

# 國立交通大學

## 電信工程學系碩士班 碩士論文

多輸入多輸出正交分頻多工系統中基於編碼可靠度  
混合重傳機制及適應性調變編碼之聯合設計



Joint Design of Reliability-Based Hybrid ARQ and  
Adaptive Modulation / Coding in MIMO-OFDM Systems

研究生：王俊傑

Student : Chun-Chieh Wang

指導教授：李大嵩 博士

Advisor : Dr. Ta-Sung Lee

中華民國九十五年六月

多輸入多輸出正交分頻多工系統中基於編碼可靠度  
混合重傳機制及適應性調變編碼之聯合設計

Joint Design of Reliability-Based Hybrid ARQ and  
Adaptive Modulation / Coding in MIMO-OFDM Systems

研 究 生：王 俊 傑

Student : Chun-Chieh Wang

指 導 教 授：李 大 嵩 博 士

Advisor : Dr. Ta-Sung Lee



A Thesis

Submitted to Institute of Communication Engineering  
College of Electrical Engineering and Computer Science

National Chiao Tung University

in Partial Fulfillment of the Requirements

for the Degree of

Master of Science

in

Communication Engineering

June 2006

Hsinchu, Taiwan, Republic of China

中 華 民 國 九 十 五 年 六 月


# 多輸入多輸出正交分頻多工系統中 基於編碼可靠度混合重傳機制 及適應性調變編碼之聯合設計

學生：王俊傑

指導教授：李大嵩 博士

國立交通大學電信工程學系碩士班

## 摘要



編碼混合重傳機制(Hybrid Automatic Repeat Request, HARQ)為一種結合順向錯誤更正(Forward Error Correction, FEC)與自動重送機制(Automatic Repeat Request, ARQ)的錯誤更正技術；並被認為是第四代高速通訊在無線通道雜訊干擾下，解決錯誤更正問題之可行技術，而編碼可靠度混合重傳(reliability-based HARQ, RB-HARQ)機制為一新式編碼混合重傳機制，僅需重傳信賴度較低的位元，較傳統機制提供更優異效能表現且更具適應性；本論文研究主題之一，即為根據不同服務品質要求，針對編碼可靠度混合重傳機制提出一套適應性選擇重傳封包長度的演算法。另一方面，多輸入多輸出正交分頻多工(Multiple-Input Multiple-Output Orthogonal Frequency Division Multiplexing, MIMO-OFDM)系統被認為是符合第四代高速通訊需求的最佳解決方案之一；MIMO 為使用多天線於傳送和接收端的可靠通訊技術，OFDM 為一種具高頻譜效益，並能有效克服多路徑衰落效應的調變技術；本論文中研究主題之二，即為針對 MIMO-OFDM 系統提出一種結合編碼可靠度混合重傳機制與調變編碼的跨層適應性收發架構，使其能夠隨時間動態地在頻率與空間通道上調整傳輸參數。吾人根據媒體存取層不同的服務品質要求，包含容許最大延遲時間及封包錯誤率，適應性的調整重傳次數、重傳封包長度、傳輸功率、傳輸速率、調變型態等系統參數，達到最佳的性能。最後，吾人藉由電腦模擬驗證上述架構在寬頻無線接取通道環境中具有優異的傳輸表現。

# Joint Design of Reliability-Based Hybrid ARQ and Adaptive Modulation / Coding in MIMO-OFDM Systems

Student: Chun-Chieh Wang

Advisor: Dr. Ta-Sung Lee

Institute of Communication Engineering  
National Chiao Tung University

## Abstract

Hybrid automatic repeat request (HARQ) that combines ARQ and forward error correction (FEC) is a promising error-correcting technique for wireless communications. In reliability-based HARQ (RB-HARQ), the bits that are to be retransmitted are adaptively selected at the receiver based on the estimated bit reliabilities at the output of a soft decoder. This technique has the potential of improving system throughput. In this thesis, we propose an adaptive algorithm which can accordingly choose the sizes of retransmissions under the quality of service (QoS) constraints such as the maximum number of retransmissions allowed per packet and packet loss probability. Multiple-input multiple-output orthogonal frequency-division multiplexing (MIMO-OFDM) is suitable for the increasing demand of the high-performance 4G broadband wireless communications with multiple antennas at both the transmitter and receiver sides. In this thesis, we then consider a new wireless communication system combining both MIMO-OFDM and RB-HARQ techniques and propose an adaptive MIMO-OFDM transceiver architecture along with a specifically designed loading procedure to dynamically adjust the transmission parameters such as retransmission size, number of retransmission, modulation order and transmit power over spatial and frequency channels, according to the instantaneous channel statistics, to meet the target QoS. Finally, we evaluate the performance of the proposed systems, and confirm that it functions well in a typical broadband wireless access environment.

# Acknowledgement

I would like to express my deepest gratitude towards my advisor, Dr. Ta-Sung Lee, for his enthusiastic guidance and great patience. His positive attitude has guided me in many areas and has propelled me in the direction of reaching my future goals. In addition, I would like to express many heartfelt thank you to all the members and staff of the Communication System Design and Signal Processing (CSDSP) Lab for their constant support and encouragement; I would not be where I am today without your help. Last but not least, I would like to show my most sincere appreciation and love for my family for their continual love and support of my pursuit of excellence. Thank you once again.



# Contents

Chinese Abstract	I
English Abstract	II
Acknowledgement	III
Contents	IV
List of Figures	VII
List of Tables	XI
Acronym Glossary	XII
Notations	XV
1 Introduction	1
2 Overview of IEEE 802.16 System	6
2.1 Review of MIMO-OFDM System .....	6
2.1.1 OFDM : Concept and Technique .....	7
2.1.2 MIMO : Concept and Technique .....	10
2.1.3 V-BLAST Based OFDM.....	17
2.2 WiMAX Overview .....	19
2.2.1 Review of IEEE 802.16 PHY .....	20
2.2.2 Review of IEEE 802.16 MAC .....	21



2.2.3	Review of IEEE 802.16-2005 .....	23
2.3	MIMO Channel Model .....	24
2.3.1	Correlation Channel Matrices .....	25
2.3.2	Generation of a MIMO Channel Using Correlation Matrix Approach .....	27
2.4	Computer Simulations .....	30
3	Reliability-Based Incremental Redundancy Hybrid ARQ Scheme with LDPC Codes .....	33
3.1	Review of LDPC Codes .....	33
3.1.1	LDPC Codes .....	34
3.1.2	Construction of LDPC Codes .....	35
3.1.3	Generator matrix of LDPC Codes .....	39
3.1.4	Decoding Algorithm of LDPC Codes .....	39
3.2	Review of ARQ Schemes .....	43
3.2.1	Conventional ARQ .....	43
3.2.2	Hybrid ARQ Type I .....	44
3.2.3	Hybrid ARQ Type II .....	45
3.2.4	Hybrid ARQ Type III .....	46
3.3	Reliability-Based HARQ .....	47
3.3.1	Review of Reliability-Based HARQ .....	48
3.3.2	System Model for RB-HARQ scheme .....	48
3.4	Proposed Adaptive RB-HARQ Algorithm .....	51
3.5	Computer Simulations .....	55
3.6	Summary .....	61

4	Combing RB-HARQ and Adaptive Modulation in IEEE	
	802.16-like MIMO-OFDM Systems	62
4.1	The Concept of Cross Layer Design .....	63
4.2	Adaptive Modulation Assisted MIMO -OFDM System.....	64
4.2.1	Adaptive Modulation .....	65
4.3	System Model for V-BLAST Based Apadtive MIMO-OFDM	
	System .....	67
4.4	V-BLAST Based Adapdtive MIMO-OFDM System.....	69
4.5	Combining HARQ with AMC Mechanism.....	75
4.5.1	System Performance Requirement at Physical Layer.....	75
4.5.2	AMC Design at the Physical Layer .....	77
4.5.3	Error Performances of AMC Design .....	78
4.6	Adaptive RB-HARQ with AMC Mechanism .....	79
4.7	Physical/MAC Cross-Layer AMC Design.....	80
4.8	Computer Simulations.....	81
4.9	Summary .....	88
5	Conclusion	89
	Bibliography	91



# List of Figures

Figure 2.1	Diagram of a MIMO wireless transmission system.....	10
Figure 2.2	Illustration of a spatial multiplexing system.....	11
Figure 2.3	Diagonal and vertical layered space-time encoding with $N_t = 3$ .....	13
Figure 2.4	Diagonal layered space-time decoding with $N_t = 3$ .....	14
Figure 2.5	Vertical layered space-time decoding with $N_t = 3$ .....	15
Figure 2.6	V-BLAST based MIMO-OFDM transmitter architecture.....	19
Figure 2.7	V-BLAST based MIMO-OFDM receiver architecture. ....	19
Figure 2.8	The flow chart for the generation of MIMO channel model coefficients.....	28
Figure 2.9	ZF V-BLAST performance with ideal detection and cancellation. QPSK modulation is used. $(N_t, N_r) = (4, 4)$ .....	31
Figure 2.10	ZF V-BLAST performance with error propagation. QPSK modulation is used. $(N_t, N_r) = (4, 4)$ .....	31
Figure 2.11	Comparison of ZF V-BLAST $(N_t, N_r) = (4, 4)$ with QPSK modulation and $(N_t, N_r) = (2, 4)$ with 16-QAM modulation. ....	32
Figure 3.1	Parity-check matrix of the (960, 640) LDPC code .....	37
Figure 3.2	Generator matrix of the (960, 640) LDPC code.....	37
Figure 3.3	Bipartite graph of (7, 4) Hamming Code .....	40
Figure 3.4	Constraint node of (7, 4) Hamming Code.....	40

Figure 3.5	Block diagram of ARQ and HARQ Type I.....	43
Figure 3.6	Block diagram of HARQ Type II.....	46
Figure 3.7	Block diagram of RB-HARQ. ....	47
Figure 3.8	System model of RB-HARQ scheme .....	49
Figure 3.9	Packet reliability versus retransmission size (bits) in RB-HARQ scheme .....	53
Figure 3.10	PER versus packet reliability in RB-HARQ scheme.....	54
Figure 3.11	PER versus average $E_b / N_0$ for (960,640) LDPC coded system in a Rayleigh channel with QPSK, 16-QAM, and 64-QAM. ....	57
Figure 3.12	Throughput (bits/symbol) versus average $E_b / N_0$ for (960,640) LDPC coded system in a Rayleigh fading channel with QPSK, 16-QAM, and 64-QAM.....	57
Figure 3.13	PER versus average <i>effective</i> $E_b / N_0$ for (960,640) LDPC coded system in a Rayleigh fading channel with QPSK by the RB-HARQ scheme, no-ARQ scheme, and IR-HARQ scheme .....	58
Figure 3.14	PER versus average SNR for (960,640) LDPC coded system in a Rayleigh fading channel with QPSK by HARQ scheme, RB-ARQ scheme, and proposed RB-HARQ scheme (required PER = $10^{-2}$ ) ....	60
Figure 3.15	Throughput (bits/symbol) versus average SNR for (960,640) LDPC coded system in Rayleigh fading channel with QPSK by HARQ scheme, RB-ARQ scheme, and proposed RB-HARQ scheme (required PER = $10^{-2}$ ) .....	61
Figure 4.1	System architecture of proposed V-BLAST based adaptive MIMO-OFDM system.....	66
Figure 4.2	Proposed V-BLAST based adaptive MIMO-OFDM system transmitter architecture .....	67

Figure 4.3	Proposed V-BLAST based adaptive MIMO-OFDM system receiver architecture.....	68
Figure 4.4	Throughput versus average SNR for the cross layer design AMC system (required $P_{loss} = 10^{-2}$ ) in IEEE 802.20 Model-C channel. $(N_t, N_r) = (2, 2)$ with RB-HARQ scheme, ARQ scheme, and AMC-only scheme. Other parameters are listed in Table 4.1.....	83
Figure 4.5	Retransmission data rate versus average SNR for the cross layer design AMC system (required $P_{loss} = 10^{-2}$ ) in IEEE 802.20 Model-C channel. $(N_t, N_r) = (2, 2)$ with RB-HARQ scheme, and ARQ scheme. Other parameters are listed in Table 4.1.....	84
Figure 4.6	PER versus average SNR for the cross layer design AMC system (required $P_{loss} = 10^{-2}$ ) in IEEE 802.20 Model-C channel. $(N_t, N_r) = (2, 2)$ with RB-HARQ scheme. Other parameters are listed in Table 4.1.....	84
Figure 4.7	Throughput versus average SNR for the cross layer design AMC system (required $P_{loss} = 10^{-2}$ ) in IEEE 802.20 Model-C channel. $(N_t, N_r) = (1, 1)$ with RB-HARQ scheme, ARQ scheme, and AMC-only scheme. Other parameters are listed in Table 4.1.....	85
Figure 4.8	Retransmission data rate versus average SNR for the cross layer design AMC system (required $P_{loss} = 10^{-2}$ ) in IEEE 802.20 Model-C channel. $(N_t, N_r) = (1, 1)$ with RB-HARQ scheme, and ARQ scheme. Other parameters are listed in Table 4.1.....	85
Figure 4.9	PER versus average SNR for the cross layer design AMC system (required $P_{loss} = 10^{-2}$ ) in IEEE 802.20 Model-C channel. $(N_t, N_r) = (1, 1)$ with RB-HARQ scheme. Other parameters are listed in Table 4.1.....	86

Figure 4.10 Throughput versus average SNR for the cross layer design AMC system (required  $P_{loss} = 10^{-2}$ ) in IEEE 802.20 Model-C channel in the presence of partial CSIT  $\Sigma_{\tilde{h}_f} = 0.01\mathbf{I}$ .  $(N_t, N_r) = (2, 2)$  with RB-HARQ scheme, ARQ scheme, and AMC-only scheme. Other parameters are listed in Table 4.1 .....87

Figure 4.11 PER versus average SNR for the cross layer design AMC system (required  $P_{loss} = 10^{-2}$ ) in IEEE 802.20 Model-C channel in the presence of partial CSIT  $\Sigma_{\tilde{h}_f} = 0.01\mathbf{I}$ .  $(N_t, N_r) = (2, 2)$  with RB-HARQ scheme, ARQ scheme, and AMC-only scheme. Other parameters are listed in Table 4.1 .....88



# List of Tables

Table 2.1	Summary of SISO link-level parameters for IEEE 802.20 channel models.....	25
Table 2.2	Summary of SISO environment parameters for IEEE 802.20 channel models .....	25
Table 2.3	Summary of MIMO link-level parameters for IEEE 802.20 channel models.....	29
Table 3.1	LDPC block sizes and code rates in IEEE 802.16-2005.....	38
Table 3.2	Notations of sum-product algorithm.....	63
Table 3.3	Simulation parameters of the adaptive RB-HARQ system .....	56
Table 4.1	Simulation parameters of the ZF-BLAST based cross layer design AMC system .....	82
Table 4.2	SNR threshold table for various M-QAM at target $PER = 10^{-2}$ .....	82

# Acronym Glossary

4G	the fourth generation
AMC	adaptive modulation and coding
ARQ	automatic repeat request
AWGN	additive white Gaussian noise
BER	bit error rate
BPSK	binary phase shift keying
BLAST	Bell Lab Layered space time
BS	base station
CP	cyclic prefix
CRC	cyclic redundancy check
CSI	channel state information
CSIT	channel state information in the transmitter
DFT	discrete Fourier transform
FDD	frequency division duplex
FFT	fast Fourier transform
HARQ	hybrid automatic repeat request
ICI	intercarrier interference
IEEE	institute of electrical and electronics engineers
IFFT	inverse fast Fourier transforms
ISI	intersymbol interference
LDPC	low density parity-check
LOS	line of sight
MAC	medium access control layer
MIMO	multiple-input multiple-output
ML	maximum likelihood

MMSE	minimum mean square error
MRC	maximal ratio combining
MS	mobile station
MUX	multiplex
OFDM	orthogonal frequency division multiplexing
OSIC	ordered successive interference cancellation
PHY	physical layer
PER	packet error rate
QAM	quadrature amplitude modulation
QoS	quality of service
QPSK	quaternary phase shift keying
RB-HARQ	reliability-based hybrid automatic repeat request
RF	radio frequency
RX	receiver
SD	spatial diversity
SM	spatial multiplexing
SNR	signal-to-noise ratio
SIC	successive interference cancellation
STC	space-time coding
TDD	time division duplex
TX	transmitter
V-BLAST	vertical Bell laboratory layered space-time
ZF	zero forcing



# Notations

$b_i$	rate at the $i$ th transmit antenna
$C$	transmission code words matrix
$f_d$	Doppler frequency
$E_b$	bit energy
$E_s$	symbol energy
$h_t^{i,j}$	channel gain between the $j$ th transmit and $i$ th receive antenna at time $t$
$H[k]$	channel frequency response on the $k$ th subcarrier
$M$	modulation order
$N_c$	number of subcarriers (FFT/IFFT size)
$N_{cp}$	number of guard interval samples
$N_t$	number of transmit antenna
$N_r$	number of receive antenna
$N_0$	noise power spectrum density
$p_n$	path metric associated with the $n$ th information bit
$P_{budget}$	power budget
$q$	antenna state
$r_t^i$	received data at the $i$ th transmit at time $t$
$S$	set of signal constellation
$s_t^j$	transmitted signal form the $j$ th transmit at time $t$
$\mathbf{T}$	set of switching levels
$T_s$	symbol duration
$T_{sample}$	sampling period
$d[k]$	input symbol on the $k$ th subcarrier
$r[k]$	received data on the $k$ th subcarrier
$\eta[k]$	additive white noise vector on the $k$ th subcarrier



$\mathbf{w}_j$	weighting vector for the $j$ th layer
$\eta_t^i$	additive white noise at the $i$ th receive antenna at time $t$
$\sigma_n^2$	noise power
$\mathcal{E}_{error}$	target BER
$\gamma$	instantaneous SNR
$\tau_{rms}$	root mean squared excess delay spread
$\rho$	average SNR at each receive antenna
$\lambda$	eigenvalue



# Chapter 1

## Introduction

Next generation broadband wireless communication systems are expected to provide users with multimedia services such as high-speed internet access, wireless television, mobile computing, and etc. The rapid growing demand for these services is driving the wireless communication technology towards higher data rates, higher mobility and higher carrier frequency. However, the physical limitation of the wireless channel, typically subject to both time-selective and frequency-selective fading that are induced by carrier phase/frequency drifts, Doppler shifts and multipath propagation, presents a fundamental challenge for reliable communications. On the other hand, the limited availability of bandwidth promotes an emerging issue of high spectral efficiency. Hence, recent research efforts are carried out to develop efficient coding and modulation schemes along with sophisticated signal processing algorithms to improve the quality and spectral efficiency of wireless communication links. Some popular examples include smart antenna, in particular multiple-input multiple-output (MIMO) technology [1]-[6], coded multicarrier modulation, adaptive modulation [7]-[10], and link-level retransmission techniques [11].

MIMO systems can be defined as follows: Given an arbitrary wireless system, we consider a link for which the transmitter side as well as the receiver side is equipped

with multiple antennas. Such setup is illustrated in Figure 2.1. The signals on the transmit antennas at one end and the receive antennas at the other end are “co-processed” in such a way that the quality (packet error rate (PER)) or the data rate (bits/sec) of the communication link is improved. A core idea in MIMO-OFDM systems is the space-time signal processing in which time is complemented with the spatial dimension inherent in the use of multiple spatially distributed antennas. A key feature of MIMO systems is to efficiently exploit the multipath, rather than mitigate it, to achieve the signal decorrelation necessary for separating the co-channel signals. Specifically, the multipath phenomenon presents itself as a source of diversity that takes advantage of random fading.

Orthogonal frequency division multiplexing (OFDM) is a multipath-friendly mechanism that treats the whole transmission band as a set of adjacent narrow sub-bands. This property leads OFDM to be chosen over a single-carrier solution to avoid using a complicated equalizer, which is usually a heavy burden in a wideband communication receiver. Moreover, with proper coding and interleaving across frequencies, multipath turns into an OFDM system advantage by yielding frequency diversity. OFDM can be implemented efficiently by using the Fast Fourier Transforms (FFTs) at the transmitter and receiver. At the receiver, FFT reduces the channel response into a multiplicative constant on a tone-by-tone basis.

In 1996, a new wireless communication scheme based on combination of the concepts of MIMO and OFDM was proposed [12]. Since then, MIMO-OFDM becomes an emerging research topic. The signaling scheme and receiver design are categorized into two categories: spatial multiplexing (SM) and spatial diversity (SD) schemes. In the former system, different data streams are transmitted from different antennas simultaneously and detected based on their unique spatial signature at the receiver. This

implies the creation of parallel spatial channels to maximize the data rate.

For the next communication, the requirement of Quality of Services (QoS) becomes more important. Two techniques are fundamental for reliability: forward error correction (FEC) and automatic repeat request (ARQ). Hybrid ARQ (HARQ) is a variation of the ARQ error control method, which gives better performance than ordinary ARQ, particularly over wireless channels, at the cost of increased implementation complexity. When the coded data block is received, the receiver first decodes the error-correction code. If the channel quality is good enough, all transmission errors should be correctable, and the receiver can obtain the correct data block. If the channel quality is bad and not all transmission errors can be corrected, the receiver will detect this situation using the error-detection code, then the received coded data block is discarded and a retransmission is requested by the receiver, similar to ARQ.

WiMAX is defined as Worldwide Interoperability for Microwave Access by the WiMAX Forum [13], [14]. The Forum describes WiMAX as "a standards-based technology enabling the delivery of last mile wireless broadband access as an alternative to cable and DSL." WiMAX uses the advanced techniques such as multi-channel scalable OFDM, HARQ, FEC, MIMO and other complementary technologies as are part of WiMAX. WiMAX is designated as the metropolitan area network (MAN) technology that can connect IEEE 802.11 (Wi-Fi) hotspots with each other and to other parts of the Internet and provide a wireless alternative to cable and DSL for last mile broadband access. It is also anticipated that WiMAX will allow inter-penetration for broadband service provision of VoIP, video, and Internet access—simultaneously.

In principle, the MIMO technologies can provide not only the antenna gain for

interference suppression, but also various point-to-point link profits for covering wider service regions and improving various QoS. High-speed data service in WMANs through MIMO largely relies on rich-scattering and reliable background channel conditions. The radio environment inside a network, however, may be time-varying, and within which high-speed transmission may lead to high frame error rates. To sustain good link services, adaptive modulation techniques are proposed to dynamically adjust transmission parameters based on the near instantaneous channel state information (CSI) [9],[10] to ease channel impairments. Also, most wireless communication transceivers have built-in modules for supporting PHY layer data processing and MAC layer resource management. As a result, cross-layer processing that exploits the joint resource for more efficient PHY layer designs and more effective MAC protocol setups will become an important issue. In this thesis, we will attempt to develop an adaptive wireless transceiver that can take advantages of the existing system jointly to effectively exploit the available degrees of freedom in the wireless communication systems. Besides, an adaptive wireless transceiver which employs smart antenna and spatial multiplexing techniques is proposed to overcome the wireless channel impairments.

This thesis is organized as follows. In Chapter 2, we describe the general data model and channel capacity of a MIMO communication link. Spatial multiplexing technique is also presented to provide a preliminary overview. In Chapter 3, we introduce the principle of algorithm of reliability-based HARQ (RB-HARQ) which is based on low density parity-check (LDPC) codes [15]-[20]; Moreover, we propose an adaptive algorithm which can accordingly choose the sizes of retransmissions under the QoS constraints such as the maximum number of retransmissions allowed per packet and packet loss probability. In Chapter 4, we develop the cross-layer design by

combining AMC at PHY layer with several MAC protection strategies; We propose an adaptive MIMO-OFDM transceiver architecture along with a specifically designed loading procedure to dynamically adjust the transmission parameters such as retransmission size, number of retransmission, modulation order and transmit power over spatial and frequency channels, according to the instantaneous channel statistics, to meet the target QoS. Finally, Chapter 5 gives concluding remarks of this thesis and leads the way to some potential future works.



# Chapter 2

## Overview of IEEE 802.16 System

IEEE 802.16 is a wireless communication system that provides high-throughput broadband connections. IEEE 802.16 can be used for a number of applications, including "last mile" broadband connections, hotspots and cellular backhaul, and high-speed enterprise connectivity for business. IEEE 802.16 uses several advanced techniques in PHY layer, such as OFDM, and MIMO. OFDM is one of the most promising PHY layer technologies for high data rate wireless communications due to its robustness to frequency selective fading, high spectral efficiency, and low computational complexity. OFDM can be used in conjunction with a MIMO transceiver to increase the diversity gain and/or the system capacity by exploiting spatial domain. Because the OFDM system effectively provides numerous parallel narrowband channels, MIMO-OFDM is considered a key technology in emerging high-data rate systems such as IEEE 802.16, IEEE 802.11n, and 4G. In this chapter, we introduce the basic ideas of the MIMO-OFDM systems and key features of IEEE 802.16.

### 2.1 Review of MIMO-OFDM System

MIMO-OFDM is considered a key technique in high-data rate systems. In this section, we introduce the basic ideas and key features of MIMO-OFDM systems.

