5.12 are the SER and BER for the three transmit antennas case and the four transmit antennas case. The comparison results we made for the two transmit antennas case also holds for the three and four transmit antennas cases. All these figures show that using two receive antennas is very helpful. The turbo receiver approaches the perfect case more rapidly. Theoretically, the performance of the perfect feedback case will be better than the one using single receive antenna. However, due to the difference in channel structure, such comparison cannot be conducted based on our simulation results.

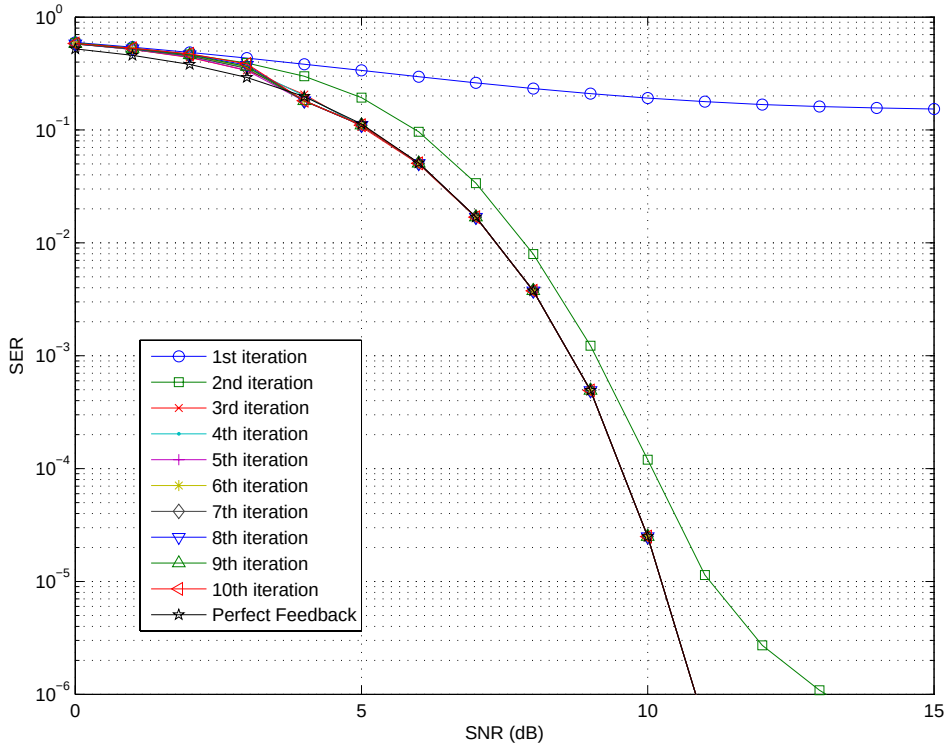

Figure 5.9 SER for 3 transmit antennas and 2 receive antenna and STTC with 32 states