## 2.1.1 OFDM : Concept and Technique

OFDM can be regarded as either a modulation or a multiplexing technique. The basic concept of OFDM is to split a high rate data stream into a number of lower rate streams that are transmitted simultaneously over subcarriers. In order to eliminate the effect of inter-symbol interference (ISI), a guard time is appended to each OFDM symbol. The guard time is chosen to be larger than the maximum delay spread such that the current OFDM symbol never hears the interference from the previous one. However, this method will cause the inter-carrier interference (ICI) due to the loss of orthogonality between subcarriers. To solve this problem, OFDM symbols are cyclically extended in the guard time to introduce cyclic prefix (CP). This ensures that the delayed replicas of an OFDM symbol always have an integer number of cycles within the FFT interval. As a result, CP resolves both ISI and ICI problems caused by multipath, as long as the delay spread of channel is smaller than the length of CP. Besides, adding CP makes the transmitted OFDM symbol appear periodic, and the linear convolution process of the transmitted OFDM symbols (containing CP) with channel impulse response will be translated into a circular convolution one. According to discrete-time linear system theory, this circular convolution is equivalent to multiplying the frequency response of the OFDM symbol with the channel's frequency response. This property can be demonstrated as follows:

Assuming that the channel length is smaller than  $N_{cp}$  (number of samples in CP), we can express the received data vector  $\mathbf{y}$  as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \boldsymbol{\eta} \quad (2.1)$$



$$\underbrace{\begin{bmatrix} y_{N_c-1} \\ \vdots \\ y_0 \end{bmatrix}}_{\mathbf{y}} = \underbrace{\begin{bmatrix} h_0 & h_1 & \cdots & h_{N_{cp}} & 0 & \cdots & 0 \\ 0 & h_0 & h_1 & \cdots & h_{N_{cp}} & 0 & \vdots \\ 0 & 0 & \ddots & \ddots & \cdots & \ddots & 0 \\ 0 & \cdots & 0 & h_0 & h_1 & \cdots & h_{N_{cp}} \end{bmatrix}}_{\mathbf{H}} \underbrace{\begin{bmatrix} x_{N_c-1} \\ \vdots \\ x_0 \\ x_{-1} \\ \vdots \\ x_{-N_{cp}} \end{bmatrix}}_{\mathbf{x}} + \underbrace{\begin{bmatrix} \eta_{N_c-1} \\ \vdots \\ \eta_0 \end{bmatrix}}_{\boldsymbol{\eta}} \quad (2.2)$$

When we use singular value decomposition (SVD), we have

$$\mathbf{H} = \mathbf{F}\mathbf{\Lambda}\mathbf{M}^H \quad (2.3)$$

where  $\mathbf{F}\mathbf{F}^H = \mathbf{I}$  and  $\mathbf{M}\mathbf{M}^H = \mathbf{I}$ . If we let  $\mathbf{x} = \mathbf{M}\mathbf{X}$  and  $\mathbf{Y} = \mathbf{F}^H\mathbf{y}$ , then we can get

$$\mathbf{Y} = \mathbf{F}^H\mathbf{y} = \mathbf{F}^H(\mathbf{H}\mathbf{x} + \boldsymbol{\eta}) = \mathbf{F}^H\mathbf{H}\mathbf{M}\mathbf{X} + \underbrace{\mathbf{N}}_{\mathbf{F}^H\boldsymbol{\eta}} = \mathbf{\Lambda}\mathbf{X} + \mathbf{N} \quad (2.4)$$

It is interesting to note that when the guard period contains a CP, that is,  $x_{-i} = x_{N-i}$  for  $i = 1, \dots, N_{cp}$ , Equation (2.2) can be rewritten in a more compact matrix form

$$\begin{bmatrix} y_{N_c-1} \\ \vdots \\ y_0 \end{bmatrix} = \begin{bmatrix} h_0 & h_1 & \cdots & h_{N_{cp}} & 0 & \cdots & 0 \\ 0 & h_0 & h_1 & \cdots & h_{N_{cp}} & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & h_0 & h_1 & \cdots & h_{N_{cp}} \\ h_{N_{cp}} & 0 & \cdots & 0 & h_0 & \cdots & h_{N_{cp}-1} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ h_1 & \cdots & h_{N_{cp}} & 0 & \cdots & 0 & h_0 \end{bmatrix} \begin{bmatrix} x_{N_c-1} \\ \vdots \\ x_0 \end{bmatrix} + \begin{bmatrix} \eta_{N_c-1} \\ \vdots \\ \eta_0 \end{bmatrix}, \quad (2.5)$$

and  $\mathbf{H}$  becomes the so called ‘‘circulant matrix’’ and has the property that  $\mathbf{H} = \mathbf{Q}^H\mathbf{\Lambda}\mathbf{Q}$ ,

where  $\mathbf{Q}$  is a discrete Fourier transform (DFT) matrix with  $kl$ th entry as

$$\mathbf{Q}_{kl} = \frac{1}{\sqrt{N_c}} e^{-j2\pi\frac{kl}{N_c}} \quad (2.6)$$

and the transformed symbol is  $\mathbf{x} = \mathbf{Q}^H\mathbf{X}$  (inverse DFT (IDFT) of  $\mathbf{x}$ ). Thus  $\mathbf{X}$  can be interpreted as symbols in the frequency domain. At the receiver, we have the received data  $\mathbf{y}$  being transformed to  $\mathbf{Y}$ .

$$\begin{aligned}
\mathbf{Y} &= \mathbf{Q}^H \mathbf{y} = \mathbf{Q}^H (\mathbf{H}\mathbf{x} + \eta) \\
&= \underbrace{\mathbf{Q}^H \mathbf{H} \mathbf{Q}^H}_{\Lambda} \mathbf{X} + \underbrace{\mathbf{Q}^H}_{\mathbf{N}} \eta \\
&= \Lambda \mathbf{X} + \mathbf{N}
\end{aligned} \tag{2.7}$$

Now, we can realize that by using CP, the OFDM modulation is equivalent to multiplying the frequency domain signals of the OFDM symbol (that is,  $\mathbf{X}$ ) with the channel's frequency response  $\Lambda$ .

Broadband transmission over multipath channels usually exhibits frequency selective fading. Since data rate requirements can be expected to increase even further in the future, this effect is likely to amplify. In OFDM, frequency diversity can be realized through coding and interleaving across subcarriers. Because information bits are separated over many subcarriers, the impairment of fading occurring at particular frequency tones can be mitigated. As a consequence, in the coded OFDM systems the presence of frequency selective fading actually saves the frequency tones at fading. Depending on the coding rate and interleaving depth, gains can be achieved at locations experiencing significant delay spread.

It can be concluded that OFDM is a powerful modulation technique that increases bandwidth efficiency and simplifies the removal of distortion due to a multipath channel. Advances in FFT algorithm enable OFDM to be efficiently implemented in hardware, even for a large number of subcarriers. The key advantages of OFDM transmission are summarized as follows:

1. OFDM deals with multipath delay channels in an efficient way. The implementation complexity is significantly lower than that of a single carrier system with an equalizer.
2. OFDM has a long symbol period (compared to an equal data-rate single-carrier system) that allows OFDM to be more robust against impulse noise.

- OFDM supports dynamic bit loading that enable different subcarriers to use different modulation modes depending on the channel characteristic or the noise level. Therefore, improved performance can be achieved in this systematic way.

## 2.1.2 MIMO : Concept and Technique

MIMO contains two important techniques: SM and diversity. Figure 2.1 shows the diagram of a MIMO wireless transmission system. SM is a technique that yields an increased bit rate by using multiple antennas at both end of the wireless link [1]-[3]. This increase comes at no extra bandwidth and power consumption. However, such a technique calls for an efficient way to map the transmit signals to individual antenna elements. At the receiver, the individual data streams are separated and demultiplexed to yield the original transmitted signals, as illustrated in Figure 2.2. The separation is made possible by the fact that the rich multipath contributes to lower correlation between MIMO channel coefficients, and hence creates a desirable coefficient matrix condition (i.e., full rank and low condition number) to resolve  $N_t$  unknowns from a linear system of  $N_r$  equations. In the following, we will introduce two SM schemes.

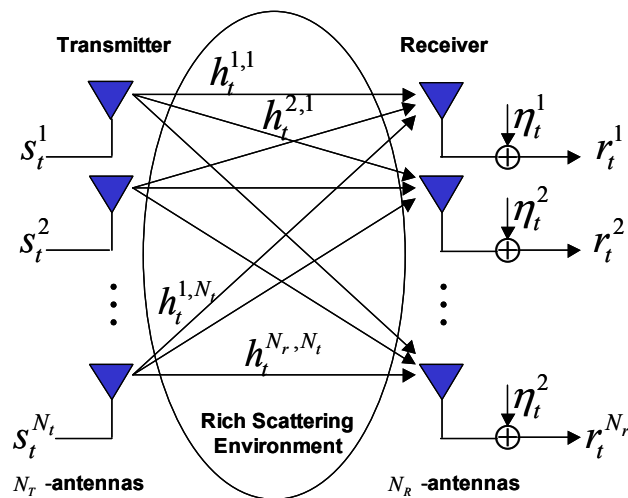


Figure 2.1: Diagram of a MIMO wireless transmission system

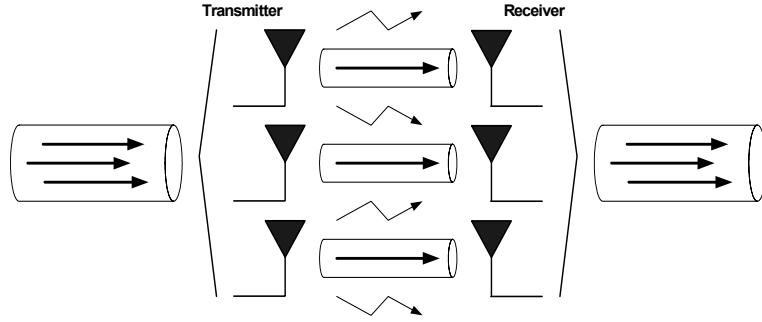


Figure 2.2: Illustration of a spatial multiplexing system

### 2.1.2.1 Diagonal Bell Labs' Layered Space-Time

The Layered Space-Time processing concept was first introduced by Foschini [1] at Bell Labs. The first version, Diagonal Bell Labs' Layered Space-Time (D-BLAST), utilizes multiple antenna arrays at both the transmitter and receiver, and an elegant diagonally-layered coding structure in which code blocks are dispersed across diagonals in space-time. The encoding and decoding procedures are described as follows:

- **Encoding:**

Considering a system equipped with  $N_t$  transmit and  $N_r$  receive antennas, the encoder applies the space-time codes to the input to generate a semi-infinite matrix  $\mathbf{C}$  of  $N_t$  rows to be transmitted. Figure 2.3 shows the encoding scheme, where  $c_\tau^k$ , representing an element in the  $k$ th row and  $\tau$ th column of  $\mathbf{C}$ , is transmitted by the  $k$ th transmit antenna at time  $\tau$ . The data received at time  $\tau$  by the  $l$ th receive antenna is  $r_\tau^l$ , which contains a superposition of  $c_\tau^k$ ,  $k = 1, 2, \dots, N_t$ , and an AWGN noise component. Each subsequence is encoded using a conventional 1-D constituent code with low decoding complexity.

- **Decoding:**

At any instance  $\tau$ , the received data vector is  $\mathbf{r}_\tau = \mathbf{H}_\tau \mathbf{c}_\tau + \boldsymbol{\eta}_\tau$ . The decoding

task is to determine  $\mathbf{c}_\tau = [c_\tau^1, c_\tau^2, \dots, c_\tau^{N_t}]^T$  with the only available information being  $\mathbf{r}_\tau$  and  $\mathbf{H}_\tau$ . The D-BLAST uses a repeated process of interference suppression, symbol detection and interference cancellation for decoding all symbols,  $c_\tau^{N_t}, c_\tau^{N_t-1}, \dots, c_\tau^1$ . Such decoding process could be expressed in a general form:

Let the QR decomposition of  $\mathbf{H}_\tau$  be  $\mathbf{Q}_\tau \mathbf{R}_\tau$ , where  $\mathbf{Q}_\tau$  is an  $N_r \times N_r$  unitary matrix and  $\mathbf{R}_\tau$  is an  $N_r \times N_t$  upper triangular matrix. We modify the received data to get

$$\begin{aligned} \mathbf{y}_\tau &= \mathbf{Q}_\tau^H \mathbf{r}_\tau = \mathbf{Q}_\tau^H \mathbf{H}_\tau \mathbf{c}_\tau + \mathbf{Q}_\tau^H \boldsymbol{\eta}_\tau = \underbrace{\mathbf{Q}_\tau^H \mathbf{Q}_\tau}_{\mathbf{I}_{N_r}} \mathbf{R}_\tau \mathbf{c}_\tau + \underbrace{\mathbf{Q}_\tau^H \boldsymbol{\eta}_\tau}_{\tilde{\boldsymbol{\eta}}_\tau} \\ &= \mathbf{R}_\tau \mathbf{c}_\tau + \tilde{\boldsymbol{\eta}}_\tau \end{aligned} \quad (2.8)$$

where

$$\mathbf{y}_\tau = \begin{bmatrix} y_\tau^1 \\ y_\tau^2 \\ \vdots \\ y_\tau^{N_r} \end{bmatrix}, \mathbf{R}_\tau = \begin{bmatrix} r_\tau^{1,1} & r_\tau^{1,2} & \dots & r_\tau^{1,N_t} \\ 0 & r_\tau^{2,2} & \dots & r_\tau^{2,N_t} \\ 0 & 0 & \ddots & \vdots \\ \vdots & 0 & \ddots & r_\tau^{N_t,N_t} \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}, \tilde{\boldsymbol{\eta}}_\tau = \begin{bmatrix} \tilde{\eta}_\tau^1 \\ \tilde{\eta}_\tau^2 \\ \vdots \\ \tilde{\eta}_\tau^{N_r} \end{bmatrix} \quad (2.9)$$

Since  $\mathbf{R}_\tau$  is upper triangular,

$$y_\tau^k = r_\tau^{k,k} c_\tau^k + \tilde{\eta}_\tau^k + \{\text{contribution from } c_\tau^{k+1}, c_\tau^{k+2}, \dots, c_\tau^{N_t}\} \quad (2.10)$$

Now, we can figure out that the interference from  $c^l$ ,  $l < k \leq N_t$ , are first suppressed in  $y^k$  and the residual interference terms in Equation (2.10) can be cancelled by the available decisions  $\hat{c}_\tau^{k+1}, \hat{c}_\tau^{k+2}, \dots, \hat{c}_\tau^{N_t}$ . Assuming all these decisions are correct, the present decision variable is

$$\tilde{c}_\tau^k = r_\tau^{k,k} c_\tau^k + \tilde{\eta}_\tau^k, \quad k = 1, 2, \dots, N_t \quad (2.11)$$

The relationship between  $c^k$  and  $\tilde{c}^k$  in Equation (2.11) can be interpreted as the input and output of a single-input and single output channel with the channel power



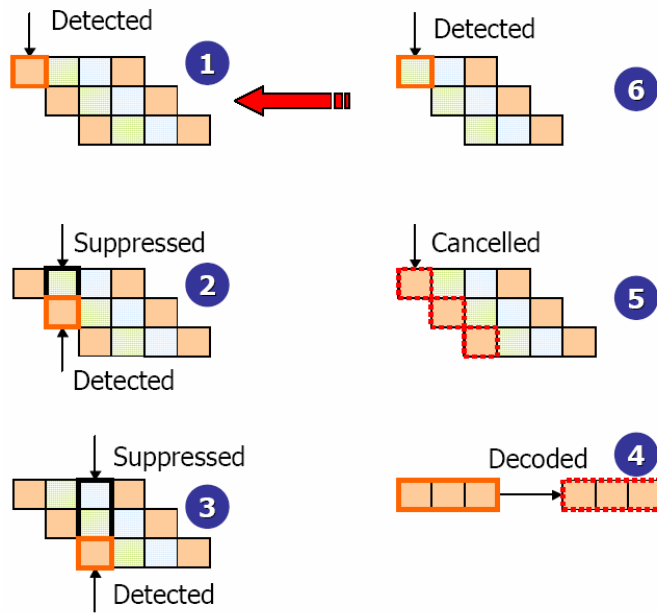


Figure 2.4: Diagonal layered space-time decoding with  $N_t = 3$

### 2.1.2.2 Vertical Bell Labs' Layered Space-Time

The diagonal approach suffers from certain implementation complexities that make it inappropriate for initial implementation. Therefore, Foschini proposed another low-complexity version of detecting the symbols transmitted synchronously over antennas, that is, V-BLAST [3]. The “V” here stands for the vertical vector mapping process, which differs from the diagonal form in D-BLAST. In V-BLAST, no inter-substream coding, or coding of any kind, is required, though conventional coding of the individual substreams may certainly be applied. In [4], a vertical-and-horizontal coding structure along with iterative detection and decoding (IDD) was promoted and showed to significantly improve the performance with limited complexity.

Figures 2.3 and 2.5 display the typical encoding and decoding steps in V-BLAST, respectively. The decoding process can also be interpreted via the general form (QR decomposition) as mentioned in decoding D-BLAST. In each step  $I$ , the signals from all but one transmit antenna are eliminated using interference suppression and

cancellation with already detected signals. Following the data model in D-BLAST, at a given time instant  $\tau$ , let  $\tilde{\mathbf{H}}^{i=1} = \mathbf{H}_\tau$  and  $\tilde{\mathbf{r}}^{i=1} = \mathbf{r}_\tau$  at the first decoding step. In each step  $i$ , the nulling matrix  $\mathbf{G}^i$  is calculated as the pseudo-inverse of  $\tilde{\mathbf{H}}^i$

$$\begin{aligned}\mathbf{G}^i &= (\tilde{\mathbf{H}}^i)^+ \\ &= \left( (\tilde{\mathbf{H}}^i)^H \tilde{\mathbf{H}}^i \right)^{-1} (\tilde{\mathbf{H}}^i)^H\end{aligned}\quad (2.12)$$

Each row of  $\mathbf{G}^i$  can be used to null all but the  $i$ th desired signal. Instead of choosing an arbitrary layer to be detected first, it was suggested to start with the layer showing the biggest post-detection signal-to-noise ratio (SNR) to efficiently reduce the error propagation effect [3]. This corresponds to choosing the row of  $\mathbf{G}^i$  with the minimum norm and defining the corresponding row,  $\mathbf{w}_{k_i}^T$ , as the nulling vector at this step:

$$k_i = \arg \min_{j \notin \{k_1, \dots, k_{i-1}\}} \|(\mathbf{G}^i)_j\|^2 \quad (2.13)$$

$$\mathbf{w}_{k_i} = (\mathbf{G}^i)_{k_i}^T \quad (2.14)$$

Multiplying  $\mathbf{w}_{k_i}$  with the vector of received data  $\tilde{\mathbf{r}}^i$  suppresses all layers but the one transmitted from antenna  $k_i$  and we get a soft decision value

$$\bar{c}_\tau^{k_i} = \mathbf{w}_{k_i}^T \tilde{\mathbf{r}}_\tau^i \quad (2.15)$$

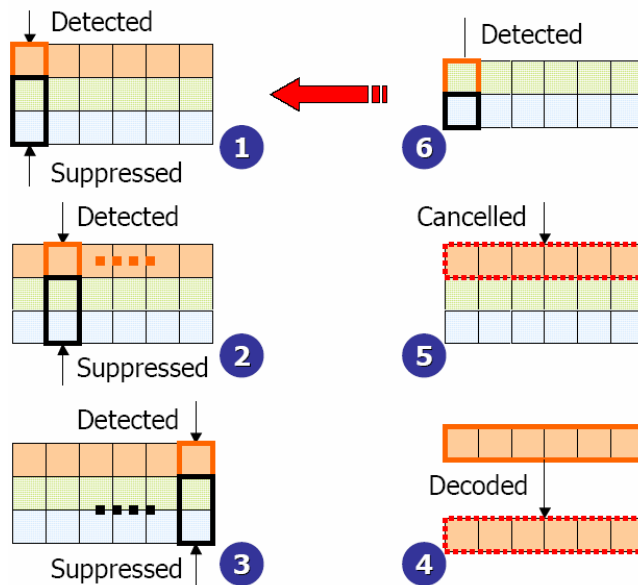


Figure 2.5: Vertical layered space-time decoding with  $N_t = 3$



Now the  $k_i$ th layer can be detected within the constellation set  $S$  that we use:

$$\hat{c}_\tau^{k_i} = \arg \min_{\tilde{x} \in S} \|\tilde{c} - \tilde{c}_\tau^{k_i}\|^2 \quad (2.16)$$

As soon as one layer is detected, we can improve the detection performance for the subsequent layers by subtracting the part of the detected signal from the received vector,

$$\tilde{\mathbf{r}}_\tau^{i+1} = \tilde{\mathbf{r}}_\tau^i - \hat{c}_\tau^{k_i} (\tilde{\mathbf{H}}^i)^{k_i} \quad (2.17)$$

where  $(\tilde{\mathbf{H}}^i)^{k_i}$  denotes the  $k_i$ th column of  $\tilde{\mathbf{H}}^i$ . After canceling out the signal from the  $k_i$ th transmit antenna, the channel matrix is reduced to

$$\tilde{\mathbf{H}}^{i+1} = (\tilde{\mathbf{H}}^i)^{\bar{k}_i} \quad (2.18)$$

where the notation  $(\tilde{\mathbf{H}}^i)^{\bar{k}_i}$  denotes the matrix obtained by zeroing columns  $k_1, k_2, \dots, k_i$  of  $\tilde{\mathbf{H}}^i$ . Since we decrease the number of layers to be nulled out in the next step by one, the diversity gain is increased by one at each step (from  $(N_r - N_t + i)$  to  $(N_r - N_t + i + 1)$ ). This can be proven by the Cauchy-Schwartz inequality [3]. The full Zero-Forcing V-BLAST detection algorithm can be summarized as follows:

*Initialization:*

$$i \leftarrow 1$$

$$\mathbf{G}^1 = (\tilde{\mathbf{H}}^1)^+$$

$$k_1 = \arg \min_j \|(\mathbf{G}^1)_j\|^2$$

*Recursion:*

$$\mathbf{w}_{k_i} = (\mathbf{G}^i)_{k_i}^T$$

$$\tilde{c}^{k_i} = \mathbf{w}_{k_i}^T \tilde{\mathbf{r}}^i$$

$$\hat{c}^{k_i} = Q(\tilde{c}^{k_i}), \quad Q(\cdot) \text{ denotes the slicing operation}$$

$$\begin{aligned}
\mathbf{r}^{i+1} &= \mathbf{r}^i - \hat{c}^{k_i} (\tilde{\mathbf{H}}^i)^{k_i} \\
\tilde{\mathbf{H}}^{i+1} &= (\tilde{\mathbf{H}}^i)^{\overline{k_i}} \\
\mathbf{G}^{i+1} &= (\tilde{\mathbf{H}}^{i+1})^+ \\
k_{i+1} &= \arg \min_{j \notin \{k_1, \dots, k_i\}} \|(\mathbf{G}^{i+1})_j\|^2 \\
i &\leftarrow i + 1
\end{aligned}$$

The post-processing SNR for the  $k_i$ th detected component of  $\mathbf{c}$  is

$$\rho_{k_i} = \frac{\langle |c^{k_i}|^2 \rangle}{\sigma^2 \|\mathbf{w}_{k_i}\|^2} \quad (2.19)$$

where the expectation value in the numerator is taken over the constellation set  $\mathcal{S}$ .

Another way to improve detection performance especially for mid-range SNR values is to replace the ZF nulling matrix by the more powerful MMSE one [3]:

$$\mathbf{G}^i = \left( (\tilde{\mathbf{H}}^i)^H \tilde{\mathbf{H}}^i + \frac{1}{SNR} \mathbf{I} \right)^{-1} (\tilde{\mathbf{H}}^i)^H \quad (2.20)$$

In this case, in addition to nulling out the interference, the noise level on the channel is taken into account. Thus, the SNR has to be estimated at the receiver. Figure 2.5 shows the typical decoding procedure in V-BLAST.

The D-BLAST code blocks are organized along diagonals in space-time. It is this coding that leads to D-BLAST's higher spectral efficiencies for a given number of transmit and receive antennas.

### 2.1.3 V-BLAST Based OFDM

Due to the scarcity of radio spectrum, high spectral efficiency becomes a must-have requirement that encourages modern wireless modems toward this trend. An evolution of the V-BLAST supporting OFDM modulation seems to be a solution that

can dramatically increase the capacity of wireless radio links with no additional power and bandwidth consumption. The core idea in such scheme is that with the aid of OFDM, the whole detection problem in MIMO-OFDM would be translated into  $N_c$  parallel sub-problems.

In the transmitter, as shown in Figure 2.6, a traditional 1-D channel encoder is used to encode the information bits. These coded bits are then mapped on the symbols of constellation adopted for each subcarrier. At a given time slot  $n$ ,  $N_c \times N_t$  bit streams  $\{c_i[n, k] : k = 0, 1, \dots, N_c\}$  for  $i = 1, 2, \dots, N_t$  are fed to the IFFT at the  $i$ th transmit antenna on the  $k$ th subcarrier to generate the  $n$ th transmitted OFDM symbols from the  $i$ th transmit antenna.

At the receiver side, as shown in Figure 2.7, receive antennas  $1 - N_r$  will receive the radiate signal from transmit antennas  $1 - N_t$ , where the V-BLAST requires  $N_r \geq N_t$  to ensure its proper working. The received data at each receive antenna will then pass through a FFT with the removal of the CP. The FFT output, at the receive antenna  $j$ , is a set of  $N_c$  signals, one for each frequency subcarrier, expressed as

$$r_j[n, k] = \sum_{i=1}^{N_t} H_{j,i}[n, k]c_i[n, k] + \eta_j[n, k] \quad \forall k = 1, 2, \dots, N_c \quad (2.21)$$

where  $H_{i,j}[n, k]$  is the flat fading coefficient representing the channel gain from the transmit antenna  $i$  to the receive antenna  $j$  at frequency  $k$ , and  $\eta_j[n, k]$  denotes the additive complex Gaussian noise at the receiver antenna  $j$  and frequency  $k$  with two sides power spectral density  $N_0 / 2$  per dimension and uncorrelated for different  $n$ 's,  $k$ 's, and  $j$ 's.