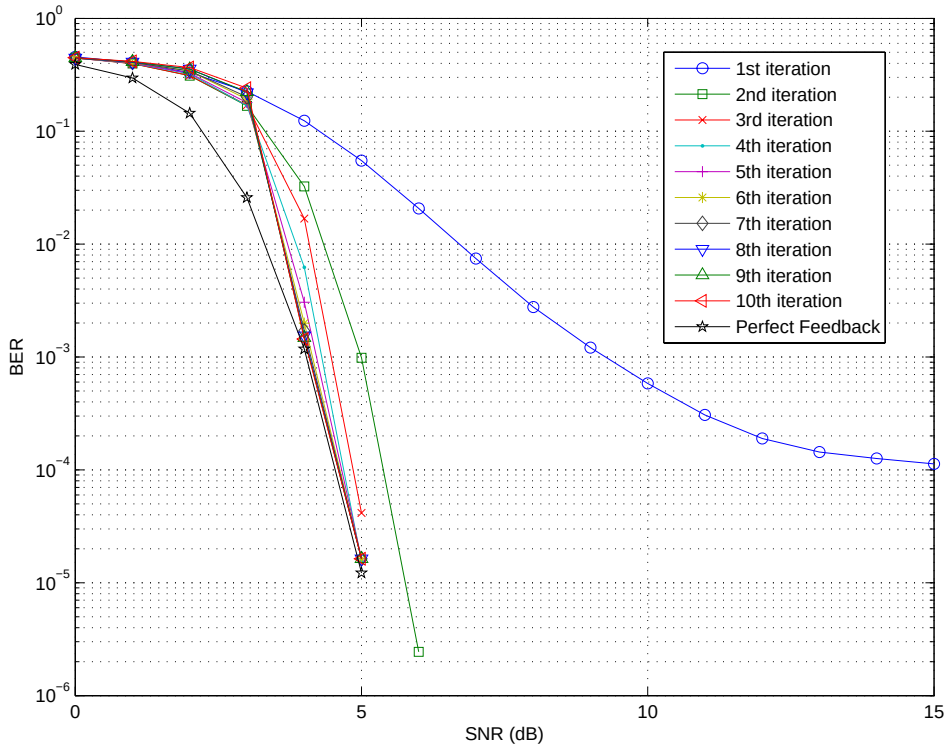

Figure 5.10 BER for 3 transmit antennas and 2 receive antenna and STTC with 32 states

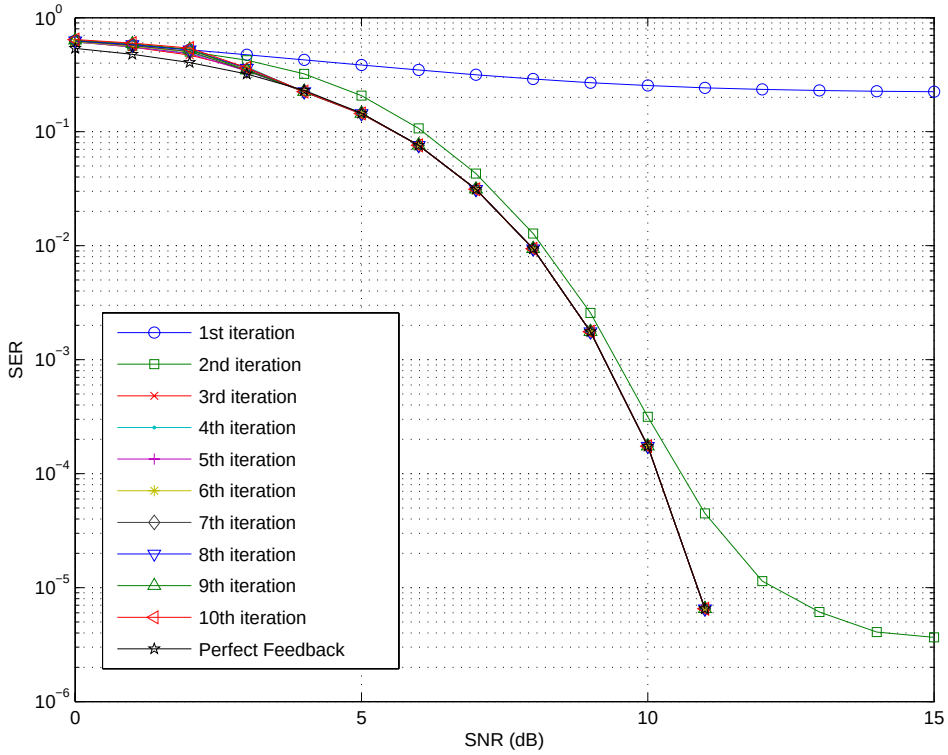

Figure 5.11 SER for 4 transmit antennas and 2 receive antenna and STTC with 32 states