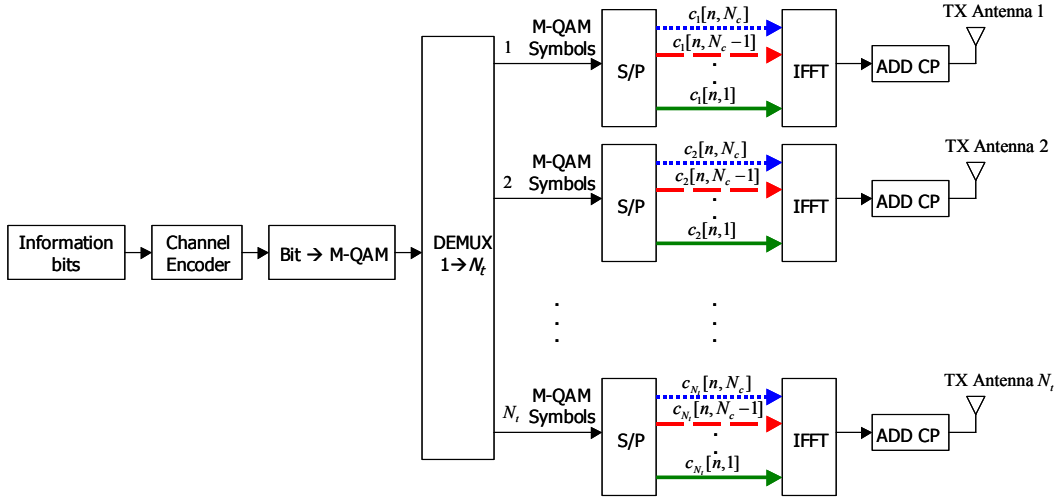


Figure 2.6: V-BLAST based MIMO-OFDM transmitter architecture

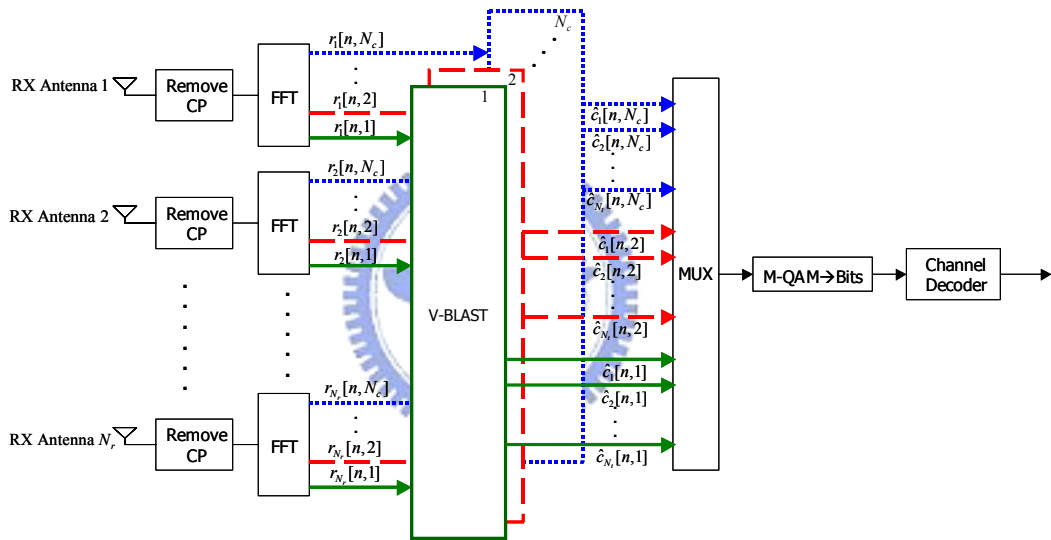


Figure 2.7: V-BLAST based MIMO-OFDM receiver architecture.

## 2.2 WiMAX Overview

WiMAX is defined as **Worldwide Interoperability for Microwave Access** by the WiMAX Forum, formed in April 2001 to promote conformance and interoperability of the standard IEEE 802.16 [13]-[14]. The Forum describes WiMAX as "a standards-based technology enabling the delivery of last mile wireless broadband access as an alternative to cable and DSL." The WiMAX Forum is "the exclusive organization dedicated to certifying the interoperability of BWA products, the WiMAX

Forum defines and conducts conformance and interoperability testing to ensure that different vendor systems work seamlessly with one another."

## 2.2.1 Review of IEEE 802.16 PHY

A recent addition to the WiMAX standard is underway which will add full mesh networking capability by enabling WiMAX nodes to simultaneously operate in "subscriber station" and "base station" mode. This will blur that initial distinction and allow for widespread adoption of WiMAX based mesh networks and promises widespread WiMAX adoption. WiMAX/802.16's use of OFDMA and scheduled MAC allows wireless mesh networks to be much more robust and reliable. These differences between and evolution of Wi-Fi and WiMAX mesh networks could serve as a separate Wikipedia topic.

The original WiMAX standard, IEEE 802.16, specifies WiMAX in the 10 to 66 GHz range. 802.16a, updated in 2004 to 802.16-2004, added support for the 2 to 11 GHz range, of which most parts are already unlicensed internationally and only very few still require domestic licenses. Most business interest will probably be in the 802.16-2004 standard, as opposed to licensed frequencies. The WiMAX specification improves upon many of the limitations of the Wi-Fi standard by providing increased bandwidth and stronger encryption. It also aims to provide connectivity between network endpoints without direct line of sight in some circumstances. The details of performance under non-line of sight (NLOS) circumstances are unclear as they have yet to be demonstrated. It is commonly considered that spectrum under 5–6 GHz is needed to provide reasonable NLOS performance and cost effectiveness for PtM (point to multi-point) deployments. WiMAX makes clever use of multi-path signals but does not defy the laws of physics.

A number of PHY considerations were taken into account for the target environment. At higher frequencies, line of sight is a must. This requirement eases the effect of multipath, allowing for wide channels, typically greater than 10 MHz in bandwidth. This gives IEEE 802.16 the ability to provide very high capacity links on both the uplink and the downlink. For sub 11 GHz non line of sight capability is a requirement. The original IEEE 802.16 MAC was enhanced to accommodate different PHYs and services, which address the needs of different environments. The standard is designed to accommodate either Time Division Duplexing (TDD) or Frequency Division Duplexing (FDD) deployments, allowing for both full and half-duplex terminals in the FDD case.

## **2.2.2 Review of IEEE 802.16 MAC**

The IEEE 802.16 media access controller (MAC) [14] is significantly different from that of IEEE 802.11 Wi-Fi MAC. In Wi-Fi, the MAC uses contention access—all subscriber stations wishing to pass data through an access point are competing for the AP's attention on a random basis. This can cause distant nodes from the AP to be repeatedly interrupted by less sensitive, closer nodes, greatly reducing their throughput. And this makes services, such as VoIP or IPTV which depend on a determined level of QoS difficult to maintain for large numbers of users.

By contrast, the IEEE 802.16 MAC is a scheduling MAC where the subscriber station only has to compete once (for initial entry into the network). After that it is allocated a time slot by the base station. The time slot can enlarge and constrict, but it remains assigned to the subscriber station meaning that other subscribers are not supposed to use it but take their turn. This scheduling algorithm is stable under overload and over-subscription (unlike IEEE 802.11). It is also much more bandwidth

efficient. The scheduling algorithm also allows the base station to control QoS by balancing the assignments among the needs of the subscriber stations.

The MAC was designed specifically for the PMP wireless access environment. It supports higher layer or transport protocols such as ATM, Ethernet or Internet Protocol (IP), and is designed to easily accommodate future protocols that have not yet been developed. The MAC is designed for very high bit rates (up to 268 mbps each way) of the truly broadband PHY layer, while delivering ATM compatible QoS; UGS, rtPS, nrtPS, and Best Effort.

The frame structure allows terminals to be dynamically assigned uplink and downlink burst profiles according to their link conditions. This allows a trade-off between capacity and robustness in real-time, and provides roughly a two times increase in capacity on average when compared to non-adaptive systems, while maintaining appropriate link availability.

The IEEE 802.16 MAC uses a variable length Protocol Data Unit (PDU) along with a number of other concepts that greatly increase the efficiency of the standard. Multiple MAC PDUs may be concatenated into a single burst to save PHY overhead. Additionally, multiple Service Data Units (SDU) for the same service may be concatenated into a single MAC PDU, saving on MAC header overhead. Fragmentation allows very large SDUs to be sent across frame boundaries to guarantee the QoS of competing services. And, payload header suppression can be used to reduce the overhead caused by the redundant portions of SDU headers.

The MAC uses a self-correcting bandwidth request/grant scheme that eliminates the overhead and delay of acknowledgements, while simultaneously allowing better QoS handling than traditional acknowledged schemes. Terminals have a variety of options available to them for requesting bandwidth depending upon the QoS and traffic

parameters of their services. They can be polled individually or in groups. They can steal bandwidth already allocated to make requests for more. They can signal the need to be polled, and they can piggyback requests for bandwidth.

### **2.2.3 Review of IEEE 802.16-2005**

IEEE 802.16-2005, approved December, 2005 (formerly named but still best known as 802.16e or Mobile WiMAX) [13]. The WiMAX mobility standard is an improvement on the modulation schemes stipulated in the original (fixed) WiMAX standard. It allows for fixed wireless and mobile Non Line of Sight (NLOS) applications primarily by enhancing the OFDMA (Orthogonal Frequency Division Multiple Access).

Many think that by stipulating a new modulation method called Scalable OFDMA (SOFDMA), 802.16-2005 will make the older 802.16-2004 which uses OFDM-256 obsolete. However, several manufacturers plan for a migration path from the older version of the standard to the more robust, mobile modulation scheme. In any case, manufacturers are working through the WiMAX Forum to achieve compatibility between similar system profiles.

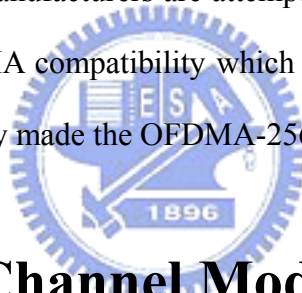
SOFDMA will improve upon OFDM-256 for NLOS applications by:

- Improving NLOS coverage by utilizing advanced antenna diversity schemes, and hybrid-Automatic Retransmission Request
- Increasing system gain by use of denser sub-channelization, thereby improving indoor penetration
- Introducing high-performance coding techniques such as Turbo coding and LDPC, enhancing security and NLOS performance



- Introducing downlink sub-channelization, allowing administrators to trade coverage for capacity or vice versa
- Improving coverage by introducing Adaptive Antenna Systems (AAS) and MIMO technology
- Eliminating channel bandwidth dependencies on sub-carrier spacing, allowing for equal performance under any RF channel spacing (1.25-14 MHz)
- Enhanced FFT algorithm can tolerate larger delay spreads, increasing resistance to multipath interference

SOFDMA and OFDMA-256 are not compatible so most equipment will have to be replaced. However, some manufacturers are attempting to provide a migration path for older equipment to SOFDMA compatibility which would ease the transition for those networks which have already made the OFDMA-256 investment.



## 2.3 MIMO Channel Model

We use IEEE 802.20 channel models in simulations [21]. In single-in single-out systems shall use the ITU model in simulations. The parameters are list in Table 2.1. Table 2.2 shows the SISO Channel Environment Parameters. We will describe a MIMO channel model that captures the above characteristics and that can be collapsed to an underlying SISO ITU channel mode by using a correlation matrix approach in simulations. The correlation matrices are only antenna system dependent.

Table 2.1: Summary of SISO link-level parameters for IEEE 802.20 channel models

Models		Case-A		Case-B		Case-C		Case-D	
PDP		Pedestrian-A		Vehicular-A		Pedestrian-B (Phase I)		Vehicular-B (Phase I)	
Number of Paths		4		6		6		6	
Relative Path power (dB)	Delay (ns)	0	0	0	0	0	0	-2.5	0
		-9.7	110	-1.0	310	-0.9	200	0	300
		-19.2	190	-9.0	710	-4.9	800	-12.8	8900
		-22.8	410	-10.0	1090	-8.0	1200	-10.0	12900
				-15.0	1730	-7.8	2300	-25.2	17100
				-20.0	2510	-23.9	3700	-16.0	20000
Speed (km/h)		3, 30, 120		30, 120, 250		3		30, 120, 250	

Table 2.2: Summary of SISO environment parameters for IEEE 802.20 channel models

Channel Scenario	Suburban Macro	Urban Macro	Urban Micro
Lognormal shadowing standard deviation	10dB	10dB	NLOS: 10dB LOS: 4dB
Pathloss model (dB), $d$ is in meters	$31.5 + 35\log_{10}(d)$	$34.5 + 35\log_{10}(d)$	NLOS: $34.53 + 38\log_{10}(d)$
			LOS: $30.18 + 26*\log_{10}(d)$

### 2.3.1 Correlation Channel Matrices

In the correlation matrix approach, the channel from any of the  $N$  transmit antennas to the  $M$  receive antenna elements is generated from  $M$  independent channels from that transmit antenna to the  $M$  receive antennas. That is, for any given

channel tap, we will have [21]

$$\mathbf{h}_i(\tau) = \mathbf{R}_r^{1/2} \cdot \mathbf{g}_i(\tau) \quad (2.22)$$

where  $\mathbf{h}_i(\tau)$  is the channel vector from the  $i$ -th transmitting antenna to the  $M$  receive antennas,  $\mathbf{g}_i(\tau)$  is the underlying independent Gaussian channel vector (i.e. it is the channel vector from the  $i$ -th transmitting antenna to the  $M$  receive antennas if the receive antennas were uncorrelated), and  $\mathbf{R}_r^{1/2}$  is the square root of the channel receive correlation matrix. Please note that the dimensions of  $\mathbf{h}_i(\tau)$ ,  $\mathbf{g}_i(\tau)$ , and  $\mathbf{R}_r^{1/2}$  are  $M \times 1$ ,  $M \times 1$ , and  $M \times M$ , respectively. In addition, we note that each ITU channel profile defines a number of taps with a corresponding tap delay and average tap power. The above description for the channel vector  $\mathbf{h}_i(\tau)$  is repeated for each channel tap. Moreover, please note that the underlying independent Gaussian channel vector  $\mathbf{g}_i(\tau)$  is completely different (i.e. independent) for each tap. Note that, when there is only one transmit antenna and one receive antenna,  $\mathbf{R}_r^{1/2}$  is simply 1 and the above reduces to the scalar ITU channel model.

In a similar fashion, let us now consider the channel from the  $N$  transmit antennas to any of the  $M$  receive antennas. The channel row-vector  $\mathbf{h}_j(\tau)$  corresponding to the channels from all the  $N$  transmit antennas to the  $j$ -th receive antenna is related to the underlying independent Gaussian channel row-vector  $\mathbf{g}_j(\tau)$  (i.e. the channel row-vector from all  $N$ -transmit antennas to the  $j$ -th receive antenna if the transmit antennas were uncorrelated) by [21]

$$\mathbf{h}_j(\tau) = \mathbf{g}_j(\tau) \cdot \mathbf{R}_t^{1/2} \quad (2.23)$$

where  $\mathbf{R}_t^{1/2}$  is the square root of the transmit array correlation matrix. We note that the dimensions of  $\mathbf{h}_j(\tau)$ ,  $\mathbf{g}_j(\tau)$ , and  $\mathbf{R}_t^{1/2}$  are  $1 \times N$ ,  $1 \times N$ , and  $N \times N$  respectively.

## 2.3.2 Generation of a MIMO Channel Using Correlation Matrix Approach

Some of the parameters that can be used in the correlation channel model are shown in Table 2.3. In order to generate a MIMO channel, we first need to have a pair of transmit and receive  $\mathbf{R}_t$  and  $\mathbf{R}_r$  correlation matrices. These are generated for each mobile station (MS) and base station (BS) based on the number of antennas, antenna spacing, number of clusters, power azimuth spectrum (PAS), azimuth spread (AS), and angle of arrival (AoA). In addition to the correlation matrices, we also need to specify an underlying ITU SISO model from Table 2.1 and choose the mobile speed.

As an example, a MIMO link with  $N$  transmit and  $M$  receive antennas can be then generated as follows: [21]

1. Generate  $N \cdot M$  SISO links based on the chosen ITU profiles as follows
  - a. Let  $A_1, A_2, \dots, A_K$  and  $\tau_1, \tau_2, \dots, \tau_K$  represent the power-delay profile for the specified ITU channel model
  - b. Generate  $K$  independent Rayleigh fading processes each having a Doppler spread  $f_d$  (function of the chosen mobile speed). The number of samples in each of the fading processes is given by the required number of symbols at the specified sampling rate
  - c. Scale the  $k$ -th Rayleigh process by  $P_k$  where

$$P_k^2 = \frac{A_k}{\sum_{k=1}^K A_k} \quad (2.24)$$

- d. Generate the pulse shaping matrix  $\tilde{\mathbf{G}}(\boldsymbol{\tau})$
- e. Compute the channels taps. This will result in  $L + 1$  Rayleigh fading processes where  $L + 1$  is the number of taps in the digital channel corresponding to the specified ITU channel model.

- Given the  $N \cdot M$  SISO links generated in step 1 above, each is described by  $L + 1$  processes, we define the following  $M \times N$   $i$ -th tap gain matrix for every channel sample (i.e. for every  $t$ ) [21]

$$\mathbf{H}_i(t) = \begin{pmatrix} h_{11}^{(i)}(t) & \dots & h_{1N}^{(n)}(t) \\ \vdots & \ddots & \vdots \\ h_{M1}^{(i)}(t) & \dots & h_{MN}^{(i)}(t) \end{pmatrix} \quad (2.25)$$

- Color the tap gain matrix by the receive and transmit correlation matrices as follows [21]

$$\hat{\mathbf{H}}_i(t) = \mathbf{R}_r^{1/2} \cdot \mathbf{H}_i(t) \cdot \mathbf{R}_t^{1/2}$$

The overall procedure for generating the channel matrices consists of three basic steps:

- Specify an environment, i.e., suburban macro, urban macro, or urban micro.
- Obtain the parameters to be used in simulations, associated with that environment.
- Generate the channel coefficients based on the parameters.

The following sections describe the details of overall procedure. Figure 2.8 provides a flow chart for generating channel coefficients.

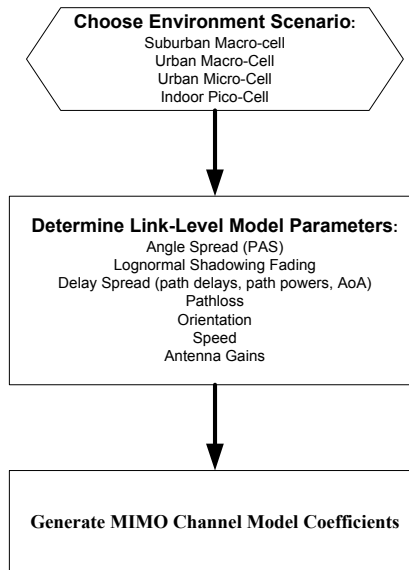


Figure 2.8: Flow chart for the generation of MIMO IEEE 802.20 Channel Models coefficients

Table 2.3: Summary of MIMO link-level parameters for IEEE 802.20 Channel Models

Models		Case-A		Case-B		Case-C		Case-D	
PDP		Pedestrian-A		Vehicular-A		Pedestrian-B		Vehicular-B	
Number of Paths		4		6		6		6	
Relative Path power (dB)	Delay (ns)	0	0	0	0	0	0	-2.5	0
		-9.7	110	-1.0	310	-0.9	200	0	300
		-19.2	190	-9.0	710	-4.9	800	-12.8	8900
		-22.8	410	-10.0	1090	-8.0	1200	-10.0	12900
				-15.0	1730	-7.8	2300	-25.2	17100
				-20.0	2510	-23.9	3700	-16.0	20000
Speed (km/h)		3, 30, 120		30, 120, 250		3		30, 120, 250	
Mobile Station	Topology	0.5 $\lambda$		0.5 $\lambda$		0.5 $\lambda$		0.5 $\lambda$	
	PAS	Fixed AoA for LOS component, remaining power has 360 degree uniform PAS.		360 degree uniform PAS		RMS angle spread of 35 degrees per path with a Laplacian distribution		RMS angle spread of 35 degrees per path with a Laplacian distribution	
	DoT (°)	0		-22.5		-22.5		22.5	
	AoA (°)	22.5 / 67.5		67.5 (all paths)		67.5 (all paths)		67.5 (all paths)	
Base Station	Topology	Reference: ULA with 0.5 $\lambda$ -spacing or 4 $\lambda$ -spacing or 10 $\lambda$ -spacing							
	PAS	Laplacian distribution with RMS angle spread of 2 degrees or 5 degrees, per path depending on AoA/AoD							
	AoD/AoA(°)	50° for 2° RMS angle spread per path or 20° for 5° RMS angle spread per path							

## 2.4 Computer Simulations

In this section, we simulate the V-BLAST performance both for the ideal and realistic case. We define the relation between SNR and  $E_b/N_0$  at each receive antenna as follows:

$$\text{SNR} = \frac{\text{signal power}}{\text{noise power}} = \frac{\frac{E_s}{T_s}}{N_0 B} = \frac{\frac{E_b \cdot N_t \cdot M}{T_s}}{N_0 \frac{1}{T_s}} = \frac{E_b}{N_0} \cdot (N_t \cdot M) \quad (2.28)$$

where  $E_s$  is the symbol energy,  $T_s$  is the symbol duration,  $B$  is the system bandwidth and  $M$  is the modulation order. Throughout the following simulations, the system transmit power is normalized to 1, and hence the noise power corresponding to a specific  $E_b/N_0$  is generated by

$$\text{noise power} = \frac{N_0}{E_b \cdot N_t \cdot M} \quad (2.29)$$

Figure 2.9 shows the BER performance of the  $(N_t, N_r) = (4, 4)$  ZF V-BLAST system with ideal detection and cancellation. It is obvious that in the ideal case, the diversity gain increases as the number of effective transmit antennas decreases. However, as shown in Figure 2.10, the realistic V-BLAST system suffers from error propagation and hence the diversity gain degrades. In Figure 2.11, we compare two equal rate V-BLAST systems. It is interesting to see that the system with fewer transmit antennas will outperform the one with more transmit antennas in the BER performance. This phenomenon hints that given a MIMO channel and some transmit power budget, we can improve the MIMO system performance by simply adjusting transmission parameters at no cost of transmission rate. So, it strongly motivates us to incorporate the concept of adaptive modulation in MIMO, which will be described in Chapter 4.

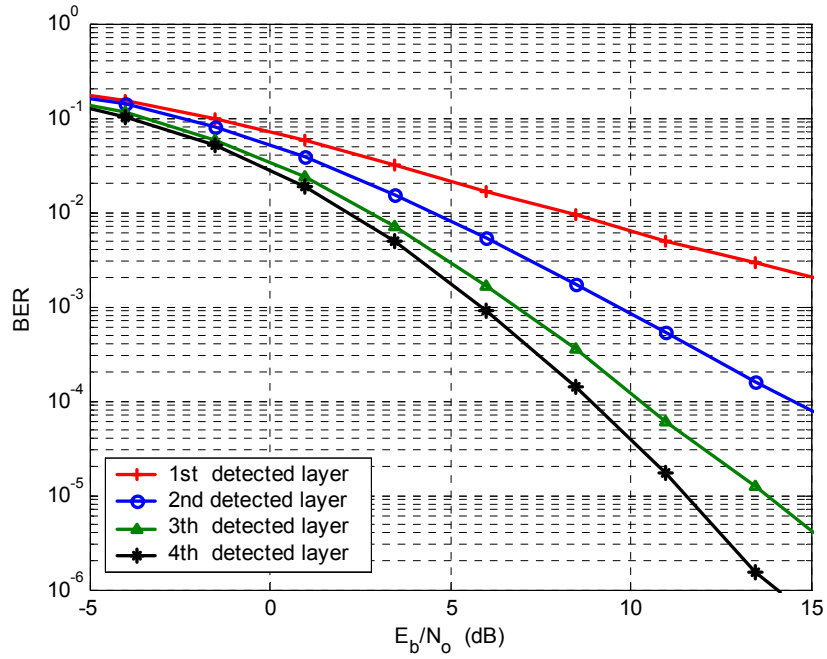


Figure 2.9: ZF V-BLAST performance with ideal detection and cancellation. QPSK modulation is used.  $(N_t, N_r) = (4, 4)$

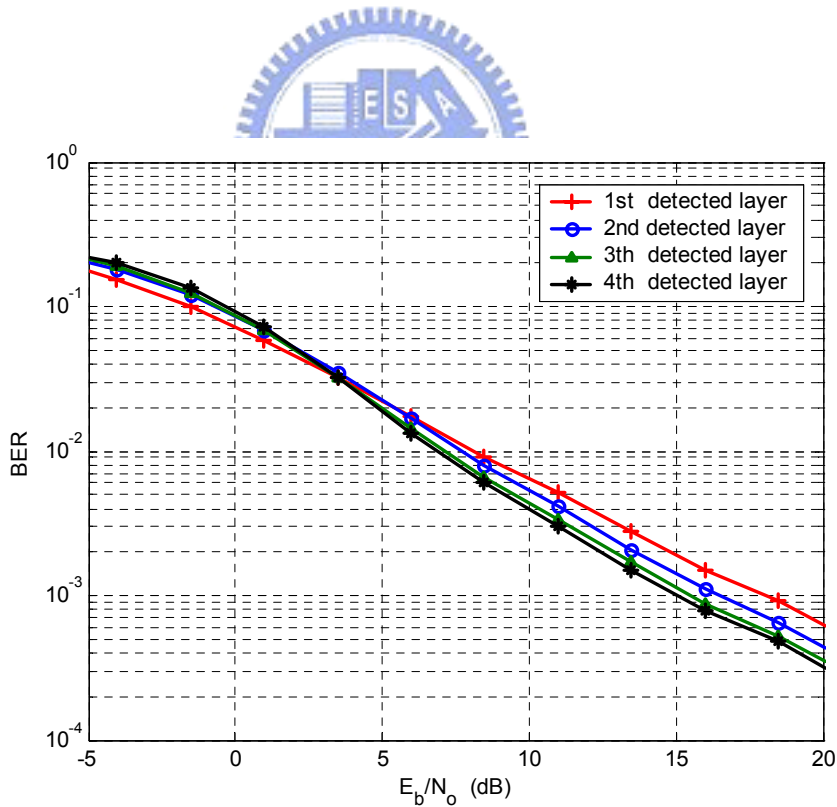


Figure 2.10: ZF V-BLAST performance with error propagation. QPSK modulation is used.  $(N_t, N_r) = (4, 4)$



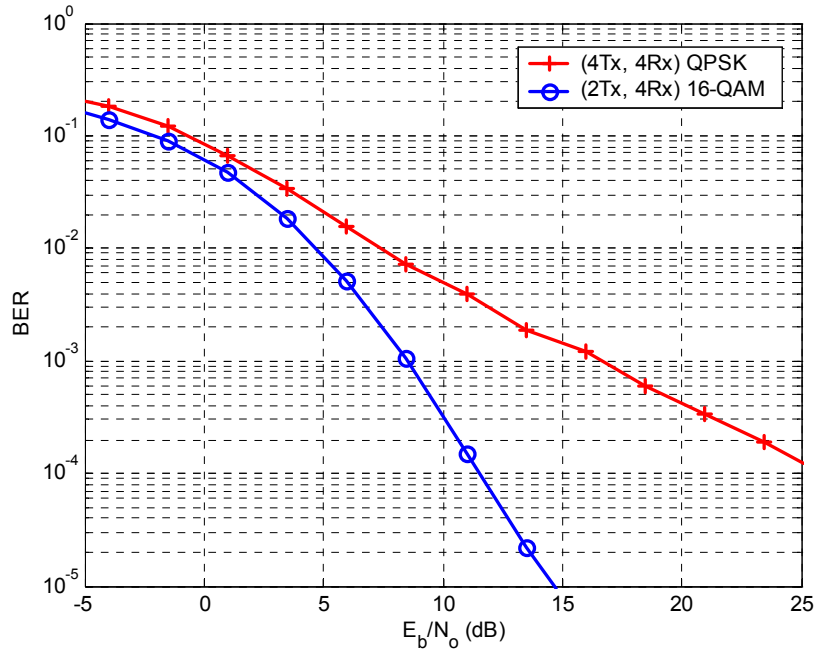


Figure 2.11: Comparison of ZF V-BLAST  $(N_t, N_r) = (4, 4)$  with QPSK modulation and  $(N_t, N_r) = (2, 4)$  with 16-QAM modulation.