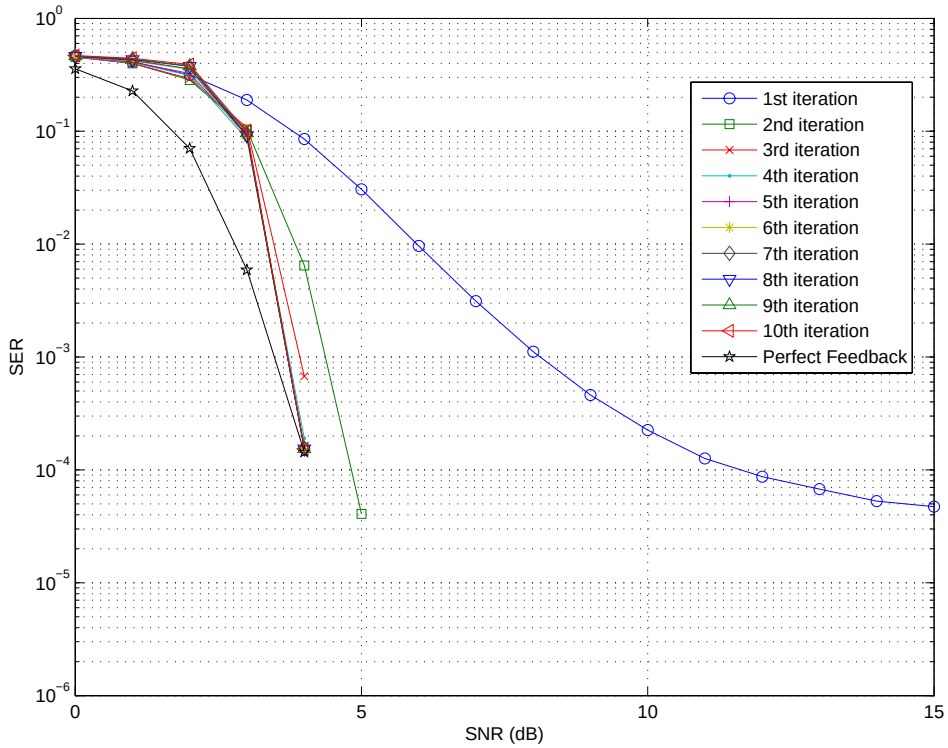

Figure 5.12 BER for 4 transmit antennas and 2 receive antenna and STTC with 32 states

5.4 Effect of Interleavers

In this section, the effect of interleavers is discussed. The selection of the interleaver is essential to turbo systems as mentioned in 4.1.2. We would like to see how the interleavers affect the performance of the turbo system. We simulate the systems whose parameters are all the same except for the block size of interleavers. The system is configured to have 2 transmit antennas and 1 receive antenna. The STTC with 32 states is adopted. Figure 5.13 shows the BER of different interleaver sizes. Figure 5.14 shows the SER of different interleaver sizes. For the perfect feedback cases, the block size of interleaver affects only the coding gain. There is no influence in the SER but a small influence in the BER. However, for turbo systems, the effect of interleaver sizes is significant. At the fifth iteration, the BERs of different interleaver sizes are quite different. The one with the larger interleaver size has smaller BER than that with the smaller interleaver size, as shown in figure 5.13 and figure 5.14. Therefore, the large interleaver size will result in a fast convergence rate and a small error rate at steady state.

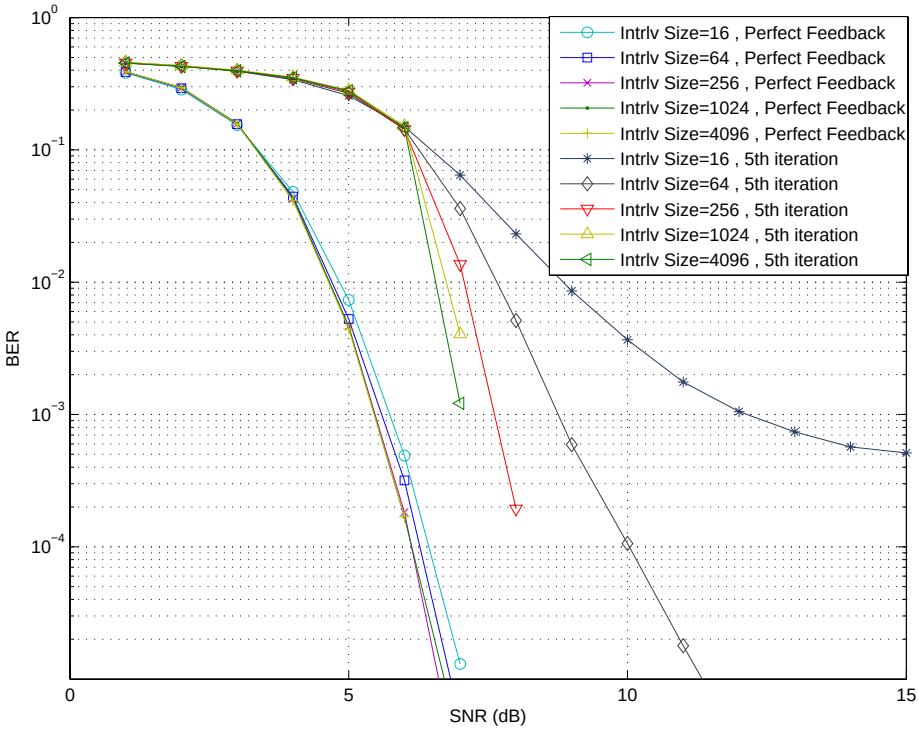

Figure 5.13 BER for different interleaver sizes at 5th iteration with 2 transmit antennas and 1 receive antenna and STTC with 32 states