# Chapter 3

## Reliability-Based Hybrid ARQ Scheme with LDPC Codes

For wireless communications on fading channel, two techniques are fundamental for reliability: FEC and ARQ. HARQ scheme that combines ARQ and FEC can offer a performance superior to either scheme. RB-HARQ, a new attractive ARQ scheme, could bring significant performance gain while only a few bits need to be retransmitted. RB-HARQ consists a soft-input, soft-output (SISO) decoder, so we will introduce the SISO decoder which is based on LDPC codes first, then introduce the basis and key fractures of the HARQ schemes and propose an adaptive algorithm of RB-HARQ scheme.

### 3.1 Review of LDPC Codes

LDPC codes were introduced along with an iterative probability-based decoding algorithm by Gallager in the early 1960's [22]. These codes were constructed using sparse random parity check matrices and showed promising distance properties. However, they went largely unnoticed until the advent of turbo codes, where they were “rediscovered” by MacKay, who showed that they perform almost as close to capacity

as turbo codes. More recently, Richardson and Urbanke have developed irregular LDPC codes that perform even better than turbo codes for very large block lengths ( $n > 10^5$ ) and can come within 0.1 dB of the Shannon capacity.

### 3.1.1 LDPC Codes

LDPC codes were originally invented by Gallager [22], However, these codes were larger ignored until the introduction of turbo codes, which rekindled some of the same ideas. LDPC codes were rediscovered by Mackay and Neal. Shortly thereafter it was recognized that these new code designs were actually reinventions of Gallager's original ideas, and much work has been devoted to finding the capacity limits, encoder and decoder designs, and practical implementation of LDPC codes for different channels.

LDPC codes are linear block codes with particular structure for the parity check matrix  $\mathbf{H}$ , which will define in next section. Since the fraction of non-zeros entries in  $\mathbf{H}$  is small, the parity-check matrix for the code has a low-density – hence the name low-density parity-check codes. Provided that the codeword length is long, LDPC codes achieve performance close to the Shannon limit and in some cases surpass the performance of parallel or serially concatenated codes [23].

The fundamental practical difference between turbo codes and LDPC codes is that turbo codes tend to have low encoding complexity (linear in block-length) but high decoding complexity (due to their iterative nature and message passing). In contrast, LDPC codes tend to have relatively high encoding complexity but low decoding complexity. The decoding algorithm of LDPC codes will present in section 3.1.3.

### 3.1.2 Construction of LDPC Codes

In this section, we will generate the LDPC codes based on the standard of IEEE 802.16-2005 [13]. The LDPC codes are based on a set of one or more fundamental LDPC codes. Each of the fundamental codes is a systematic linear block code. How to adjust the various code rates and block sizes based on fundamental codes will present in the below.

Each LDPC code in the set of LDPC codes is defined by a matrix  $\mathbf{H}$  of size  $m$ -by- $n$ , where  $n$  is the length of the code and  $m$  is the number of parity check bits in the code. The number of systematic bits is  $k = n - m$ .

The matrix  $\mathbf{H}$  is defined as: [13]

$$\mathbf{H} = \begin{bmatrix} \mathbf{P}_{0,0} & \mathbf{P}_{0,1} & \mathbf{P}_{0,2} & \cdots & \mathbf{P}_{0,n_b-2} & \mathbf{P}_{0,n_b-1} \\ \mathbf{P}_{1,0} & \mathbf{P}_{1,1} & \mathbf{P}_{1,2} & \cdots & \mathbf{P}_{1,n_b-2} & \mathbf{P}_{1,n_b-1} \\ \mathbf{P}_{2,0} & \mathbf{P}_{2,1} & \mathbf{P}_{2,2} & \cdots & \mathbf{P}_{2,n_b-2} & \mathbf{P}_{2,n_b-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{P}_{m_b-1,0} & \mathbf{P}_{m_b-1,1} & \mathbf{P}_{m_b-1,2} & \cdots & \mathbf{P}_{m_b-1,n_b-2} & \mathbf{P}_{m_b-1,n_b-1} \end{bmatrix} = \mathbf{P}^{H_b} \quad (3.1)$$

where  $\mathbf{P}_{i,j}$  is one of a set of  $z$ -by- $z$  permutation matrices or a  $z$ -by- $z$  matrix. The matrix  $\mathbf{H}$  is expanded from a binary  $\mathbf{H}_b$  of size  $m_b$ -by- $n_b$ , where  $n = z \times n_b$ , and  $m = z \times m_b$ , with  $z$  is a positive integer. The base matrix is expanded by replacing each 1 in the base matrix with a  $z$ -by- $z$  permutation matrix, and each 0 with a  $z$ -by- $z$  zero matrix. The base matrix size  $n_b$  is an integer equal to 24.

The permutations used are circular right shifts, and the set of permutation matrices contains the  $z \times z$  identity matrix and circular right shifted versions of the identity matrix. Because each permutation matrix is specified by a single circular right shift, the binary base matrix information and permutation replacement information can be combined into a single model  $\mathbf{H}_{bm}$ . The model matrix  $\mathbf{H}_{bm}$  is the same size as the binary matrix  $\mathbf{H}_b$ , with each binary entry  $(i, j)$  of the base matrix  $\mathbf{H}_b$  replaced to create the

model matrix  $\mathbf{H}_{bm}$ . Each 0 in the  $\mathbf{H}_b$  is replaced by a blank or negative (e.g., by -1) to denote a  $z \times z$  all zero matrix, and each 1 in  $\mathbf{H}_b$  is replaced by a circular shift size  $p(i, j) \geq 0$ . The model matrix can then be directly expanded to  $\mathbf{H}$ . Figure 3.1 and Figure 3.2 shows of parity check matrix and generator matrix for the (960,640) LDPC code, respectively. The points show the non-zero terms in the parity check matrix. We can see the non-zeros terms is about 0.355 percentage of the parity check matrix.

$$\begin{aligned}
 \mathbf{S}_0 &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{S}_3 = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\
\mathbf{S}_7 &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{S}_{-1} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}
 \end{aligned}$$

$\mathbf{H}_b$  is partitioned in two sections, where  $\mathbf{H}_{b1}$  correspond to the systematic bits and  $\mathbf{H}_{b2}$  corresponds to the parity-check bits, such that  $\mathbf{H}_b = \left[ (\mathbf{H}_{b1})_{m_b \times k_b} \mid (\mathbf{H}_{b2})_{m_b \times m_b} \right]$ .

$\mathbf{H}_{b1}$  is partitioned into two sections, where vector  $\mathbf{h}_b$  has odd weight, and  $\mathbf{H}'_{b2}$  has a dual-diagonal structure with matrix elements at row  $i$ , column  $j$  equal to 1 for  $i = j$ , for  $i = j + 1$ , and 0 elsewhere.

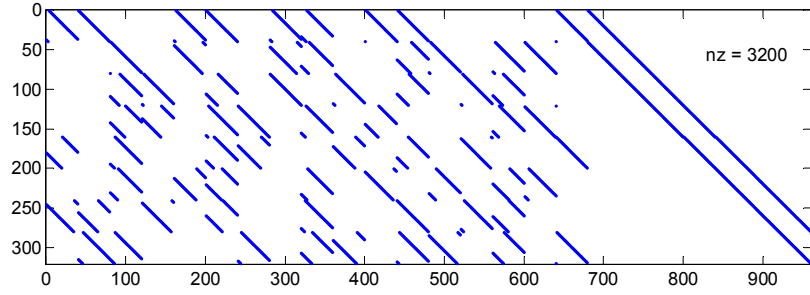


Figure 3.1: Parity-check matrix of the (960,640) LDPC code.

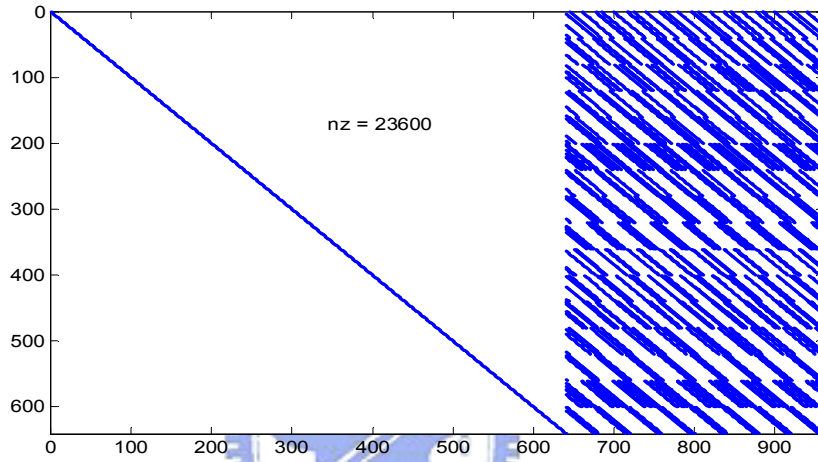


Figure 3.2: Generator matrix of the (960,640) LDPC code.

$$\begin{aligned}
 \mathbf{H}_{b_2} &= [\mathbf{h}_b \mid \mathbf{H}'_{b_2}] \\
 &= \left[ \begin{array}{c|cccc}
 h_b(0) & 1 & & & \\
 h_b(1) & 1 & 1 & & 0 \\
 \cdot & & 1 & \ddots & \\
 \cdot & & & \ddots & 1 \\
 \cdot & & 0 & 1 & 1 \\
 h_b(m_b - 1) & & & & 1
 \end{array} \right]. \tag{3.2}
 \end{aligned}$$

A based model matrix is defined for the largest code length ( $n = 2304$ ) of each code rate. The set of shifts  $\{p(i, j)\}$  in the base model matrix are used to determine the shift sizes for all other code lengths of the same code rate. Each base model matrix has  $n_b = 24$  columns, and the expansion factor  $z_f$  is equal to  $n/24$  for code

length  $n$ , Here  $f$  is the index of the code lengths for a given code rate,  $f = 0, 1, 2, \dots, 18$ . For code length  $n = 2304$  the expansion factor is designated  $z_0 = 96$ . The shift size  $\{p(i, j)\}$  for a code size corresponding to expansion factor  $z_f$  are derived from  $\{p(i, j)\}$  by scaling  $\{p(i, j)\}$  proportionally, [13]

$$p(f, i, j) = \begin{cases} p(i, j), & p(i, j) \leq 0 \\ \left\lfloor \frac{p(i, j) z_f}{z_0} \right\rfloor, & p(i, j) > 0 \end{cases} \quad (3.3)$$

where  $\lfloor x \rfloor$  denotes the flooring function that gives the nearest integer towards  $-\infty$ .

Table 3.1 shows the LDPC block sizes and code rates which are used by IEEE 802.16-2005.

Table 3.1: LDPC block sizes and code rates in IEEE 802.16-2005

$n$ (bits)	$n$ (bytes)	$z$ factor	$k$ (bytes)				Number of subchannels		
			R = 1/2	R = 2/3	R = 3/4	R = 5/6	QPSK	16QAM	64QAM
576	72	24	36	48	54	60	6	3	2
672	84	28	42	56	63	70	7	—	—
768	96	32	48	64	72	80	8	4	—
864	108	36	54	72	81	90	9	—	3
960	120	40	60	80	90	100	10	5	—
1056	132	44	66	88	99	110	11	—	—
1152	144	48	72	96	108	120	12	6	4
1248	156	52	78	104	117	130	13	—	—
1344	168	56	84	112	126	140	14	7	—
1440	180	60	90	120	135	150	15	—	5
1536	192	64	96	128	144	160	16	8	—
1632	204	68	102	136	153	170	17	—	—
1728	216	72	108	144	162	180	18	9	6
1824	228	76	114	152	171	190	19	—	—
1920	240	80	120	160	180	200	20	10	—
2016	252	84	126	168	189	210	21	—	7
2112	264	88	132	176	198	220	22	11	—
2208	276	92	138	184	207	230	23	—	—
2304	288	96	144	192	216	240	24	12	8

### 3.1.3 Generator matrix of LDPC Codes

A characteristic feature of the considered LDPC codes is that combining rows of the parity-check matrix ( $\mathbf{H}$ ) for the lowest rate code produces  $\mathbf{H}$  for higher rates. This is equivalent to replacing a group of check nodes with a single check node that sums all the edges coming into each of the original check nodes. Direct consequences are that all code rates have the same block length and the same variable node degree distribution.

The LDPC codes encoder is systematic, i.e. encodes an information block of size  $k$ ,  $\mathbf{m} = (m_0, \dots, m_{(k-1)})$  into a codeword  $\mathbf{C}$ .  $\mathbf{C} = (m_0, \dots, m_{(k-1)}, p_0, \dots, p_{(n-k-1)})$  by adding  $n - k$  parity bits obtained so that  $\mathbf{H}\mathbf{c}^T = \mathbf{0}$ .

General encoding of a  $(n, k)$  systematic LDPC code

$$\mathbf{c}_{(1 \times n)} = \left[ \mathbf{m}_{(1 \times k)} \quad \mathbf{p}_{(1 \times (n-k))} \right] = \mathbf{m} \times \mathbf{G}_{(k \times n)}. \quad (3.4)$$

If we have a parity check matrix  $\mathbf{H}$

$$\mathbf{H}_{(n-k) \times n} = \left[ \mathbf{A}_{(n-k) \times k} \quad \mathbf{B}_{(n-k) \times (n-k)} \right] \quad (3.5)$$

where  $\mathbf{B}$  is lower triangular matrix, then we can find the generator matrix  $\mathbf{G}$  [24]

$$\begin{aligned} \mathbf{H} \times \mathbf{c}^T &= [\mathbf{A} : \mathbf{B}] \times [\mathbf{m} : \mathbf{p}]^T \\ &= \mathbf{A}\mathbf{m}^T + \mathbf{B}\mathbf{p}^T = \mathbf{0} \end{aligned} \quad (3.6)$$

$$\Rightarrow \mathbf{p}^T = \mathbf{B}^{-1}\mathbf{A}\mathbf{m}^T$$

$$\Rightarrow \mathbf{G} = \left[ \mathbf{I}_k : (\mathbf{B}^{-1}\mathbf{A})^T \right] \quad (3.7)$$

Note that all the operations are operating in GF(2).

### 3.1.4 Decoding Algorithm of LDPC Codes

At the transmitter, a message vector  $\mathbf{m}$  is encoded into a code vector  $\mathbf{c} = \mathbf{m}\mathbf{G}$ , where  $\mathbf{G}$  is the generator matrix for a given parity-check matrix  $\mathbf{H}$ . The vector  $\mathbf{c}$  is



transmitted over a noisy channel to produce the received vector

$$\mathbf{r} = \mathbf{c} + \mathbf{e} \quad (3.8)$$

where  $\mathbf{e}$  is the error vector due to channel noise. By construction,  $\mathbf{H} \cdot \mathbf{c}^T = \mathbf{0}$ . Given the received vector  $\mathbf{r}$ , the bit-by-bit decoding algorithm problem is to find the most probable vector  $\hat{\mathbf{c}}$  that satisfies the condition  $\mathbf{H} \cdot \hat{\mathbf{c}}^T = \mathbf{0}$ .

The famous decoding algorithm is sum-product algorithm which is also called message passing algorithm [25]-[26]. The sum-product algorithm is based on bipartite graph-based codes. In bipartite graph, we can divide the parity-check matrix  $\mathbf{H}$  into variable nodes and constraint nodes. Figure 3.3 shows the example of the bipartite graph. Here is the example of an (7,4) Hamming code: [24]

$$\mathbf{H} = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \quad (3.9)$$

Each constraint node represents a parity check equation which is displayed in Figure 3.4.

$$v_0 + v_1 + v_2 + v_4 = 0 \quad (3.10)$$

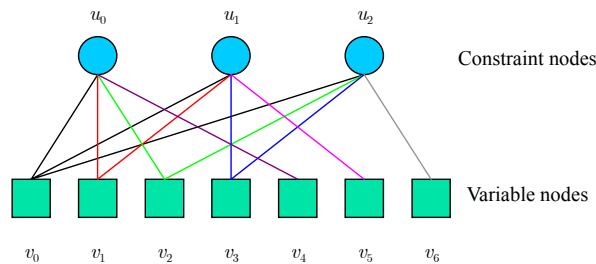


Figure 3.3: Bipartite graph of (7,4) Hamming code

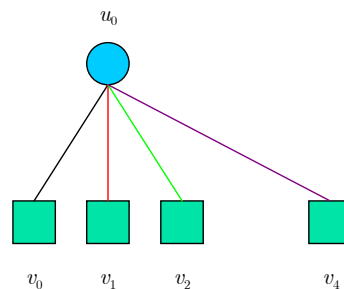


Figure 3.4: Constraint node of (7,4) Hamming code

On each iteration, each constraint node provides a probability for each variable nodes with which it shares an edge. These probabilities are then combined for the computation of the new variable nodes probability [24]-[26]

$$L_i = \frac{\Pr[c_i = 1 | y_i]}{\Pr[c_i = 0 | y_i]} \quad (\text{ML decoding})$$

$$\text{constraint} \begin{cases} \hat{c}_i = \begin{cases} 1 & \text{if } L_i \geq 0 \\ 0 & \text{if } L_i < 0 \end{cases} \\ \mathbf{H}\hat{\mathbf{c}}^T = \mathbf{0} \end{cases} \quad (3.11)$$

The sum-production algorithm proceeds as follows:

1. Initialize:

$$q_{ij}(1) = p_i = \Pr(c_i = 1 | y_i) = \frac{1}{1 + e^{2y_i/\sigma^2}} \quad (3.12)$$

$$q_{ij}(0) = 1 - p_i \quad (3.13)$$

We only computes it for  $\forall i, j$  that satisfies  $h_{ij} = 1$

2. First half round iteration:

$$r_{ji}(0) = \frac{1}{2} + \frac{1}{2} \prod_{i' \in R_j \setminus i} (1 - 2q_{i'j}(1)) \quad (3.14)$$

$$r_{ji}(1) = \frac{1}{2} - \frac{1}{2} \prod_{i' \in R_j \setminus i} (1 - 2q_{i'j}(1)) \quad (3.15)$$

3. Second half round iteration:

$$q_{ij}(0) = K_{ij} (1 - p_i) \prod_{j' \in C_i \setminus j} r_{j'i}(0) \quad (3.16)$$

$$q_{ij}(1) = K_{ij} p_i \prod_{j' \in C_i \setminus j} r_{j'i}(1) \quad (3.17)$$

where constant  $K_{ij}$  are selected to ensure  $q_{ij}(0) + q_{ij}(1) = 1$ .

4. Soft decision:

$$Q_i(0) = K_i (1 - p_i) \prod_{j \in C_i} r_{ij}(0) \quad (3.18)$$

$$Q_i(1) = K_i p_i \prod_{j \in C_i} r_{ij}(1) \quad (3.19)$$

where constant  $K_j$  are selected to ensure  $Q_i(0) + Q_i(1) = 1$ .

5. Hard Decision:

$$L_i = \log \left( \frac{Q_i(1)}{Q_i(0)} \right) \quad (3.20)$$

$$\hat{c}_i = \begin{cases} 1 & \text{if } L_i \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (3.21)$$

6. if  $\hat{\mathbf{c}}\mathbf{H}^T = 0$  or number of iterations exceeds limitation then stop, else go to Step 2.

The denotations which we used for sum-product algorithm are listed in Table 3.2. Because they are too many multiplications in sum-product algorithm, a log-domain algorithm is desirable. Multiplications may cause overflow or saturation with large numbers of iterations. To reduce the complexity, we can use log-likelihood ratio to represent the information [26]

$$L(c_i) \triangleq \log \frac{\Pr(c_i = 0 | y_i)}{\Pr(c_i = 1 | y_i)} = \log \frac{1 - p_i}{p_i} \quad (3.22)$$

Table 3.2: Notations of sum-product algorithm

Notations	Meaning
$q_{ij}$	Messages to be passed from bit node $c_i$ to check nodes $f_j$
$r_{ji}$	Messages to be passed from check node $f_j$ to bit nodes $c_i$
$R_j = \{ i : h_{ji} = 1 \}$	The set of column locations of the 1's in the $j$ -th row
$R_{j \setminus i} = \{ i' : h_{ji'} = 1 \} \setminus \{ i \}$	The set of column locations of the 1's in the $j$ -th row, excluding location $i$
$C_i = \{ j : h_{ji} = 1 \}$	The set of row locations of the 1's in the $i$ -th column
$C_{i \setminus j} = \{ j' : h_{ji'} = 1 \} \setminus \{ j \}$	The set of row locations of the 1's in the $i$ -th column, excluding location $j$

## 3.2 Review of ARQ Schemes

ARQ is an Error control method for data transmission in which the receiver detects transmission errors in a message and automatically requests a retransmission from the transmitter [27]-[29]. Usually, when the transmitter receives the ARQ, the transmitter retransmits the message until it is either correctly received or the error persists beyond a predetermined number of retransmissions.

### 3.2.1 Conventional ARQ

In most digital communication systems, whenever error events occur in the transmitted messages, some action must be to correct these events. This action may take the form of an error correction procedure. In some applications where a two-way communication link exists between sender and the receiver, the receiver may inform the sender that a message has been received in error and, hence, request a repeat of that message. In principle, the procedure may be repeated as many times as necessary until that message is received error free. An error control system in which the erroneously received messages are simply retransmitted is called ARQ [27]. The procedure of ARQ is listed in Figure 3.5.

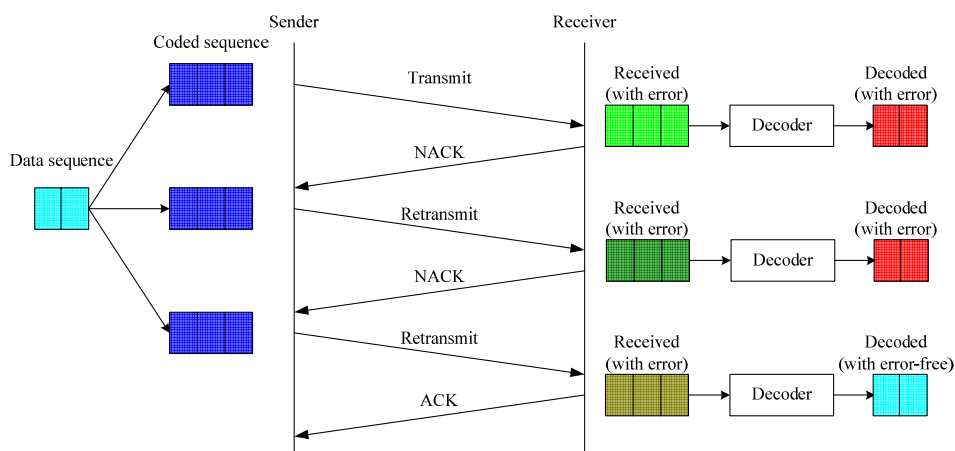


Figure 3.5: Block diagram of ARQ and HARQ type I

In ARQ systems, the receiver must perform only an error detection procedure on the received messages without attempting to correct the errors. Hence, an error detecting code, in the form of specific redundant or parity-check symbols, must be added to the information-bearing sequence. In general, as the error detecting capability of the code increases, the number of added redundant symbols must also be increased. Clearly, with such a system, an erroneously received message is delivered to the user only if the receiver fails to detect the presence of errors. Since error detection coding is simple, powerful, and quite robust, ARQ systems constitute a simple and efficient method for providing highly reliable transfer of messages used in wireless communication.

ARQ refers to retransmission techniques, which basically operate as follows:

- Erroneously received packets are retransmitted until they are received or detected as being error-free.
- Errors are detected using a simple detection code.
- Positive (ACK) or negative (NAK) acknowledgements are sent back by the receiver to the sender over a reliable feedback channel in order to report whether a previously transmitted packet has been received error-free or with errors.

### 3.2.2 Hybrid ARQ Type I

Hybrid ARQ schemes combines the conventional ARQ with FEC and they can be regarded as an implicit link adaptation technique [28]. A hybrid ARQ system consists of an FEC subsystem contained in an ARQ system. The function of the FEC system is to reduce the frequency of retransmission by correcting the error patterns that occur most frequency, thus ensuring a high system throughput. When a less frequent error pattern occurs and is detected, the receiver requests a retransmission instead of passing

the erroneous data to the user. This increases the system reliability. HARQ schemes offer the potential of better performance and lower implementation cost if appropriate ARQ and FEC schemes are properly combined.