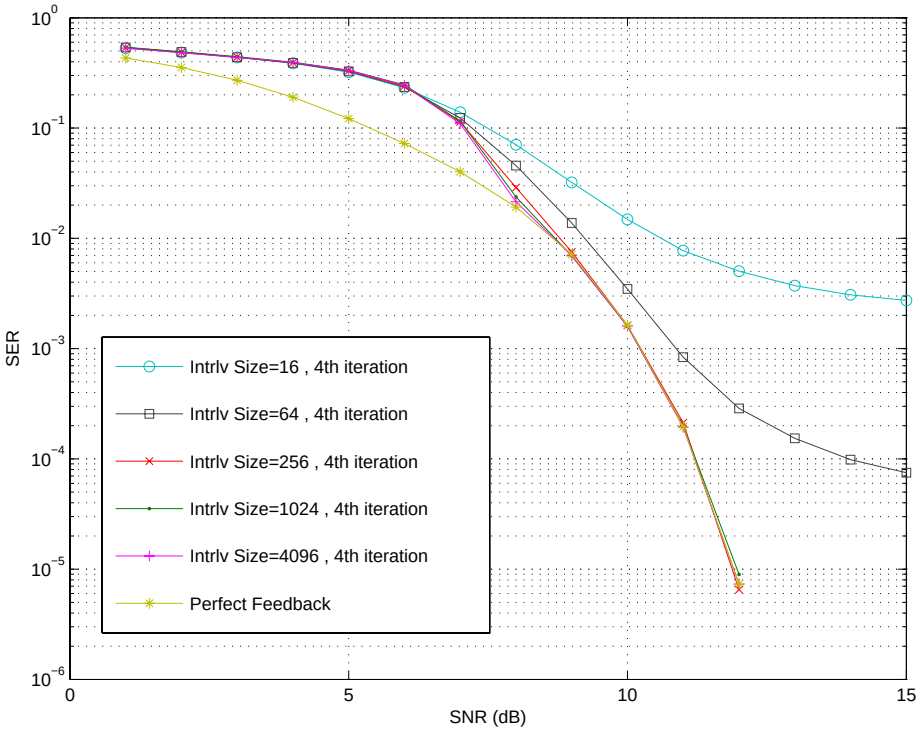

Figure 5.14 SER for different interleaver sizes at 4th iteration with 2 transmit antennas and 1 receive antenna and STTC with 32 states

CHAPTER 6

Conclusions and Perspectives

6.1 Conclusions

In this thesis, we propose a low-complexity turbo equalizer structure for multiple-antenna systems. Through the iterative detection, the simple equalizer is capable of removing inter-symbol interference and co-channel interference. The cross-canceling filters in the feedback part remove the co-channel interference effectively. The simulations of the systems with two, three and four transmit antennas show that the equalizer can successfully recover the transmitted symbols in spite of the increase in the antenna number, which implies the increase of co-channel interference,. As long as the a priori information from the previous iteration is not too erroneous, the cross-canceling filters can remove co-channel interference well. When two receive antennas are adapted instead of a single antenna, the extra cross-filters in the feedforward part can boost the performance of the equalizer. The simulations show that the performance improvement caused by increasing the number receive antennas are dramatic at the first iteration. With the lower error rate at the first iteration, the convergence to the perfect feedback solution becomes faster, that is, less iterations are required. In addition, two receive antennas lower the error rate of the perfect feedback case. The use of interleavers brings the function of

de-correlation to the data sequence and results in the performance gain. In a turbo system, this de-correlation is especially important. Through simulations, we show the relationship between the interleaver size and the required number of iterations, which provides a trade-off to the system designer. As a conclusion, the turbo equalization we propose is capable of combating frequency-selective channels with the appropriate interleaver size.

6.2 Perspective

From the fundamental view point, the equalizer we proposed is to recover each transmitted sequences of each transmit antenna. When working with one sequence, the other sequences become co-channel interference. The simulations shows that the equalizer can still recover the sequence successfully even when there are four transmit antennas. This idea can be extended to all kinds of co-channel interference including multi-user interference. In addition, the proposed receiving scheme can incorporate with other advanced techniques into the turbo system. With more blocks that generate soft information, the turbo system can be more effective and have faster convergence.

Bibliography

[1] V. Tarokh, N. Seshadri, and A. R. Calderbank, “Space–time codes for high data rate wireless communication: Performance criterion and code construction,” *IEEE Trans. Inform. Theory*, vol. 44, pp. 744–765, Mar. 1998.