In HARQ Type I, the same PDU is retransmitted until the receiver accepts it as error-free or until the maximum number of allowed retransmission attempts is reached. HARQ Type I scheme is best suited for communications systems in which a fairly constant level of noise and interference is anticipated in the channel. In the case, an adequate amount of redundancy can be built into the system to correct the vast majority of errors so the number of retransmissions can be kept to a minimum. For non-stationary channels, such as the one encountered in wireless communications, the HARQ Type I can be very inefficient. When the channel condition is good, the large number of redundancy bits built into the FEC subsystem results in a waste of bandwidth. On the other hand, when channel condition is poor, the FEC may be not powerful enough and too many retransmissions may be needed.

### **3.2.3 HARQ Type II**

HARQ Type II is particularly suited for time-varying channels. In this scheme, the concept of incremental redundancy is employed and the received PDUs are concatenated to form corrupted codewords from increasingly longer and lower rate codes [29]-[30]. In the first transmission, the PDU may be coded with a high-rate code (low redundancy) for error detection and correction. If the receiver detects the presence of errors in the PDU, it saves the erroneous PDU in a buffer and at the same time requests a retransmission. Unlike HARQ Type I, what is retransmitted in HARQ Type II is not the original PDU but a block of new data. The new data are formed based on the original PDU and the error correcting code used. When the new PDU is received, it

is used to correct the errors in the erroneous PDU stored previously in the buffer. If the second attempt fails again, the receiver will request a further retransmission and this process continues until satisfactory results are achieved. The procedure of HARQ Type II is listed in Figure 3.6.

### 3.2.4 HARQ Type III

HARQ Type III also belongs to the class of incremental redundancy ARQ schemes [28]. With HARQ Type III, however, each retransmission is self-decodable. Chase combining involves the retransmission by the transmitter of the same coded data packet. The decoder at the receiver combines these multiple copies of the transmitted packet weighted by the SNR of the received signal for each attempt. As a result, diversity gain in the time domain is obtained and the temporal diversity may also include contributions of spatial diversity implicitly.

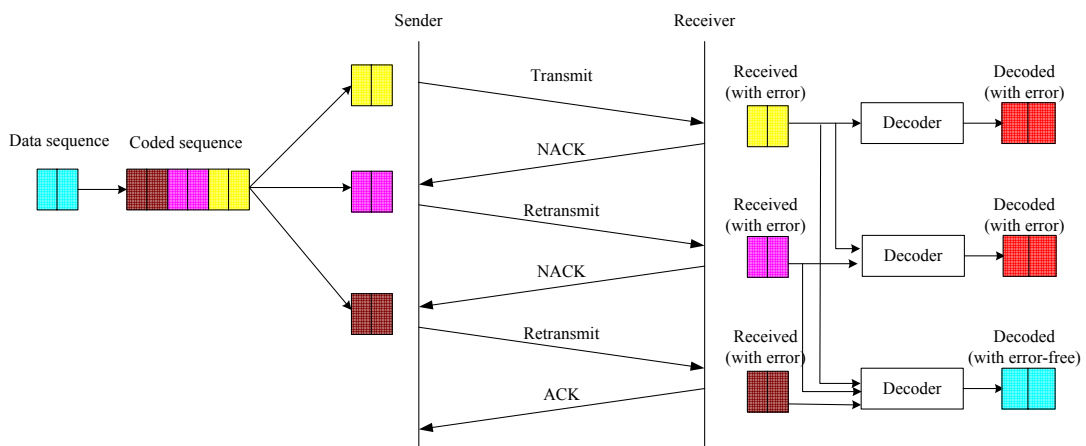


Figure 3.6: Block diagram of HARQ type II

### 3.3 Reliability-Based HARQ

The RB-HARQ scheme, which can be used with error correcting codes using SISO decoders such as convolutional codes and turbo codes has been proposed [15]-[20]. In the RB-HARQ scheme, the error rate performance is improved by selecting the retransmission bits based on LLR of each bit in the receiver. However, the receiver has to send the bit positions of retransmission bits to the transmitter. Figure 3.7 shows the block diagram of RB-HARQ. On the other hand, LDPC codes are attracting a lot of interest, recently. Hence, LDPC codes can achieve near Shannon limit performance and be decoded easily compared to turbo code. In this section, we evaluate the RB-HARQ scheme using LDPC code. Moreover, we propose an adaptive RB-HARQ scheme that adjusts the retransmission sizes according the QoS request.

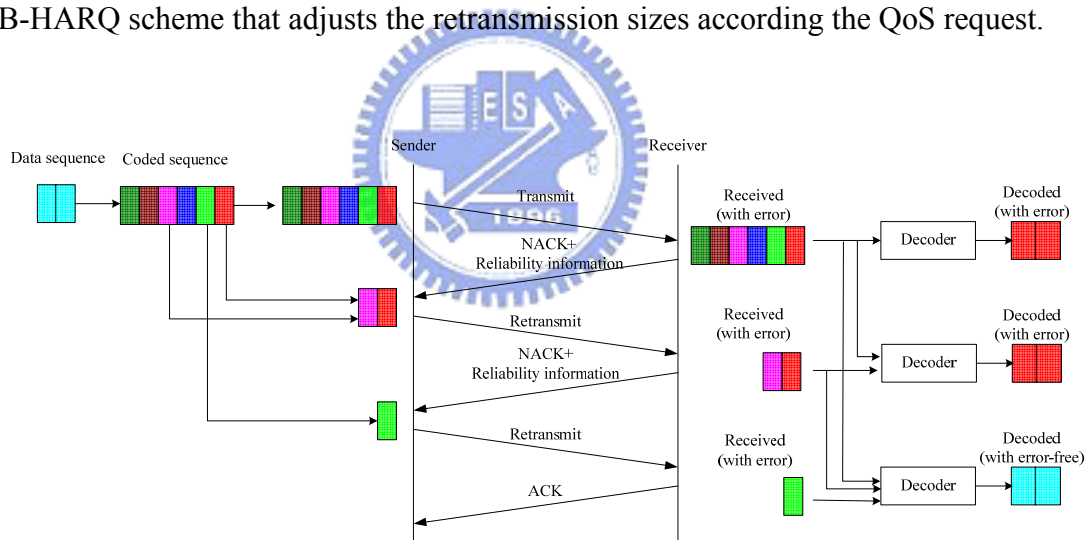


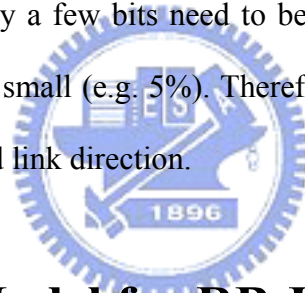
Figure 3.7: Block diagram of RB-HARQ.



### 3.3.1 Review of Reliability-Based HARQ

In [15], Shea proposed a new HARQ scheme that utilizes reliability estimates generated by SISO decoders. The proposed HARQ scheme transmits additional information for the unreliable bits, and this information is used to perform additional decoding. The bits that are to be retransmitted are adaptively selected at the receiver based on the estimated bit reliabilities at the output of the SISO decoder so the scheme is called RB-HARQ.

Unlike other ARQ schemes, only the weak bits would be retransmitted. When decoding attempt fails, the receiver selects the most unreliable bits and sends their position information to the source. According to [20], it could bring significant performance gain while only a few bits need to be retransmitted in forward link, and the increment may be quite small (e.g. 5%). Therefore, RB-HARQ is considered to be quite efficient in the forward link direction.



### 3.3.2 System Model for RB-HARQ scheme

Considering a single-transmit single-receive antenna system in Figure 3.8. The source radio  $S$  and the destination radio  $D$  are linked by a data channel through which a packet of information is to be delivered from  $S$  to  $D$ . The data bits in  $S$  are encoded by LDPC codes and the resulting code bits are modulated using  $M$ -QAM. The encoded packet is then transmitted over a fading channel. The destination  $D$  attempts to decode the packet and sends a retransmission request through the feedback channel if an error is detected. The retransmission-request packet, sent from  $D$  to  $S$ , contains a list of the least-reliable code bits based on the magnitudes of the soft output of LDPC decoder. The source  $S$  then retransmits the code bits. Noisy versions of the retransmitted code bits are received at  $D$  and they are added to the previously received

values of the same code bits. In our study, we assume perfect error detection and the presence of a highly reliable feedback channel from  $D$  to  $S$  [20].

In this system,  $N$  coded bits are obtained by encoding  $K$  information bits with a LDPC code of rate  $R = K / N$ . Let information bits  $b_1$  through  $b_K$  be encoded into a codeword  $c$  with bits  $c_1$  through  $c_N$ . The coded bits are mapped to transmission symbol  $s_1$  through  $s_{N/M}$ .  $M$  represents to the  $M$ -QAM. In this study, we only use  $\{4,16,64\}$ -QAM. Every  $\log_2(M)$  encoded bits can be viewed as a  $\log_2(M)$ -dimensional vector  $u_k = \{c_{(k-1) \times \log_2(M)+1}, c_{(k-1) \times \log_2(M)+2}, \dots, c_{k \times \log_2(M)}\}$ ,  $k = 1, 2, \dots, \lfloor N / \log_2(M) \rfloor$ . Thus, coded symbol will be mapped to one of the  $M$  points on the  $M$ -QAM constellation, which is represented by a couple of real-value symbols  $(A_p, B_p)$ .

Considering a coherent receiver, the received signal can be written as  $(X_k, Y_k)$ ,

$$X_k + jY_k = a_k (A_k + jB_k) + n_k \quad (3.23)$$

where  $a_k$  is the complex channel fading gain.  $n_k$  is complex Gaussian noise with zeros mean, variance  $\sigma_n^2$ . After the soft-demodulator, we define the bit level LLR as the soft value corresponding to the  $c_i$  coded bit is given by  $y_i$ .

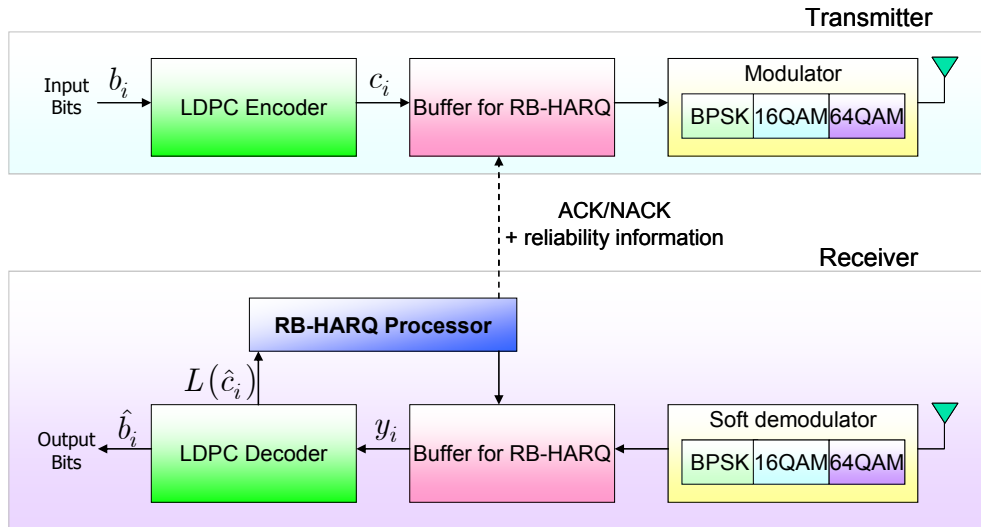


Figure 3.8: System model of RB-HARQ scheme

Using Bayes' rule, and supposing  $P\{c_i = 1\} = P\{c_i = 0\}$ , we can obtain:

$$y_i \triangleq \log \frac{P\{c_i = 1 | X_i, Y_i\}}{P\{c_i = 0 | X_i, Y_i\}} = \log \frac{P\{X_i, Y_i | c_i = 1\}}{P\{X_i, Y_i | c_i = 0\}} \quad (3.24)$$

As we know,  $c_i = 1$  and  $c_i = 0$  are corresponding to  $M/2$  different constellation points respectively. That is to say, for every bit  $c_i$ , the M-QAM constellation is split into two partitions. Suppose that  $C_1(i)$  comprises the  $M/2$  symbol points  $(X_n, Y_n)$  with  $c_i = 1$ , and  $C_0(i)$  comprises the  $M/2$  symbol points  $(X_n, Y_n)$  with  $c_i = 0$ ,  $n = 1, 2, \dots, M/2$ . Thus we can rewrite Equation (3.24) as: [17]

$$y_i \triangleq \log \frac{\sum_{(X_n, Y_n) \in C_1(i)} P\{X_k + jY_k = a_k(X_n + jY_n) + n_k\}}{\sum_{(X_n, Y_n) \in C_0(i)} P\{X_k + jY_k = a_k(X_n + jY_n) + n_k\}} \quad (3.25)$$

We use the  $y_i$  as the input message of the LDPC decoder. The LDPC decoder uses the sum-production algorithm which we have illustrated in section 3.1.3. After decoding, the *a posteriori* LLR of code bit can define as

$$L_i = \log \frac{Q_i(1)}{Q_i(0)}. \quad (3.26)$$

We use the syndrome  $\hat{\mathbf{c}}\mathbf{H}^T$  to replace the error detection. When  $\hat{\mathbf{c}}\mathbf{H}^T = 0$ , we believe there is no error in the error detection. Otherwise, there are some errors in the error detection. If the received packet is in error after decoding, the receiver decides the retransmission bits based on the LLRs that are outputs of the SISO decoder. For each error packet, the bits are ranked by the magnitude of decoder outputs  $L_i$ . The bit with the smallest  $|L_i|$  is considered to be least reliable, denoted by “reliability (1),” and the bit with the largest  $|L_i|$  is considered to be most reliable, denoted by “reliability (N).” If the decoding fails, the receiver feeds back the indices of the bits with the reliability (1)~(T) to the transmitter where  $T$  denotes the number of retransmission bits. The transmitter then retransmits the code bits corresponding to those requested information

bits. Retransmitted code bits corrupted by noise are received. When a new copy of the same coded bits (either information or parities) is received, old copies are not discarded, but are combined with the new ones to facilitate decoding. In general, packet combining is done by averaging the soft decision values from the multiple copies. In our study, we assume the presence of a highly reliable

### 3.4 Proposed Adaptive RB-HARQ Algorithm

In [16], the simulations show that the performance of RB-HARQ is related to two factors: number of retransmissions and retransmission sizes. Increasing the number of retransmissions or retransmission size will decrease the probability of packet error but also decrease the system throughput. Decreasing the number of retransmissions or retransmission size will increase the probability of packet error but also decrease the system data rate. How to efficiently select the number of retransmissions and retransmission size is a tradeoff problem which dominates the system performance.

Unfortunately, the previous works don't have any efforts dealing with the tradeoff problem. So, we will propose the adaptive RB-HARQ algorithm which can adaptively adjust the number of retransmissions and retransmission sizes to satisfy system QoS. The proposed algorithm is based on the reliability of the posteriori LLR.

In the system, the decoding is performed jointly by combining soft inputs from the previous retransmission with the soft inputs derived from the current retransmission. It is assumed that a maximum of  $R^{\max}$  retransmissions of a packet can be jointly decoded. Let  $j$  denote the transmission index,  $\Pi_j$  denote the set of transmitted bits in the  $j$ th transmission,  $T_j$  denote the size of  $j$ th transmission and  $\mathbf{y}_j$  denote the corresponding soft input vector. For joint decoding, the soft input vectors are combined across

transmissions to obtain  $\tilde{\mathbf{y}}_j = \sum_{k=1}^j \mathbf{y}_k$ . The effective code rate  $R_j$  attained by combining the transmission is given by [30]

$$R_j = \frac{K}{\sum_{k=1}^j T_k} \quad (3.27)$$

$T_1 = N$  is the size of the first transmission. First, we define the reliability of a codeword,  $\mu$ , as the expected value of the magnitude of the *a posteriori* LLRs of the information bits, averaged over realizations of noise, sets of transmitted bits and channel taps at a fixed SNR. Following the  $j$ th transmission, codeword reliability  $\mu_j$  is then given by

$$\mu_j = E_{\Pi, n, \mathbf{h}} [L_i^j] \quad (3.28)$$

where the expectation is taken over noise realizations, the cumulative set of transmitted bits for the  $j$ th transmission and channel taps. In AWGN and static channel environments, the expectation with respect to channel taps is unnecessary. For a given sequence of information bit LLRs following the  $j$ th transmission, an estimate of the codeword reliability  $\hat{\mu}_j$  can be obtained at the receiver by computing an average packet reliability

$$\hat{\mu}_j = \frac{1}{K} \sum_{i=1}^K |L_i^j| \quad (3.29)$$

This computation can be performed at the receiver in static as well as fading channels.

We denote by  $PER = P(\mu)$  the reliability to PER mapping and by  $\mu = G(T)$  the retransmission size to reliability mapping. Based on the numerical results, both mappings are one-to-one, so that the inverse mapping  $\mu = P^{-1}(PER)$  and  $T = G^{-1}(\mu)$  are well defined [31]. In our simulation, we can find that the mapping are almost the same in different SNR, because the channel tap dominates the system performance.

Next, we will demonstrate how the reliability mappings  $G(\cdot)$  and  $P(\cdot)$  can be used to optimize the size of retransmission. The size of the first packet transmission is determined by the MCS selection algorithm which we will illustrate in chapter 4. In this

section, we focus on the adaptive RB-HARQ algorithm, which adaptively optimizes sizes  $T_2$  through  $T_{R^{\max}}$ .

Figure 3.10 shows an example of the reliability to PER mapping  $P(\cdot)$ . Suppose that after the first transmission, the packet is received with reliability  $\hat{\mu}_1$ , corresponding to a probability of decoding failure  $PER_1 = P(\hat{\mu}_1)$ . Transmitting additional coded bits and combining them with the initial transmission will raise the reliability to  $\hat{\mu}_2$ , corresponding to a probability of decoding failure  $PER_2 = P(\hat{\mu}_2)$ . Additional transmissions will continue to increase the reliability and move the probability of a failure down the curve, as shown in Figure 3.10.

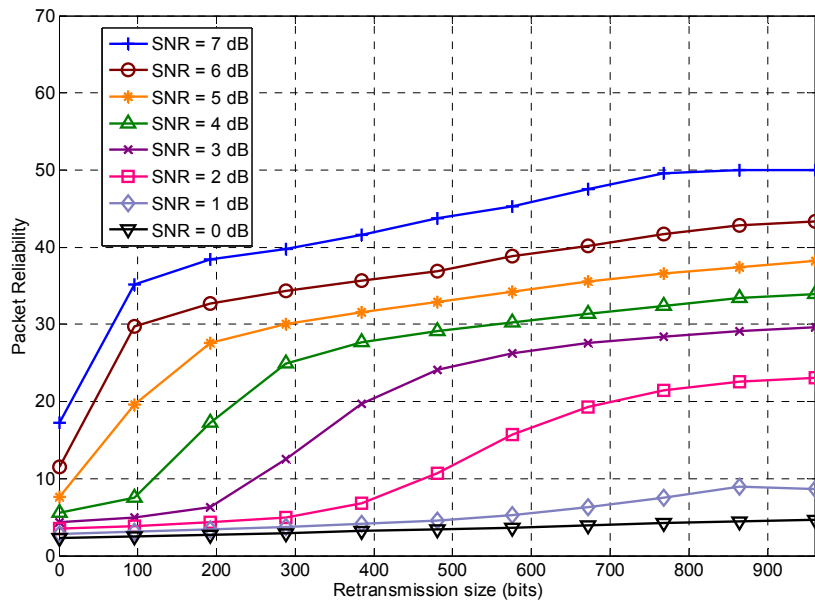


Figure 3.9: Packet reliability versus retransmission size (bits) in RB-HARQ scheme

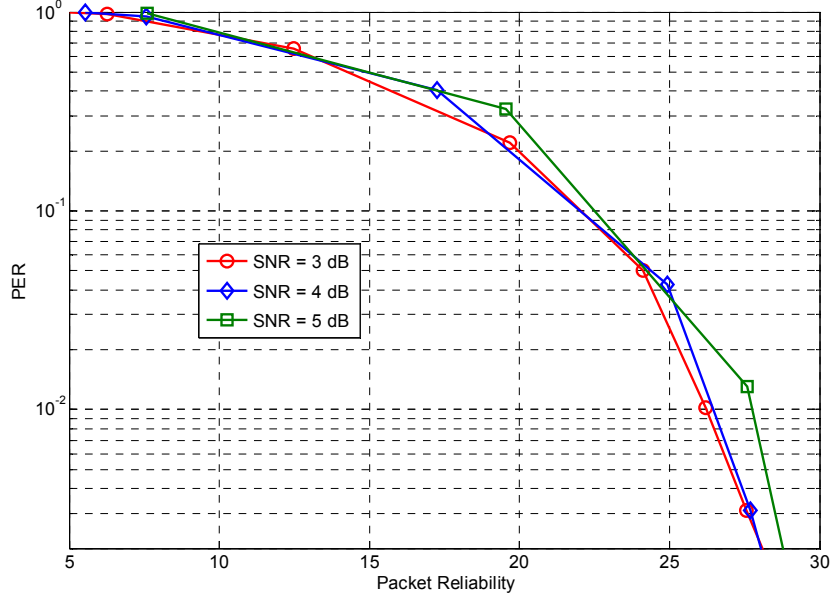


Figure 3.10: PER versus packet reliability in RB-HARQ scheme

The goal of the reliability-based approach is to determine reliability levels for the transmissions, which minimize the use of the channel response. First, to simplify the adaptive problem, we only consider that the number of retransmission  $R^{\max}$  equals 1. Suppose that a packet is in error after the first transmission and has an estimated packet reliability level  $\hat{\mu}_1$ . The size of the second transmission is chosen to raise the packet reliability to  $\bar{\mu}_2$ , corresponding to a target probability of decoding failure  $PER_{\text{target}} = F(\bar{\mu}_2)$ . Specifically, by inverting the retransmission size to packet reliability mapping  $G^{-1}(\cdot)$ , the size is computed as

$$T_2 = \max\left(G^{-1}(\bar{\mu}_2) - G^{-1}(\hat{\mu}_1), 0\right) \quad (3.30)$$

When condition  $G^{-1}(\hat{\mu}_1) \geq G^{-1}(\bar{\mu}_2)$  holds, the instantaneously PER is lower than the target probability of decoding failure, but we will still drop the packet and transmit the next packet. In a practical system,  $T_2$  would be further quantized to the nearest size from the set of permissible transmission sizes. The index of the quantized size with the unreliability information would then be signaled back to the transmitter.

The adaptive algorithm can be extended when the number of retransmission is more than two. The steps of the algorithm:

1. Initialization,  $j = 1$
2. If decoding errors occur,  $\hat{\mu}_j < \bar{\mu}_{\text{target}}$  and  $j < J$

$$T_{j+1} = G^{-1}(\bar{\mu}_{\text{target}}) - G^{-1}(\hat{\mu}_j) \quad (3.31)$$

Retransmission size =  $T_{j+1}$ .

Go back to step 2,

3. Transmit the next packet.

Note that a packet is dropped if it received incorrectly after  $R^{\text{max}}$  transmissions and  $\hat{\mu}_j \geq \bar{\mu}_{\text{target}}$ . When  $\hat{\mu}_j \geq \bar{\mu}_{\text{target}}$ , We suppose that the instantaneous PER is guaranteed to be not greater than  $P_{\text{loss}}$ .

Based on the adaptive algorithm, we can adaptively adjust the retransmission sizes which can minimize the number of retransmission and satisfy the QoS requirements. The performance of the adaptive algorithm will present in the next section.

## 3.5 Computer Simulations

In this section, computer simulations are conducted to evaluate the performance of the proposed adaptive RB-HARQ algorithm using LDPC codes. In our simulation, we use a (960, 640) LDPC code with code rate 2/3. The QPSK modulation is used, and a Rayleigh fading channel and an ideal feedback channel are assumed. In the sum-product algorithm, the maximum number of iterations is set to 20. Table 3.4 lists all parameters used in our simulation.



Table 3.4: Simulation parameters of the adaptive RB-HARQ system

Number of TX/RX antennas	1/1
$2^M$ -QAM available	2, 4, and 6
Channel coding	LDPC (960,640)
Packet size	640 data bits
Maximum number of transmissions per packet	$R^{max} = 1$
Channel model	Rayleigh fading
LDPC decoding algorithm	sum-product algorithm
	20 iterations
Retransmission sizes available	0, 96, 192, 288, 384, 480 576, 672, 768, 864, 960
ARQ Types	ARQ, Fixed RB-HARQ, Adaptive RB-HARQ
$P_{loss}$	$10^{-2}$

Each retransmission effectively reduces the code rate and hence increases the  $E_b/N_0$  at the receiver. We account this additional received energy by defining the *effective*  $E_b/N_0$  as the average  $E_b/N_0$  at the receiver, taking into account the average number of incremental redundancy transmissions.

Figure 3.11 and Figure 3.12 show the PER and throughput (bits per symbol) performances of a system using a (960,640) LDPC code with QPSK, 16-QAM, and 64-QAM in a Rayleigh fading channel, respectively. As shown in Figure 3.11, the PER curves are similar to the waterfall curve in an AWGN channel. Based on the property, LDPC code shows the inherent feature which eliminates the need of the channel interleaver. From the results, we can observe that the LDPC code provides excellent PER performance. This is the reason why we use LDPC code in the HARQ schemes.