[2] J. Guey, M. P. Fitz, M. R. Bell, and W. Kuo, “Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels,” in *Proc. IEEE VTC*, vol. 1, 1996, pp. 136-140.

[3] S.S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1987

[4] R. Horn and C. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.

[5] H. Boleskei, D. Gesbert and A. J. Paulraj, “On the capacity of OFDM-based multi-antenna systems”, in Proc. ICASSP’00, 2000, pp. 2569-2572.

[6] A. Wittneben, “A new bandwidth efficient transmit antenna modulation diversity scheme for linear digital modulation,” in Proc. IEEE’ICC, 1993, pp. 1630-1634.

[7] N. Seshadri and J. H. Winters, “Two signaling schemes for improving the error performance of frequency-division-duplex (FDD) transmission systems using transmitter antenna diversity,” Int. J. Wireless Inform. Networks, vol. 1, no. 1, 1994.

[8] V. Tarokh, H. Jafarkhani and A. R. Calderband, “Space-time block codes from orthogonal designs”, *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1456-1467, July 1999.

[9] S. M. Alamouti, “A simple transmit diversity technique for wireless communications”, *IEEE Journal Select. Areas Commun.*, vol. 16, no. 8, pp. 1451-1458, Oct. 1998.

[10] J. Grimm, *Transmitter diversity code design for achieving full diversity on*

Rayleigh fading channels, Ph.D. dissertation, Purdue University, 1998.

[11] S. Baro, G. Bauch, and A. Hansmann, “Improved codes for space-time trellis-coded modulation,” IEEE Comm. Lett., vol. 4, pp. 20-22, January 2000.

[12] A. R. Hammons and H. E. Gamal, “On the theory of space-time codes for PSK modulation,” IEEE Trans. Information Theory, vol. 46, no. 2, pp. 524-542, March 2000.

[13] V. Tarokh, A. Naguib, N. Seshadri, and A. R. Calderbank, “Space-time codes for high data rate wireless communication: performance criteria in the presence of channel estimation errors, mobility, and multiple paths,” IEEE Trans. Communications, vol. 47, no. 2, pp. 199–207, 1999.

[14] S. Baro, G. Bauch and A. Hansmann, “Improved codes for space-time trellis coded modulation”, IEEE Commun. Lett., vol. 4, no. 1, pp. 20-22, Jan. 2000.

[15] A. Leon-Garcia, Probability and random processes for electrical engineering, Addison Wesley, 1994

[16] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, “Optimal decoding of linear codes for minimizing symbol error rate,” IEEE Trans. Inform. Theory, vol. 20, pp. 284–287, Mar. 1974.

[17] C.E. Shannon, “A mathematical theory of communication”, Bell Syst. Tech. J., vol. 27, pp. 379-423 (Part one), pp. 623-656(Part two), Oct. 1948, reprinted in book form, University of Illinois Press, Urbana, 1949.

[18] Y. Gong and K. B. Letaief, “Performance evaluation and analysis of space-time coding in unequalized multipath fading links,” IEEE Trans. Communications, vol. 48, no. 11, pp. 1778– 1782, 2000.

[19] A. Gersho and T. L. Lim, “Adaptive cancellation of intersymbol interference for data transmission,” Bell Syst. Tech. J., vol. 60, o. 11, pp. 1997–2021, Nov. 1981.

[20] M. S. Muller and J. Salz, “A unified theory of data-aided equalization,” Bell Syst. Tech. J., vol. 60, no. 9, pp. 2023–2038, Nov. 981.

[21] C. Douillard, M. Jezequel, C. Berrou, A. Picart, P. Didier, and A. Glavieux, “Iterative correction of intersymbol interference: Turbo equalization,” European Trans. Telecomm., vol. 6, pp. 507–511, Sept.–Oct 1995.

[22] R. R. Lopes, “Iterative estimation, equalization and decoding,” Ph.D. dissertation, Georgia Inst. Technol., Atlanta, GA, 2003.

[23] M. Tüchler, R. Koetter, and A. C. Singer, “Turbo equalization: Principles and new results,” IEEE Trans. Commun., vol. 50, no. 5, pp. 754–767, May 2002.

[24] R. R. Lopes, and J. R. Barry, “The Soft-Feedback Equalizer for turbo Equalization of Highly Dispersive Channels,” IEEE Trans. Commun., vol. 54, no. 5, pp. 783– 788, 2006.