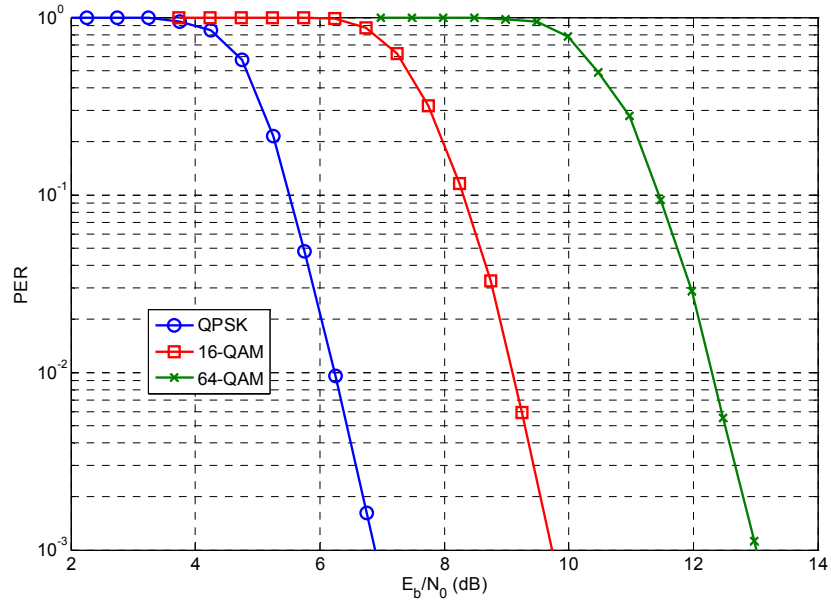


Figure 3.11: PER versus average  $E_b/N_0$  for (960,640) LDPC coded system in a Rayleigh channel with QPSK, 16-QAM, and 64-QAM

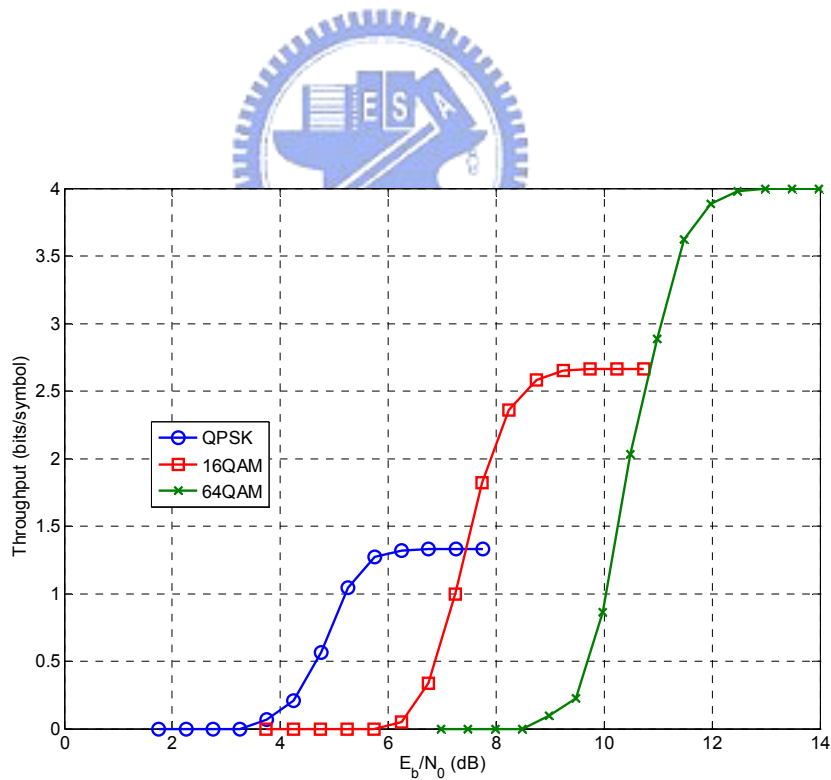


Figure 3.12: Throughput (bits/symbol) versus average  $E_b/N_0$  for (960,640) LDPC coded system in a Rayleigh fading channel with QPSK, 16-QAM, and 64-QAM

Figure 3.13 shows the comparison of the PER performances among the RB-HARQ scheme, the IR-HARQ scheme, and the no-ARQ scheme. We observe that the RB-HARQ scheme ( $T_2 = 384$  bits) requires 0.7 dB *effective*  $E_b / N_0$  lower than the IR-HARQ scheme, and 2.2 dB *effective*  $E_b / N_0$  lower than the no-ARQ scheme to achieve a packet error probability of  $10^{-2}$ . In the RB-HARQ scheme, by retransmitting the most unreliable bits, the PER performance can be improved more quickly than that in the IR-HARQ scheme, which retransmits the redundant code bits. We can observe that when the number of retransmission bits increases, the probability of packet error decreases. However, when the number of retransmission bits becomes greater, the probability of packet error saturates.

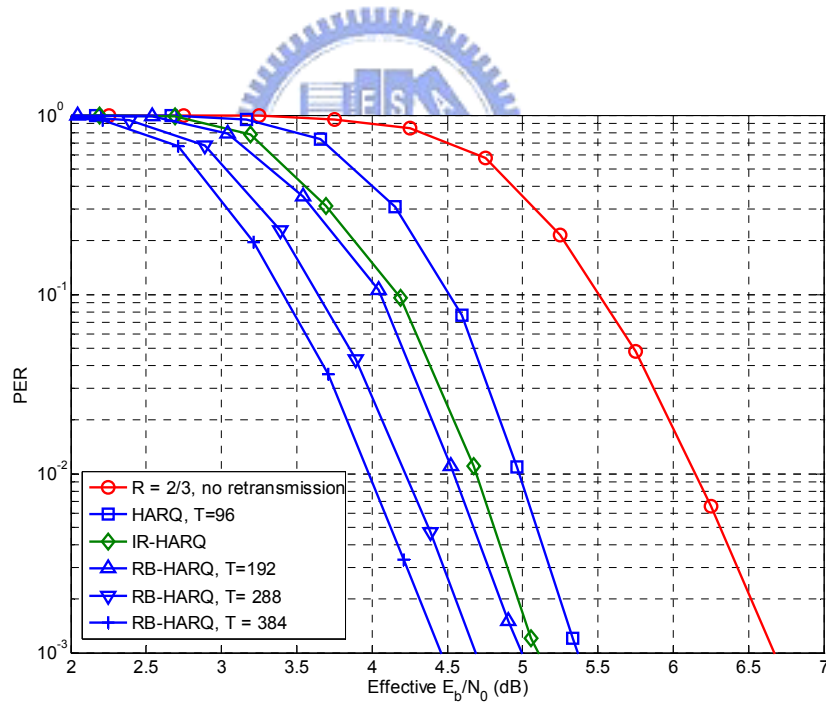


Figure 3.13: PER versus average *effective*  $E_b / N_0$  for (960,640) LDPC coded system in a Rayleigh fading channel with QPSK by the RB-HARQ scheme, no-ARQ scheme, and IR-HARQ scheme

Figure 3.14 presents the comparison of the PER performances among the conventional ARQ scheme, the fixed RB-HARQ scheme ( $T_2 = 288$  bits), and the proposed adaptive RB-HARQ scheme targeted on  $PER \leq 10^{-2}$ . As shown in the figure, the average PER of the proposed RB-HARQ scheme is lower than the required  $P_{loss}$  for  $SNR > 3.0$  dB. However, the fixed RB-HARQ scheme satisfies the  $P_{loss}$  requirement for  $SNR > 4.4$  dB, and the conventional ARQ scheme satisfies the  $P_{loss}$  requirement for  $SNR > 6.8$  dB.

Figure 3.15 compares the throughput performances of the adaptive RB-HARQ scheme, the fixed RB-HARQ scheme and the conventional ARQ scheme. As seen in the figure, the proposed RB-HARQ scheme provides significantly better performance than others, since it can adaptively choose the number of retransmission bits based on the packet reliability. In the low SNR region, the proposed RB-HARQ scheme chooses the larger number of retransmission bits, and in the high SNR region, it chooses the smaller one. Compared with the fixed RB-HARQ scheme, it leads to an increase of about 0.4 bits/symbol in the throughput at the  $SNR = 3$  dB. For example, in IEEE 802.16 standard with 7 MHz bandwidth where the symbol rate is 5.93 Msymbols/s, the proposed scheme leads to an increase of about 2.37 Mbps in the throughput over the fixed RB-HARQ scheme.

In the above simulations, the proposed adaptive RB-HARQ scheme can improve the performance in wireless environments by simply adjusting the number of retransmission bits, and also satisfy the PER requirement. So, it strongly motivates us to incorporate the concept of the adaptive RB-HARQ scheme in MIMO-OFDM systems, which is described in Chapter 4.

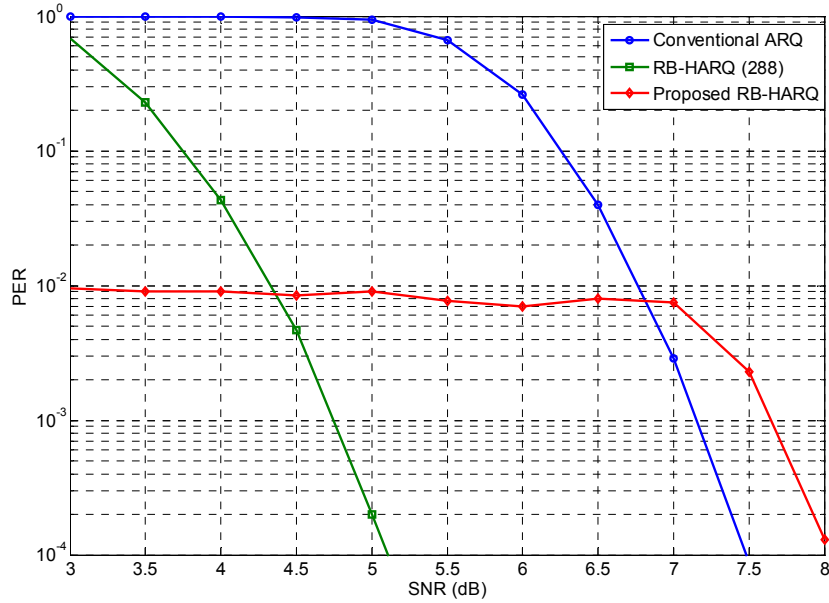


Figure 3.14: PER versus average SNR for (960,640) LDPC coded system in a Rayleigh fading channel with QPSK by the HARQ scheme, RB-ARQ scheme, and proposed RB-HARQ scheme (require PER = 10<sup>-2</sup>)

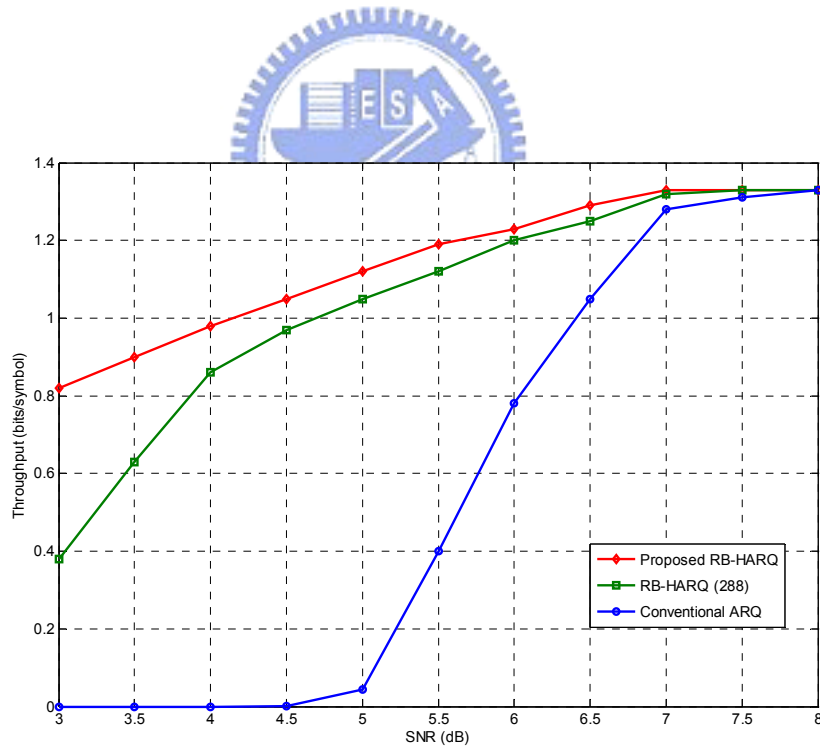
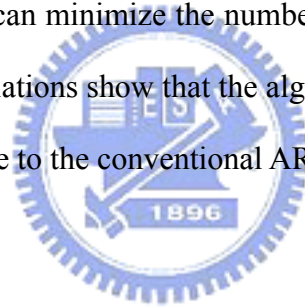


Figure 3.15: Throughput (bits/symbol) versus average SNR for (960,640) LDPC coded system in Rayleigh fading channel with QPSK by HARQ scheme, RB-ARQ scheme, and proposed RB-HARQ scheme (required PER = 10<sup>-2</sup>)

## 3.6 Summary

In this Chapter, we evaluated the performance of the RB-HARQ scheme using LDPC codes. Moreover, we proposed an adaptive RB-HARQ algorithm for communication systems employing LDPC coding. The algorithm is based on the notion of codeword reliability, and its relationship to retransmission sizes and PER. The algorithm adaptively computes the size of the ARQ transmission with the goal of minimizing the retransmission size under the prescribed QoS requirement. Conventional RB-HARQ scheme specifies the retransmission bits by considering each unreliable bit. How to determine the retransmission size will deeply affect the performance. Based on the adaptive algorithm, we can adaptively adjust the retransmission sizes which can minimize the number of retransmissions and satisfy the QoS requirement. The simulations show that the algorithm has the potential of attaining high throughput gain relative to the conventional ARQ scheme.



## Chapter 4

# Combining RB-HARQ and Adaptive Modulation in IEEE 802.16-like MIMO-OFDM Systems

The evolution of the communications space will rely on the design of next generation wireless communication systems based on a whole new concept of fast, reconfigurable networks, supporting features such as high data rates, QoS, adaptability to varying channel conditions, and integration of a number of wireless access technologies, depending on the exploitation of new resources such as cross-layer and contextual information. The system performance of future wireless networks will be enhanced by cross-layer joint design between PHY layer and MAC layer.

In this chapter, instead of considering AMC at the PHY layer and ARQ at the MAC layer separately, we pursue a cross-layer design that combines these two layers judiciously to maximize spectral efficiency, or throughput, under prescribed delay and error performance constraints. We also evaluate different error control and adaptation mechanisms available in the different layers for transmission, namely MAC retransmission strategy, modulation order and channel coding.

## 4.1 The Concept of Cross Layer Design

The system performance of future wireless networks will be enhanced by cross-layer joint design between multiple protocol layers [32]. To this end, current research efforts focus on identifying the most promising approaches, and the requirements and challenges associated with their incorporation in future wireless systems design. In order to allow wireless communication transceivers to operate in a multi-parametric continuously changing environment, reconfigurable adaptive techniques to adjust the transmit parameters of transceivers and achieve the best performance need to be devised. Reconfigurability in smart antenna (a special form of MIMO) transceivers can be viewed as the capability of intelligent switching between transceiver architectures with varying performance in a certain parameter of interest. One example could be the design of an algorithm that exploits the fundamental trade-off between spatial diversity and multiplexing in MIMO channels. Novel approaches have been proposed recently that achieve reconfigurability by introducing parameterization in the transceiver design with respect to the parameters against which reconfiguration is to be performed, such as antenna correlation and CSI reliability [32][33].

The system performance can be enhanced by interacting with the higher layers of the open systems interconnection model of the International Standards Organization (OSI/ISO) protocol stack. Smart antenna techniques can be developed combining parameters in the PHY, link (MAC, data link control (DLC), scheduling, etc.), and network layers (radio resource management, routing, transport, etc.); that is, in a cross-layer fashion rather than attempting to optimize the designs in isolation from one another. A layer-isolated approach often proves inefficient when the performance evaluation takes into account higher layers. Furthermore, layer-isolated approaches can



prevent the PHY layer innovations from being incorporated in the standardization effort and eventually adopted in the implementation phase. The information to be exchanged among the functionalities residing in different OSI layers can be classified as follows [33]:

- **CSI**, that is, estimates for channel impulse response, location information, vehicle speed, signal strength, interference level, and so on.
- **QoS-related parameters** including delay, throughput, bit error rate (BER), packet error rate (PER) measurements, and so on.
- **PHY layer resources** including spatial processing schemes, number of antenna elements, battery depletion level, and so on.

It is also important to carefully consider the cross-layer optimization criteria. In practical systems, the link quality when employing smart antenna techniques is determined not only by the performance of data detection methods, but also by the specific coding scheme being used, the MAC/DLC functionalities adopted in the link layer, or even the performance of protocols in the higher layers of the protocol stack. Thus, all these blocks should optimally be designed to achieve the highest possible overall system throughput including all links instead of the highest data rate for a single link [33].

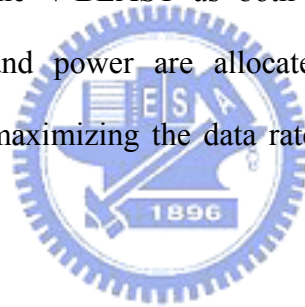
## **4.2 Adaptive Modulation Assisted MIMO-OFDM System**

The combined application of MIMO and OFDM (MIMO-OFDM) yields a noticeable PHY layer capable of meeting the requirements for 4G broadband wireless systems [34]. Thanks to OFDM, which is characterized by possessing multi-channels over frequencies, the signal processing techniques involved in MIMO-OFDM could be

borrowed from the sophisticated space-time ones by admitting the virtual equivalence between time and frequency in some particular scenarios.

From the analysis of MIMO channel capacity, we figure out that the waterfilling distribution of power over channels with different SNR values achieves the optimal transmission scheme [35]. However, while the waterfilling distribution will indeed yield the optimal solution, it is difficult to compute, and also assumes infinite granularity in the constellation size, which is not practically realizable; the waterfilling solution is optimal in capacity sense, not guaranteed the throughput being optimal, and it may cause the error probability enhancement.

In this section, we introduce a practical adaptive loading procedure for MIMO-OFDM that uses the V-BLAST as both its channel quality indicator and detection algorithm. Bit and power are allocated in a manner to fix the total transmission power while maximizing the data rate and yet still maintaining a target system performance.



## 4.2.1 Adaptive Modulation

Wireless communication channels typically exhibit time-variant quality fluctuations and hence conventional fixed-mode modems usually suffer from bursts of error (see Figure 4.1 to get the idea). This defect could be somewhat mitigated if the system was designed to provide a high link margin, that is, determine transmission parameters based on the worst-case channel conditions to render the immunity against channel impairments. However, it results in insufficient utilization of the full channel capacity. An efficient approach of avoiding these detrimental effects is to dynamically adjust transmission parameters based on the near instantaneous channel quality information. In general, the following steps have to be taken to react to the change in

channel condition for an adaptive wireless transceiver.

1. Channel quality estimation: If the communication between the two stations is bi-directional and the channel can be considered reciprocal, then each station can estimate the channel quality on the basis of received symbols, and adapt the parameters of the local transmitter to this estimation in an open-loop manner.
2. Choice of the appropriate parameters for the next transmission: The transmitter has to select the appropriate modulation mode for each sub-channel based on the prediction of the channel quality for the next timeslot.
3. Signaling of the employed parameters: The receiver has to be informed, as to which demodulator parameters to employ for the received packet. This information can be either conveyed by the transmitted signal itself, at the cost of a loss of effective data throughput, or estimated by a blind detection mechanism at the receiver.

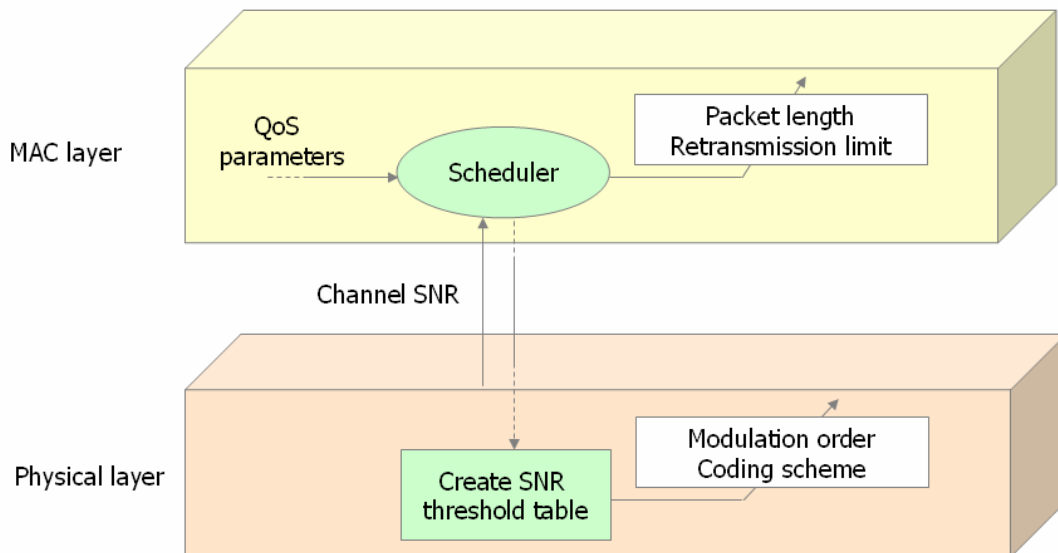


Figure 4.1: System architecture of the proposed V-BLAST based adaptive MIMO-OFDM system

## 4.3 System Model for V-BLAST Based Adaptive MIMO-OFDM system

Figures 4.2 and 4.3 show the system architecture of the proposed closed-loop V-BLAST based adaptive MIMO-OFDM system. It consists of a joint adaptive modulation and coding module at the PHY layer, and link adaptation module at the MAC layer.

At the PHY layer, we have  $N_c \times N_t$  sub-channels, where  $N_c$  is the number of subcarriers and  $N_t$  is the number of spatial channels per subcarrier. For a given time slot  $n$ ,  $N_c \times N_t$  bit streams,  $\{\mathbf{b}_i[n, k]: k = 0, 1, \dots, N_c\}$  for  $i = 1, 2, \dots, N_t$ , are encoded separately in different puncturing rates into  $N_c \times N_t$  coded bit streams,  $\{\mathbf{c}_i[n, k]: k = 0, 1, \dots, N_c\}$  for  $i = 1, 2, \dots, N_t$ , respectively, and then modulated into QAM symbols,  $\{t_i[n, k]: k = 0, 1, \dots, N_c\}$  for  $i = 1, 2, \dots, N_t$ , where the length of  $\mathbf{b}_i[n, k]$ , coding rate of  $\mathbf{c}_i[n, k]$ , and the modulation level of  $t_i[n, k]$  are informed by the adaptive loading processor at the transmitter after receiving the CSI. Eventually, the  $n$ th OFDM symbols at the  $i$ th transmit antenna is generated by passing  $\{t_i[n, k]: k = 0, 1, \dots, N_c\}$  through an IFFT. A CP will be added before transmission to reduce the effect of ISI and ICI.

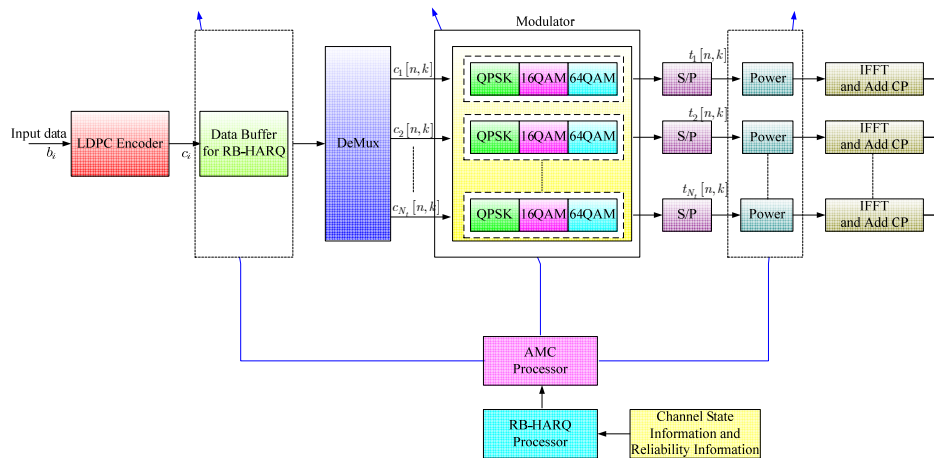


Figure 4.2: Proposed V-BLAST based adaptive MIMO-OFDM system transmitter architecture



unnecessary step because those ill channels are assigned more robust transmission modes to eliminate error bursts. In this way, the traditional interleaving techniques can be omitted as well. Besides, we assume that multiple transmission modes are available. Based on CSI acquired at the receiver, the AMC selector determines the modulation-coding pair (mode), which is sent back to the transmitter through a feedback channel. The AMC controller then updates the transmission mode at the transmitter. The decoded bit streams are mapped to packets, which are pushed upwards to the MAC layer.

We next list the operating assumptions adopted in this thesis.

1. Perfect CSI is available at the receiver using training-based channel estimation. The corresponding mode selection is fed back to the transmitter without error and latency. The assumption that the feedback channel is error free and has no latency, could be at least approximately satisfied by using a fast feedback link with powerful error control for feedback information.
2. Error detection based on LDPC code is perfect. For flat fading channels, the channel quality can be captured by a single parameter, namely the received SNR.
3. At the PHY layer, we deal with packet by packet transmissions.

## **4.4 Adaptive MIMO-OFDM System**

In a system with multiple antennas at the transmitter and/or receiver, the SNR not only varies over time and frequency but also depends on a number of parameters including the way the transmitting signals are mapped and weighted onto the transmit antennas, the processing techniques used at the receiver, and some propagation related parameters such as the pairwise antenna correlation. In this section, we consider a

space-time-frequency scheme, that is V-BLAST based MIMO-OFDM as describe before, and develop a practical LA procedure that integrate temporal, spatial, and spectral components together.

Our problem is to assign bits and power to each sub-channel to maximize data rate for some given power budget and a target BER. The main feature of our approach is to separate the joint space-frequency problem into two separated ones by treating each OFDM tone as a narrow band MIMO. We demonstrate by computer simulations that the V-BLAST based MIMO-OFDM system works well to satisfy the system target BER even in the low SNR scenarios at the cost of degrading transmission rate.

Seeing that a discrete rate set should be used in a practical communication system, a loading criterion for discrete rate is required instead of using the water-filling solution derived by maximizing Shannon capacity. However, a close form solution for optimal discrete rate and power control could not be found and an exhaustive search over the set of rates and powers is too complicated to be conducted in a real time. Hence, some bit loading algorithms were proposed to obtain an optimal (or near optimal) solution [36]-[39]. On account of simplicity and capability, when applying bit loading in the MIMO-OFDM system, we promote that Campello's loading criteria [39] could be somewhat modified and extended to the V-BLAST based adaptive MIMO-OFDM system with reasonable computation complexity.

The joint space-frequency bit loading problem should be taken apart into two separated sub-problems for the following reasons:

1. The active sub-channels should be predetermined before a full search over  $N_c \times N_t$  sub-channels to avoid the unpredictable manner introduced by the V-BLAST detection algorithm.
2. By taking the joint loading problems apart into two smaller ones, the sorting complexity is significantly reduced.

Hence, at the first stage, the loading algorithm is applied to each subcarrier to obtain an optimal bit and power allocation over its  $N_t$  spatial channels. At the second stage, the same loading algorithm is processed over those active sub-channels surviving from the first stage (at most  $N_t \times N_c$ ). The procedures are described as follows:

For each sub-band  $k$  (containing  $N_t$  spatial channels), our allocation problem could be stated as follows:

$$\max \sum_{i=1}^{N_t} c_i[k] \quad \text{subject to} \quad \begin{cases} \sum_{i=1}^{N_t} P_i[k] \leq P_{budget} \\ \varepsilon_i[k] \leq \varepsilon_{error} \quad \forall i \end{cases} \quad (4.2)$$

where  $c_i[k]$ ,  $P_i[k]$ , and  $\varepsilon_i[k]$  are the rate (in bits/symbol), allocated power, and error probability, respectively, of the  $i$ th transmit antenna at the  $k$ th subcarrier,  $\varepsilon_{error}$  is the target BER, and  $P_{budget}$  is the total power constraint ( $P_{budget}$  is normalized to 1 in our simulation to guarantee a fair comparison between systems equipped with different transmit antennas). In general,  $c$  should be restricted to an integer number when the practicability is under consideration. However, if we define  $c$  as data rate, a system using a specified channel encoder along with different puncturing rate will make  $c$  equivalent to some fraction numbers. To facilitate our descriptions, we start with the case without channel coding; in this example,  $c$  is equal to  $b$ .

#### ***A. Initialization***

1. Set  $q$ , defined as the state of the  $N_t$  transmit antennas according to their active modes, to be  $2^{N_t} - 1$  at the first iteration. To realize what  $q$  denotes, we give an example. Assuming that four antennas are available at the transmitter side, if we select three of them, e.g. the 1st, 2nd, and 4th antennas to be active, the state  $q$  will become 13, which is the result of converting the corresponding active-mode vector  $[1, 1, 0, 1]$  to a decimal number. Let  $\text{Rate} = 0$  and  $P_{residual\_final} = P_{budget}$ .



2. Load equal power on those  $N_{t,active}[k]$  active transmit antennas such that

$$P_{i\_active}[k] = \frac{P_{budget}}{N_{t\_active}[k]} \quad \forall 1 \leq i\_active \leq N_{t,active} \quad (4.3)$$

3. Calculate the post-processing SNR of each active layer

$$\rho_{i\_active}[k] = \frac{P_{i\_active}[k]}{\sigma^2 \|\mathbf{w}_{i\_active}[k]\|^2} \quad (4.4)$$

4. Clip the power of each layer and thereby reduce its SNR to fit the nearest threshold below it by consulting the threshold table, and collect the residual power

$$P_{i\_active}[k] \leftarrow \frac{P_{i\_active}[k] \times 10^{(t_k/10)}}{\rho_{i\_active}[k]} \quad (4.5)$$

$$P_{residual}[k] = P_{budget} - \sum_{i=1}^{N_{t,active}[k]} P_{i,active}[k] \quad (4.6)$$

## B. Power-Tightening

1. Since this step is valid for all subcarriers, we drop the index  $k$  for the following expressions. Let  $\Delta p_i(b_i)$  be the power required for  $i$ th layer to increase rate from  $b_i - 1$  to  $b_i$ .  $\Delta p_i(b_i)$  is defined as

$$\Delta p_i(b_i) = \begin{cases} P_i \times 10^{\frac{(t_{b_i} - t_{b_i-1})}{10}} - P_i, & b_i > 1 \\ 10^{\frac{t_{b_i}}{10}}, & b_i = 1 \end{cases} \quad (4.7)$$

We construct a power increment table containing  $\Delta p_i(b_i + 1)$  and  $\Delta p_i(b_i)$   $\forall 1 \leq i \leq N_{t,active}$  to record the least amount of power needed to step from current transmission mode into the next higher-rate mode.

2. Following the principle: A bit distribution is said to be Power-tighten if

$$0 \leq P_{budget} - \sum_{i=1}^{N_t} P_i \leq \min_{1 \leq i \leq N} [\Delta p_i(b_i + 1)] \quad (4.8)$$

3.  $m = \arg \left\{ \min_{1 \leq i \leq N_{t,active}} [\Delta p_i(b_i + 1)] \right\}$
4. While  $p_{residual} > \Delta p_m(b_m + 1)$ 
  - (a)  $P_m \leftarrow P_m + \Delta p_m(b_m + 1)$
  - (b)  $p_{residual} \leftarrow [p_{residual} - \Delta p_m(b_m + 1)]$
  - (c)  $b_m \leftarrow b_m + 1$
  - (d)  $m = \arg \left\{ \min_{1 \leq i \leq N_{t,active}} [\Delta p_i(b_i + 1)] \right\}$

### C. Power-Efficientizing

1. Following the principle: A distribution is said to be efficient if

$$\max_i [\Delta p_i(b_i)] \leq \min_i [\Delta p_i(b_i + 1)] \quad (4.9)$$

2.  $m = \arg \left\{ \min_{1 \leq i \leq N_{t,active}} [\Delta p_i(b_i + 1)] \right\}$ , and  $n = \arg \left\{ \max_{1 \leq i \leq N_{t,active}} [\Delta p_i(b_i)] \right\}$

3. If  $n = m$

- (a)  $l = \arg \left\{ \min_{1 \leq i \neq m \leq N_{t,active}} [\Delta p_i(b_i + 1)] \right\}$ ,

$$\text{and } j = \arg \left\{ \max_{1 \leq i \neq n \leq N_{t,active}} [\Delta p_i(b_i)] \right\}$$

- (b) if  $\Delta p_n(b_n) - \Delta p_l(b_l + 1) > \Delta p_j(b_j) - \Delta p_m(b_m + 1)$ ;  $m \leftarrow l$   
else  $n \leftarrow j$

4. While  $\Delta p_n(b_n) > \Delta p_m(b_m + 1)$

- (a)  $b_m \leftarrow b_m + 1, b_n \leftarrow b_n - 1$

- (b)  $P_m \leftarrow P_m + \Delta p_m(b_m + 1), P_n \leftarrow P_n - \Delta p_n(b_n)$

- (c)  $p_{residual} \leftarrow [p_{residual} + \Delta p_n(b_n) - \Delta p_m(b_m + 1)]$

- (d)  $n \leftarrow \arg \left\{ \max_{1 \leq i \leq N_{t,active}} [\Delta p_i(b_i)] \right\}$ ;  $m \leftarrow \arg \left\{ \min_{1 \leq i \leq N_{t,active}} [\Delta p_i(b_i + 1)] \right\}$

(e) If  $n = m$

$$\text{i. } l = \arg \left\{ \min_{1 \leq i \neq m \leq N_{t,active}} [\Delta p_i(b_i + 1)] \right\},$$

$$\text{and } j = \arg \left\{ \max_{1 \leq i \neq n \leq N_{t,active}} [\Delta p_i(b_i)] \right\}$$

ii. if  $\Delta p_n(b_n) - \Delta p_l(b_l + 1) > \Delta p_j(b_j) - \Delta p_m(b_m + 1)$ ;  $m \leftarrow l$

else  $n \leftarrow j$

Since it makes no sense to add and remove one bit in the same channel at the same time, the case  $n = m$  should be discarded in (e) to avoid this contradiction.

#### **D. Comparing and Recording**

1. If (all the predetermined active layers remain surviving)

$$\text{If } \sum_{i=1}^{N_t} b_i \geq \text{Rate}$$

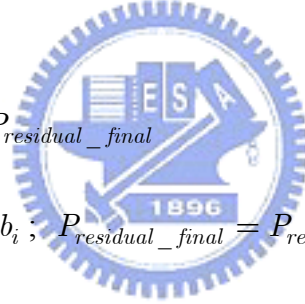
$$\text{If } P_{residual} > P_{residual\_final}$$

$$\text{Rate} = \sum_{i=1}^{N_t} b_i ; P_{residual\_final} = P_{residual}$$

$$\text{else } q \leftarrow q - 1$$

$$\text{else } q \leftarrow q - 1$$

$$\text{else } q \leftarrow q - 1$$



In this stage, we do an exhaustive search to determine which transmit antennas should be active to support the optimal bit and power allocation for every subcarrier. Therefore, given a subcarrier, we should examine  $2^{N_t} - 1$  possible combinations, which seem to be a time-consuming task. Fortunately, the needed effort can be eased due to the causality between every possible combination. For example, given a subcarrier, if we assume that the procedure starts with searching the antenna state  $[1, 1, 1, 1]$  and finally converges to the state  $[1, 0, 1, 0]$ , then the following search starts with the antenna state  $[1, 0, 1, 0]$  will lead to the same result, thus can be omitted.

### Stage 2:

In this stage, the algorithms  $B$  and  $C$  are reused for those sub-channels (both frequency and spatial channels) surviving from the first stage to further exhaust the total residual power, that is  $\sum_{k=1}^{N_c} P_{residual}[k]$ , to get a rate enhancement.

#### B. Power-Tightening

$$0 \leq N_c \cdot P_{budget} - \sum_{i=1}^{N_t} \sum_{k=1}^{N_c} P_i[k] \leq \min_{1 \leq i \leq N_t; 1 \leq k \leq N_c} [\Delta p_i[k](b_i[k] + 1)] \quad (4.10)$$

#### C. Power-Efficientizing

$$\max_{i,k} [p_i[k](b_i[k])] \leq \min_{i,k} [p_i[k](b_i[k] + 1)] \quad (4.11)$$

## 4.5 Combining HARQ with AMC Mechanism

In this section, we develop our cross-layer design, which combines AMC at the PHY layer with RB-HARQ at the MAC layer [40]. Since only finite delays and buffer sizes can be afforded in practice, and the maximum transmission delay depends on the length of a packet and the maximum number of retransmissions. Therefore, the maximum number of ARQ retransmissions has to be bounded. This number can be specified by dividing the maximum allowable system delay over the round trip delay required for each retransmission. Formally, we adopt the following delay constraint.

1. The maximum number of retransmissions allowed per packet is  $R^{\max}$ . If a packet is not received correctly after  $R^{\max}$  retransmissions, we will drop it, and declare packet loss. On the other hand, the error packets can also be utilized if the receiver decides to do so. To maintain an acceptable packet stream, we impose the following performance constraint:
2. The transmission duration of packet after  $R^{\max}$  retransmissions is no longer than maximum delay limit.

3. The probability of packet loss after  $R^{\max}$  retransmissions is no larger than  $P_{\text{loss}}$ .

Hence, constraint 1, 2, and 3 can be derived from the required quality of service in the application at hand. The delay constraint 1 dictates that truncated ARQ with up to  $R^{\max}$  retransmissions should be performed at the data link layer. Having specified ARQ at the data link layer, we next design the AMC at the PHY layer. In other words, we address the following interesting question: with the aid of  $R^{\max}$ -truncated ARQ at the data link layer, how can we optimally design the AMC at the PHY layer to maximize throughput, while guaranteeing the overall system performance dictated by the three constraints.

## 4.5.1 System Performance Requirement at PHY Layer



We first determine how reliable performance is needed at the PHY layer to meet constraint 1, given that  $R^{\max}$ -truncated ARQ is implemented at the data link layer. Notice that a packet is dropped if it is received incorrectly after a maximum number of  $(R^{\max} + 1)$  transmissions; i.e., after  $R^{\max}$  retransmissions. Let us suppose that the instantaneous PER is guaranteed to be no greater than  $P_0$  for each chosen AMC mode at the PHY layer. Then, the packet loss probability at the data link layer is no greater than  $P_0^{R^{\max}+1}$ . To satisfy C2, we need to impose [40]

$$P_0^{R^{\max}+1} \leq P_{\text{loss}} \quad (4.12)$$

From Equation (4.12), we obtain

$$P_0 \leq P_{\text{loss}}^{(1/R^{\max}+1)} = P_{\text{target}} \quad (4.13)$$

Therefore, if we design AMC to satisfy a PER upper-bound as in Equation (4.13) at the PHY layer, and implement a  $R^{\max}$ -truncated ARQ at the data link layer, both delay and performance requirements 1 and 2 will be satisfied. Our remaining problem is to design AMC to maximize spectral efficiency while maintaining Equation (4.13).

## 4.5.2 AMC Design at the PHY Layer

Our objective here is to maximize the data rate, while maintaining the required performance, through AMC at the PHY layer. As we already mentioned, the transmission modes are arranged so that the rate is increasing as the mode index  $n$  creases. Let  $N$  note the total number of transmission modes available. We partition the total SNR range into  $N + 1$  overlapping consecutive intervals, with boundary points denoted as  $\{\gamma_n\}_{n=0}^{N+1}$ . Specifically, [40] mode  $n$  is chosen, when

$$\gamma \in [\gamma_n, \gamma_{n+1}) \quad (4.14)$$

To avoid deep channel fades, no payload bits will be sent when  $\gamma_0 \leq \gamma < \gamma_1$ . What remains now is to determine the boundary points  $\{\gamma_n\}_{n=0}^{N+1}$ .

### 4.5.2.1. Error Performances of PHY Modes

The boundary points are specified in [40] for a given target BER. Finding the target BER through the required PER, and then specifying the boundaries as in [40]. Exact closed-form PERs for the uncoded modulations are provided in [41], while exact closed-form PERs for the coded modulations are not available. To simplify the AMC design, will rely on the following approximate PER expression: [40]

$$\text{PER}_n(\gamma) \approx \begin{cases} 1, & \text{if } 0 < \gamma < \gamma_{pn} \\ a_n \exp(-g_n \gamma), & \text{if } \gamma \geq \gamma_{pn} \end{cases} \quad (4.15)$$

where  $n$  is the mode index and  $\gamma$  is the received SNR. Parameters  $a_n$ ,  $g_n$ , and  $\gamma_{pn}$  in Equation (4.15) are mode-dependent, and are obtained by fitting Equation (4.15) to the exact PER, where the accuracy of this PER approximation is also verified. Using the approximate yet simple expression Equation (4.15) facilitates the mode selection. We set the region boundary (or the switching threshold)  $\gamma_n$  for the transmission mode  $n$  to be the minimum SNR required to achieve  $P_{\text{target}}$ . In general, the required PER in Equation (4.13) satisfies  $P_{\text{target}} < 1$ . Inverting the PER expression in Equation (4.15), we obtain

$$\begin{aligned}\gamma_0 &= 0 \\ \gamma_n &= \frac{1}{g_n} \ln \left( \frac{a_n}{P_{\text{target}}} \right), n = 1, 2, \dots, N \\ \gamma_{N+1} &= +\infty\end{aligned}\quad (4.16)$$

With the  $\gamma_n$  specified by Equation (4.16), one can verify that the AMC in Equation (4.14) guarantees Equation (4.13). Maintaining the target performance, the proposed AMC with the Equation (4.14) and Equation (4.16) then maximizes the spectral efficiency, with the given finite transmission modes.

### 4.5.3 Error Performances of AMC Design

Since the instantaneous PER is upper-bounded by  $P_{\text{loss}}$  in our AMC design, the average PER at the PHY layer will be lower than  $P_{\text{loss}}$ . First, we evaluate this average PER at the PHY layer. According to the AMC rule in Equation (4.14), the transmission mode, and thus the instantaneous PER, depend on the received SNR  $\gamma$ . In each mode  $n$  will be chosen with probability [40]

$$\Pr(n) = \int_{\gamma_n}^{\gamma_{n+1}} p_\gamma(\gamma) d\gamma \quad (4.17)$$

Let  $\overline{PER}_n$  denote the average packet error rate for mode  $n$ . From Equations (4.15) and (4.16), we can derive  $\overline{PER}_n$  as

$$\begin{aligned}
\overline{PER}_n &= \frac{1}{\Pr(n)} \int_{\gamma_n}^{\gamma_{n+1}} PER_n(\gamma) p_\gamma(\gamma) d\gamma \\
&= \frac{1}{\Pr(n)} \int_{\gamma_n}^{\gamma_{n+1}} a_n \exp(-g_n \gamma) p_\gamma(\gamma) d\gamma
\end{aligned} \tag{4.18}$$

The average PER of AMC can then be computed as the ratio of the average number of incorrectly received packets over the total average number of transmitted packets.

$$\overline{PER} = \frac{\sum_{n=1}^N R_n \Pr(n) \overline{PER}_n}{\sum_{n=1}^N R_n \overline{PER}_n} \tag{4.19}$$

where  $R_n$  is the code rate of mode  $n$ .

## 4.6 Adaptive RB-HARQ with AMC Mechanism

In this section we have taken a general look at the throughput by considering the adaptive RB-HARQ scheme and the adaptive modulation assisted MIMO-OFDM system. We separate the problem into two parts. First, we use the Equations (4.12)-(4.16) to determine the modulation order. Second, if the receiver fails in decoding, then we use the adaptive RB-HARQ algorithm to decide the retransmission size of the next transmission.

The throughput  $\eta$  is defined as the average number of information bits per second received successfully. The throughput in bits per second is given by [31]

$$\eta = \frac{I}{T_{trans}} = \frac{I(1-p_J)\tilde{M}}{T_s \left( N_1 + \sum_{j=2}^J N_j p_{j-1} \right)} \tag{4.20}$$

where  $T_{trans}$  is the average cumulative time duration of the packet transmissions until successful decoding and  $T_s$  is the symbol period.  $\tilde{M}$  is the average modulation order



which can obtain from AMC scheme.  $p_j$  is the probability of decoding failure in transmission 1 through  $j$ . That is, [31]

$$\begin{aligned} p_j &= P\left[\bigcap_{k=1}^j \text{Decoding failure in the } k\text{th transmission}\right] \\ &= P[\text{Decoding failure in the } j\text{th transmission}] = BLER_j \end{aligned} \quad (4.21)$$

where  $BLER_j$  denotes the block error rate achieved after the  $j$  transmission.

The average packet delay,  $\delta$ , is expected time that elapses from the moment the packet is first transmitted over the channel to the moment the packet is first transmitted over the channel to the moment the packet is successfully decoded. It is given by [31]

$$\delta = T_{trans} + T_{wait} = T_s \frac{N_1 + \sum_{j=2}^J N_j p_{j-1}}{(1 - p_J) \tilde{M}} + \frac{T_r \sum_{j=1}^J p_j}{1 - p_J} \quad (4.22)$$

where  $T_{trans}$  is defined above,  $T_{wait}$  is the average cumulative time between transmission of a packet until successful decoding, and  $T_r$  is the time interval between transmissions. For simplicity, the delay expression is derived assuming that the interval between attempts to transmit a packet is equal to the time interval between transmissions.

## 4.7 PHY/MAC Cross-Layer AMC Design

Based on the above analysis, the run-time optimal cross layer bit allocation algorithm can be summarized as follows:

1. Estimate the channel condition (various methods exist to estimate the channel condition based on the receiver feedback; these are not discussed here)
2. For the estimated channel SNR and given QoS parameters, determine the target packet error probability  $P_{target}$  from Equation (4.13).
3. Choose maximum retransmission number  $R_{max}$  to obtain the packet error probability to lower than the target packet error probability  $P_{target}$

4. Base on the parameters selected by the previous step, construct SNR threshold table  $\{\gamma_n\}_{n=0}^{N+1}$  by Equation (4.16).
5. When error occurs in receiver, determine the codeword reliability. Using Equation (3.30) to choose the optimal retransmission size.

## 4.8 Computer Simulations

In this section, computer simulations are conducted to evaluate the performance of the proposed V-BLAST based cross layer design AMC system. Throughout the simulation, we only deal with discrete time signal processing in the baseband, hence pulse-shaping and matched-filtering are removed from consideration for simulation simplicity. Also, channel estimation and timing synchronization are assumed to be perfect. Table 4.1 lists all parameters used in our simulation. The configuration we consider here is a MIMO-OFDM system with a bandwidth of 7 MHz and 256 subcarriers. The set of QAM constellation used in the simulation is  $\{0, 4, 16, \text{ and } 64\}$ . The system is an IEEE 802.16-like OFDM mode system. We use the IEEE 802.20 MBWA Channel Model-C which is suited for MIMO broadband wireless access environment.

Firstly, a look-up table that contains the SNR threshold values of each modulation mode should be established. These threshold values could be obtained by dynamically evaluating each PER curve to find the corresponding SNR value that meets the PER requirement ( $10^{-2}$  in our simulations). The look-up table is listed in Table 4.2. In the low SNR scenario, our cross layer design AMC forces some of transmit antennas to be blocked frequently to avoid inefficient or unreliable transmission in order to meet the target PER requirement.

Table 4.1: Simulation parameters of the ZF-BLAST based cross layer design AMC system

Number of TX/RX antennas	1/1, 2/2
Bandwidth	7 MHz
Number of carriers	256
OFDM symbol duration	16 + 2 (CP) $\mu$ s
$2^M$ -QAM available	0, 2, 4, and 6
Channel coding	LDPC code (960, 640) + 20 iterations
Packet loss probability	$10^{-2}$
ARQ types	No ARQ, Conventional ARQ, RB-HARQ
Packet size	640 bits
Channel model	802.20-C , 6 paths model, $\tau_{rms} = 1.27 \mu$ s

Table 4.2: SNR threshold table for various M-QAM at the target PER= $10^{-2}$

SNR (dB)	No TX	QPSK	16-QAM	64-QAM
AMC-only	$-\infty$	3.91	9.92	15.01
ARQ, RB-HARQ	$-\infty$	3.47	9.41	14.43

Figure 4.4 shows the throughput performances for the cross layer design AMC system with the proposed RB-HARQ scheme, the ARQ scheme and the AMC-only scheme under  $(N_t, N_r) = (2, 2)$ . As shown in the figure, when the SNR increases, the throughput performance of each scheme increases proportionally. Moreover, we can observe that the throughput performance with the ARQ scheme is close to that of the AMC-only scheme. When the transmitter has perfect CSI, it can always choose a suitable mode which can satisfy the prescribed QoS requirement. Compared with the AMC-only scheme, the ARQ scheme has better error-correcting capability, which relieves the PHY layer from stringent error correction requirements. With a lower PER requirement, the transmission rate can be increased at the PHY layer. This in turn leads to the improvement of the transmission rate and the deterioration of the PER performance. This is the reason why the throughput performance of the ARQ scheme is close to that of the AMC-only scheme. The RB-HARQ scheme leads to an increase of

about 5 Mbps in the throughput over the ARQ scheme and the AMC-only scheme because only a few bits are needed to be retransmitted in the RB-HARQ scheme. The comparison of retransmission data rates between the RB-HARQ scheme and the ARQ scheme is shown in Figure 4.5. The retransmission data rate is defined as the total data rate which contains the retransmission information. As shown in Figure 4.6, the average PER of the RB-HARQ scheme is lower than the required  $P_{loss}$ . Figures 4.7–4.9 show the performances of throughput, retransmission data rate, and PER respectively under  $(N_t, N_r) = (1, 1)$ . The results are similar to the case under  $(N_t, N_r) = (2, 2)$ .

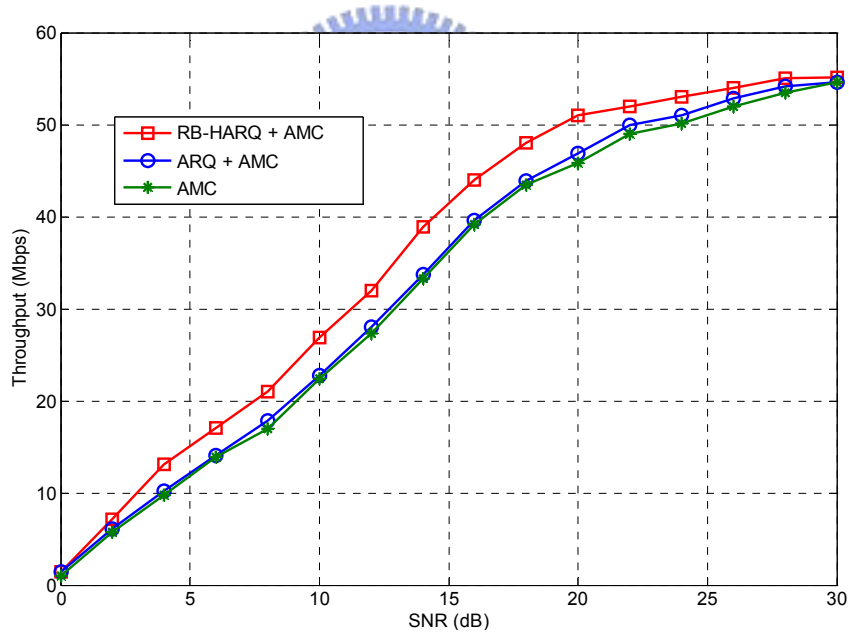


Figure 4.4: Throughput versus average SNR for the cross layer design AMC system (required  $P_{loss} = 10^{-2}$ ) in IEEE 802.20 Model-C channel.  $(N_t, N_r) = (2, 2)$  with RB-HARQ scheme, ARQ scheme, and AMC-only scheme. Other parameters are listed in Table 4.1

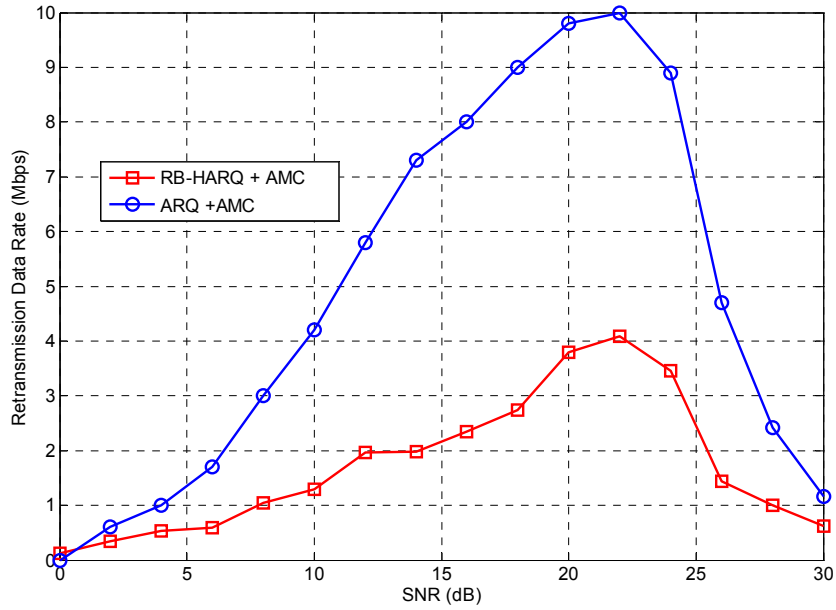


Figure 4.5: Retransmission data rate versus average SNR for the cross layer design AMC system (required  $P_{loss} = 10^{-2}$ ) in IEEE 802.20 Model-C channel.  $(N_t, N_r) = (2, 2)$  with RB-HARQ scheme, and ARQ scheme. Other parameters are listed in Table 4.1

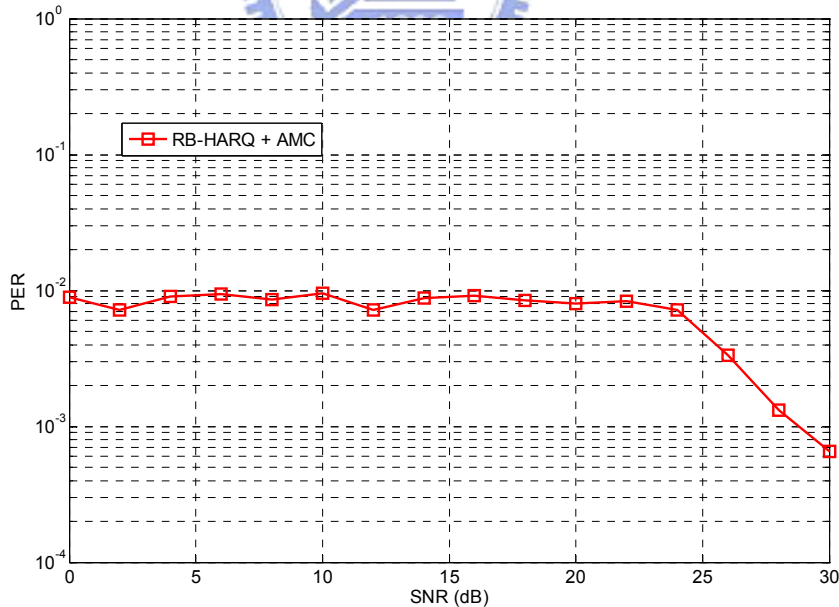


Figure 4.6: PER versus average SNR for the cross layer design AMC system (required  $P_{loss} = 10^{-2}$ ) in IEEE 802.20 Model-C channel.  $(N_t, N_r) = (2, 2)$  with RB-HARQ scheme. Other parameters are listed in Table 4.1

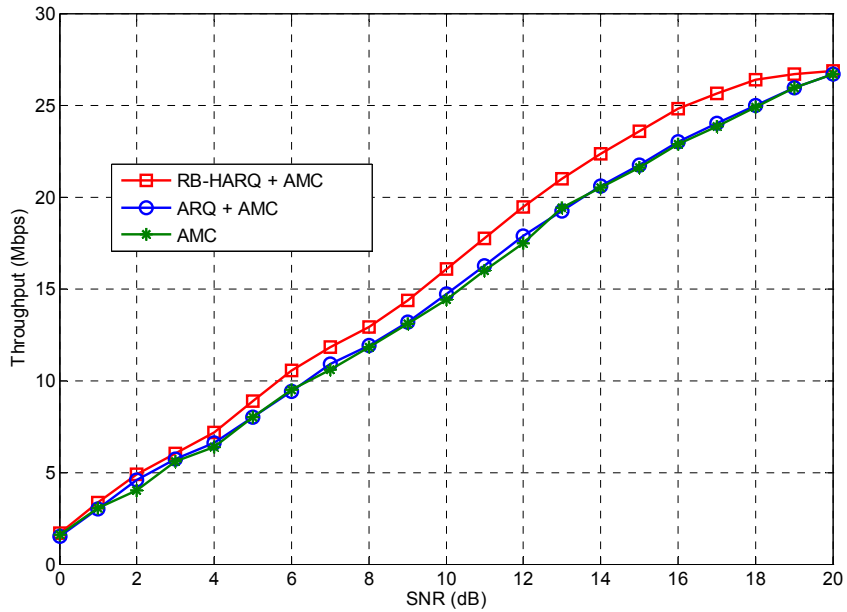


Figure 4.7: Throughput versus average SNR for the cross layer design AMC system (required  $P_{loss} = 10^{-2}$ ) in IEEE 802.20 Model-C channel.  $(N_t, N_r) = (1, 1)$  with RB-HARQ scheme, ARQ scheme, and AMC-only scheme. Other parameters are listed in Table 4.1

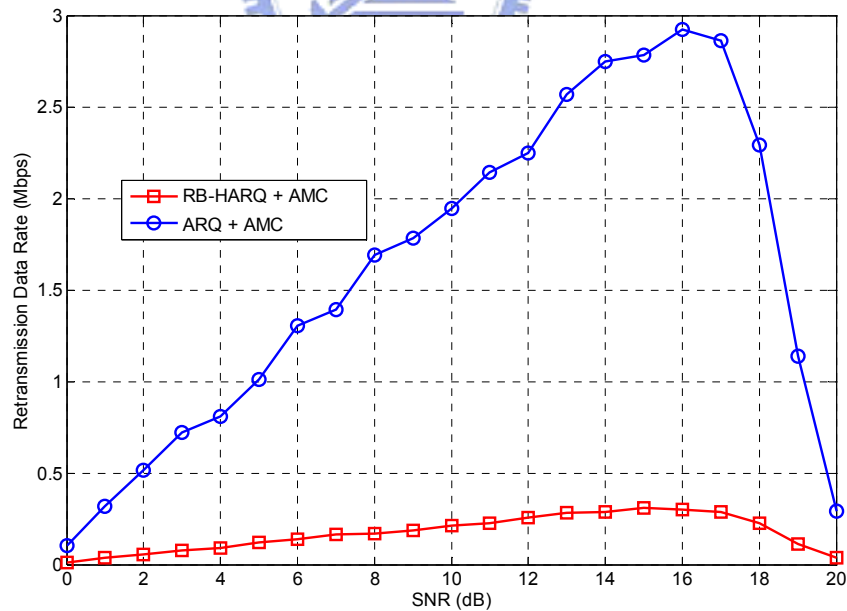


Figure 4.8: Retransmission data rate versus average SNR for the cross layer design AMC system (required  $P_{loss} = 10^{-2}$ ) in IEEE 802.20 Model-C channel.  $(N_t, N_r) = (1, 1)$  with RB-HARQ scheme, and ARQ scheme. Other parameters are listed in Table 4.1

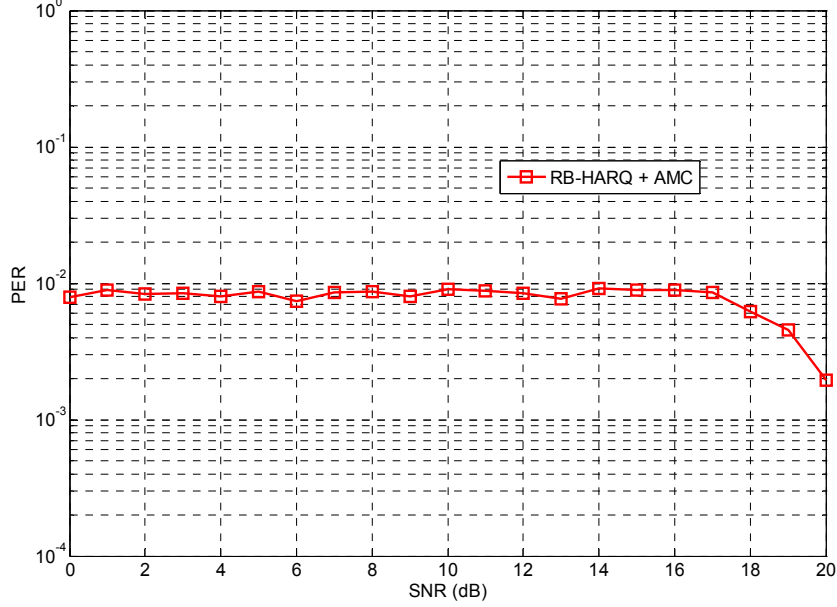


Figure 4.9: PER versus average SNR for the cross layer design AMC system (required  $P_{loss} = 10^{-2}$ ) in IEEE 802.20 Model-C channel.  $(N_t, N_r) = (1, 1)$  with RB-HARQ scheme. Other parameters are listed in Table 4.1

In the above simulations, we assume that perfect CSI is available at the transmitter. However, it may not always hold true due to many different reasons, e.g. channel estimation error, and/or feedback delay combined with Doppler spread and quantization error. Then, we consider a system with partial (imperfect) CSI at the transmitter (CSIT), but with full (perfect) CSI at the receiver. The transmitter acquires channel knowledge either via a feedback channel, or, by channel estimation in a TDD operation. The partial CSIT includes the frequency domain channel feedback  $\bar{\mathbf{h}}_{f,0}$  that is treated as deterministic mean plus a perturbation (error) term  $\tilde{\mathbf{h}}_f$  with known pdf to account for various sources of uncertainty [42], i.e.,

$$\hat{\mathbf{h}}_f = \bar{\mathbf{h}}_{f,0} + \tilde{\mathbf{h}}_f \quad (4.23)$$

where  $\hat{\mathbf{h}}_f \sim CN(\bar{\mathbf{h}}_{f,0}, \Sigma_{\tilde{\mathbf{h}}_f})$ . In the simulation, we assume the error covariance matrix  $\Sigma_{\tilde{\mathbf{h}}_f} = 0.01\mathbf{I}$ .

Figure 4.10 shows the comparison of the throughput performances among the RB-HARQ scheme, the ARQ scheme, and the AMC-only scheme under partial CSIT. We observe that the RB-HARQ scheme leads to an increase of about 5 Mbps in the throughput over the ARQ scheme and the AMC-only scheme. Unlike the perfect CSIT case, the PER performance of the ARQ scheme is better than that of the AMC-only scheme since the ARQ scheme can use ACK/NACK to mitigate the effect of partial CSIT. The PER performance of the RB-HARQ scheme is still the best one. The PER performance in the presence of partial (imperfect) CSI at the transmitter is shown in Figure 4.11. In the presence of partial CSIT, all the schemes cannot meet the required  $P_{loss}$ .

In the RB-HARQ scheme, the packet reliability will become smaller when the perfect CSIT is not available at the transmitter. With the smaller packet reliability, the receiver request the greater number of retransmission bits. So the effect of partial CSIT degrades the PER performance slightly.

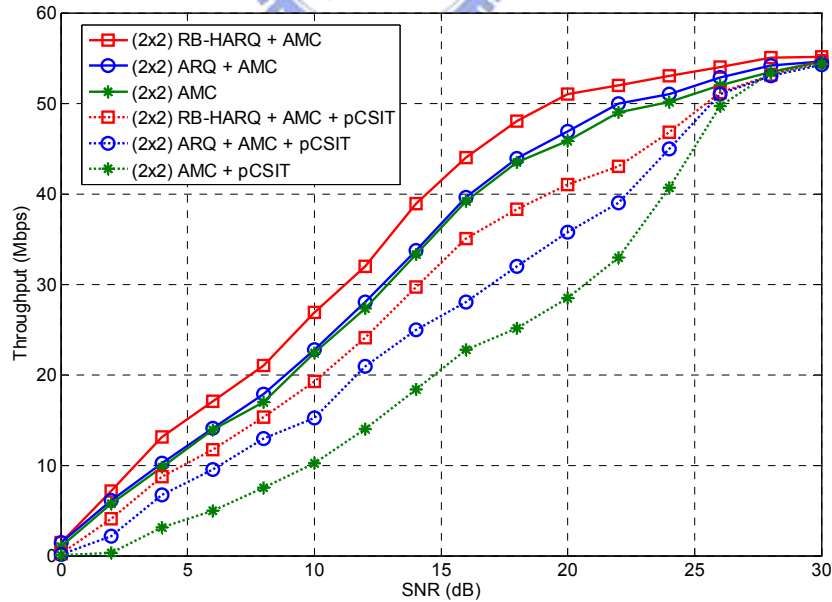


Figure 4.10: Throughput versus average SNR for the cross layer design AMC system (required  $P_{loss} = 10^{-2}$ ) in IEEE 802.20 Model-C channel in the presence of partial CSIT  $\Sigma_{h_f} = 0.01\mathbf{I}$ .  $(N_t, N_r) = (2, 2)$  with RB-HARQ scheme, ARQ scheme, and AMC-only scheme. Other parameters are listed in Table 4.1



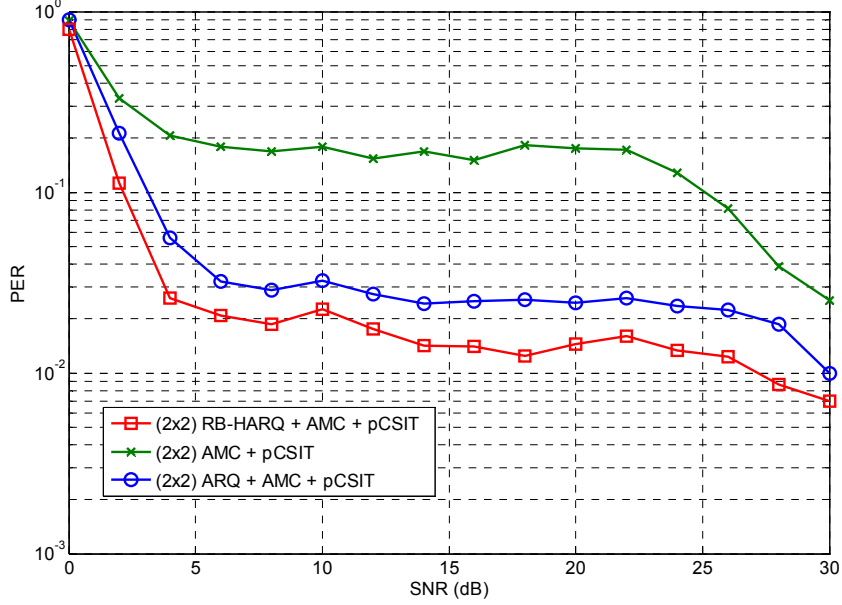


Figure 4.11: PER versus average SNR for the cross layer design AMC system (required  $P_{loss} = 10^{-2}$ ) in IEEE 802.20 Model-C channel in the presence of partial CSIT  $\Sigma_{\tilde{h}_f} = 0.01\mathbf{I}$ .  $(N_t, N_r) = (2, 2)$  with RB-HARQ scheme, ARQ scheme, and AMC-only scheme. Other parameters are listed in Table 4.1

## 4.9 Summary

Maximizing throughput in a wireless channel is a very important issue for the quality of multi-media or data transmission. The cross layer design AMC system combines the AMC scheme at the PHY layer and the RB-HARQ at the MAC layer, in order to maximize the system throughput under prescribed delay and performance constrains. The RB-HARQ scheme can alleviate stringent error-control requirements on modulation and coding, and bring considerable throughput performance gain. Computer simulations results demonstrate the throughput and PER performance improvement of the cross-layer design AMC with RB-HARQ.

# Chapter 5

## Conclusion

In this thesis, we survey the MIMO techniques and explore their applications in the OFDM systems. With the aid of OFDM, we show that MIMO can find a feasible way in extending itself to wideband transmission based on the flat fading channels introduced by OFDM tones. In Chapter 2, we present V-BLAST based MIMO-OFDM architecture, aimed to provide high spectral efficiency.

In Chapter 3, we compare several HARQ schemes and conclude that RB-HARQ is the best scheme. We propose an adaptive algorithm which can accordingly choose the sizes of retransmissions under the QoS constraints such as the maximum number of retransmissions allowed per packet and packet loss probability. In the simulations, the proposed adaptive algorithm has a significant performance gain over other schemes.

Although the adaptive MIMO-OFDM systems can enjoy both the diversity and multiplexing gains in a flexible manner, it does not consider the requirement of higher layer. By analytical throughput computation, the adaptive MIMO-OFDM systems at PHY layer will not lead to a significant throughput increase at higher layers.

Observing that the network performance in mobile wireless applications is determined largely by a complex interaction between PHY layer, and MAC layer, a joint design of these layers was proposed in Chapter 4. The cross layer design AMC system combines adaptive modulation and coding scheme at the PHY layer and MAC

protection strategies at the MAC layer to maximize the system throughput under prescribed delay and performance constraints. It presents a new point of view in which SM and SD are complementary rather than competing approaches. By adjusting the transmission parameters to prevent ill-conditioned sub-channels from dominating the system performance, an indirect form of diversity is also drawn. Computer simulation results demonstrated the throughput and PER performance improvement of the cross-layer design AMC.

As a remark for addressing the ability of system extension, the proposed adaptive MIMO transceiver is an example of link adaptation which can be achieved by adjusting the transmission parameters. It is noted that the proposed adaptive MIMO transceiver is essentially an adaptive modulation and coding scheme at the PHY layer, and integrated with RB-HARQ at the MAC layer to achieve the QoS demands under prescribed delay and error performance constraints. Finally, the assumptions that perfect CSI is available at the receiver, and the feedback channel has zero delay and is error free, may not always hold true. One possible extension of this work is to design and analyze the cross layer design with imperfect CSIT. Besides, the impact of cross layer design on other parameters at the PHY and higher layers is also worth investigating.

# Bibliography

- [1] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Labs Syst. Tech. J.*, vol. 1, pp. 41-59, Autumn 1996.
- [2] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, no. 3, pp. 311-335, 1998.
- [3] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel," *URSI International Symposium*, pp. 295-300, Oct. 1998.
- [4] X. Li, H. Huang, G. J. Foschini, and R. A. Valenzuela, "Effects of iterative detection and decoding on the performance of BLAST," *IEEE GLOBECOM*, vol. 2, pp. 1061-1066, Nov. 2000.
- [5] E. Biglieri, G. Taricco and A. Tulino, "Decoding space-time codes with BLAST architectures," *IEEE Trans. Signal Processing*, vol. 50, no. 10, pp. 2547-2552, Oct. 2002.
- [6] Y. Li, J. H. Winters, and N. R. Sollenberger, "MIMO-OFDM for wireless communications: signal detection with enhanced channel estimation," *IEEE Trans. Commu.*, vol. 50, no. 9, pp. 1471-1477, Sep. 2002.
- [7] S. Catreux, D. Gesbert, and V. Erceg, "Adaptive modulation and MIMO coding for broadband wireless data networks," *IEEE Commu. Mag.*, vol. 40, no.6, Jun. 2002.
- [8] S. Shim, J. S. Choi, C. Lee, and D. H. Youn, "Rank adaptive transmission to improve the detection performance of the BLAST in spatially correlated MIMO channel," *IEEE VTC 2002-Fall*, vol. 1, pp. 195-198, Sep. 2002.
- [9] A. Goldsmith and S. Chua, "Variable-rate variable-power MQAM for fading channels," *IEEE Trans. Commu.*, vol. 45, no. 10, pp. 1218-1230, Oct. 1997.

- [10] W. T. Webb and R. Steele, "Variable rate QAM for mobile radio," *IEEE Trans. Commu.*, vol. 43, no. 7, pp. 2223-2230, Jul. 1995.
- [11] Helmut Bölcskei, A. J. Paulraj, et al, "Fixed broadband wireless access: state of the art, challenges, and future directions," *IEEE Commu. Mag.*, vol. 39, no. 1, pp. 100-108, Jan. 2001.
- [12] G. G. Raleigh and J. M. Cioffi, "Spatial-temporal coding for wireless communications," *Proc. IEEE 1996 Global Commu. Conference*, pp. 1809-1814, Nov. 1996.
- [13] IEEE Std 802.16e™-2005 and IEEE Std 802.16™-2004/Cor1-2005, "Part 16: air interface for fixed and mobile broadband wireless access systems," Feb. 2006.
- [14] IEEE Std 802.16™-2004, "Part 16: air interface for fixed broadband wireless access systems," Oct. 2004.
- [15] J. M. Shea, "Reliability-based hybrid ARQ," *IEE Electronics Letters*, vol. 38, no. 13, pp. 644-645, Jun. 2002.
- [16] A. Roongta, J. W. Moon and J. M. Shea, "Reliability-based hybrid ARQ as an adaptive response to jamming," *IEEE JSAC*, vol.23, no. 5, pp. 1045-1055, May 2005.
- [17] Y. Cao, Y. Wang, and H. Yang, "Performance analysis of RB-HARQ with LDPC," *International Conference on Wireless Communications, Networking and Mobile Computing*, vol. 1, pp. 411-414, Sep. 2005.
- [18] A. Avudainayagam, J.M. Shea, and A. Roongta, "Improving the efficiency of reliability-based Hybrid-ARQ with convolutional codes," *IEEE MILCOM 2004*, pp. 1-7, Oct. 2005.
- [19] A. Roongta and J.M. Shea, "Reliability-based hybrid ARQ and rate-compatible punctured convolutional codes," *IEEE WCNC. 2004*, vol.4 pp. 2105-2109, Mar. 2004.
- [20] A. Roongta and J.M. Shea, "Reliability-based hybrid ARQ using convolutional codes," *IEEE ICC'03*, vol. 4, pp. 11-15, May 2003.
- [21] "802.20 Channel Models Document for IEEE 802.20 MBWA System Simulations," 802.20-PD-08, Sep. 2005.
- [22] R. G. Gallager, "Low-density parity-check codes," *IRE Trans. inform. Theory*, pp. 21-28, Jan. 1962.

- [23] T. K. Moon, *Error Correction Coding, Mathematical Methods and Algorithms*, Wiley, 2005.
- [24] S. Haykin, *Communication Systems 4<sup>th</sup> edition*, Wiley, 2000.
- [25] P. Berlin and D. Tuninetti, "LDPC codes for fading Gaussian broadcast channels," *IEEE Trans. Info. Theory*, vol. 51, no. 6, pp. 2173-2182, Jun. 2005.
- [26] D. J. C. Mackay, "Good error-correcting codes based on very sparse matrices," *IEEE Trans. Info. Theory*, vol. 45, no. 2, pp. 399-431, Mar. 1999.
- [27] A. Goldsmith, *Wireless Communications*, Cambridge University Press, 2005.
- [28] J. D. Gibson and E. Gibson, *The Communications Handbook*, CRC Press, 1997.
- [29] Y. J. Guo, *Advances in Mobile Radio Access Networks*, Artech House Publishers, 2004.
- [30] S. Sesia, G. Caire and G. Vivier, "Incremental redundancy hybrid ARQ schemes based on low-density parity-check codes," *IEEE Trans. Commun.*, vol. 52, no. 8, pp. 1311-1321, Aug. 2004.
- [31] E. Visotsky, S. Yakun, V. Tripathi, M. L. Honig, and R. Peterson, "Reliability-based incremental redundancy with convolutional codes," *IEEE Trans. Commun.*, vol.53, no.6, pp. 987-997, Jun. 2005.
- [32] M. V. D. Schaar, S. Krishnamachari, S. Choi, and X. Xu, "Adaptive cross-layer protection strategies for robust scalable video transmission over IEEE 802.11 WLANs," *IEEE JSAC*, vol. 21, no. 10, pp. 1752-1763, Dec. 2003.
- [33] A. Alexiou, and M. Haardt, "Smart antenna technologies for future wireless systems: trends and challenges," *IEEE Commun. Mag.*, vol. 42, no. 9, pp. 90-97, Sep. 2004.
- [34] H. Sampath, S. Talwar, J. Tellado, V. Erceg and A. Paulraj, "A fourth-generation MIMO-OFDM broadband wireless systems: design, performance, and field trial results," *IEEE Commun. Mag.*, vol. 40, no. 9, pp. 143-149, Sep. 2002.
- [35] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Commun*, vol. 10, no. 6, pp. 585-595, 1999.
- [36] Helmut Bölcskei, A. J. Paulraj, et al, "Fixed broadband wireless access: state of the art, challenges, and future directions," *IEEE Commun. Mag.*, vol. 39, no. 1, pp. 100-108, Jan. 2001.

- [37] D. Dardari, "Ordered subcarrier selection algorithm for OFDM-based high-speed WLANs," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1452-1458, Sep. 2004.
- [38] Q. Liu, S. Zhou, and G. B. Giannakis, "Cross-layer combining of adaptive modulation and coding with truncated ARQ over wireless Links," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 1746-1755, Sep. 2004.
- [39] J. Campello and D. Souza, "Discrete bit loading for multicarrier modulation systems," PhD. Dissertation, Stanford University, 1999.
- [40] A. L. Toledo, X. Wang, and B. Lu, "A cross-layer TCP modeling framework for MIMO wireless systems," *IEEE Trans. Wireless Commun.*, vol. 5, no. 4, pp. 920-929, Apr. 2006.
- [41] A. M. Wyglinski, F. Labeau, and P. Kabal, "Bit loading with BER-constraint for multicarrier systems," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1383-1387, Jul. 2005.
- [42] Y. Yingwei, and G. B. Giannakis, "Rate-maximizing power allocation in OFDM based on partial channel knowledge," *IEEE Trans. Wireless Commun.*, vol. 4, no. 3, pp. 1073-1083, May 2005.

