國 立 交 通 大 學

電信工程學系碩士班
碩士論文

動態副載波配置與適應性調變技術在多用
戶多輸入多輸出正交分頻多工系統的應用

Applications of Dynamic Subcarrier Allocation

and Adaptive Modulation in Multiuser

MIMO-OFDM Systems

研 究 生：王建中　　　　　　Student: Jiann-Jong Wang

指導教授：李大嵩 博士　　　　Advisor: Dr. Ta-Sung Lee

中 華 民 國 九 十 三 年 六 月

動態副載波配置與適應性調變技術在多用戶
多輸入多輸出正交分頻多工系統的應用

Applications of Dynamic Subcarrier Allocation
and Adaptive Modulation in Multiuser
MIMO-OFDM Systems

研 究 生：王建中　　　　　Student: Jiann-Jong Wang

指導教授：李大嵩 博士　　　Advisor: Dr. Ta-Sung Lee

國立交通大學

電信工程學系碩士班

碩士論文

中 華 民 國 九 十 三 年 六 月

動態副載波配置與適應性調變技術在多用戶多輸入多輸出正交分頻多工系統的應用

學生：王建中　　　　　　　　　　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

　　多輸入多輸出為使用多天線於傳送和接收端的可靠通訊技術，並被認為是符合第四代高速通訊需求的最佳方案之一。透過空間多工的方式，多輸入多輸出技術可在空間中的獨立平行通道傳送不同資料串流，藉以提昇系統的整體傳輸速率。另一方面，正交分頻多工為一種具高頻譜效益，並能有效克服多重路徑衰減效應的調變技術，尤其適用於多用戶系統中。在本論文中，吾人將探討結合多輸入多輸出技術與多用戶正交分頻多工系統的通訊系統架構。基於相同副載波對於不同用戶會展現不同通道條件的現象，吾人將針對多用戶正交分頻多工系統提出一種動態副載波配置演算法。此演算法考慮個別用戶對於服務品質與傳輸速率不同的需求，配置一組最適當的副載波給每一用戶，藉以提昇系統的整體傳輸速率。此外，吾人更進一步針對多用戶多輸入多輸出正交分頻多工系統提出一種適應性傳收架構及位元負載演算法，使系統能夠隨時間動態地在頻率與空間通道上調整傳輸參數—例如，調變階數與傳輸能量—以便充分地利用空間、時間以及頻率通道上的特性以維持系統的目標錯誤率，同時更進一步提昇系統的整體傳輸速率。最後，吾人藉由電腦模擬驗證上述架構在無線通訊環境中具有優異的傳輸表現。

Applications of Dynamic Subcarrier Allocation and Adaptive Modulation in Multiuser MIMO-OFDM Systems

Student: Jiann-Jong Wang   Advisor: Dr. Ta-Sung Lee

Institute of Communication Engineering

National Chiao Tung University

Abstract

Multiple-input multiple-output (MIMO) is a promising technique suited to the increasing demand for high-speed 4G broadband wireless communications. Through spatial multiplexing, the MIMO technology can transmit multiple data streams in independent parallel spatial channels, thereby increasing the total transmission rate of the system. On the other hand, orthogonal frequency division multiplexing (OFDM) is a high spectral efficiency modulation technique that can deal efficiently with multipath fading effects especially suited to multiuser systems. In this thesis, a new wireless communication system combining MIMO and multiuser OFDM system is considered, called the multiuser MIMO-OFDM system. On account of the same subcarrier experiencing different channel conditions for different users, a dynamic subcarrier allocation algorithm is proposed. This algorithm considers user-specific quality of service (QoS) and transmission rate requirements and allocates to each user the most appropriate subcarriers, thereby enhancing the overall transmission rate of the system. Furthermore, an adaptive multiuser MIMO-OFDM transceiver architecture along with a bit loading algorithm is proposed, which dynamically adjusts the transmission parameters such as modulation order and transmit power over spatial and frequency channels to fully exploit the properties of the space-time-frequency channels to meet the target bit error rate (BER) and further enhance the overall transmission rate of the system. Finally, the performance of the proposed systems is evaluated by computer simulations, confirming that they work well in wireless communication environments.

Acknowledgement

I would like to express my deepest gratitude to my advisor, Dr. Ta-Sung Lee, for his enthusiastic guidance and great patience. Heartfelt thanks are also offered to all members in the Communication Signal Processing (CSP) Lab for their constant encouragement and help. Finally, I would like to show my sincere thanks to my parents and a special friend for their inspiration and love.

Contents

List of Figures

List of Tables

Acronym Glossary

ADC	analog-to-digital conversion
AWGN	additive white Gaussian noise
BER	bit error rate
BLAST	Bell Lab Layered space time
BPSK	binary phase shift keying
BS	base station
CCI	co-channel interference
CNR	channel gain-to-noise ratio
CP	cyclic prefix
CRC	cyclic redundancy check
CSI	channel state information
D-BLAST	diagonal Bell labs' layered space-time
DFT	discrete Fourier transform
FDMA	frequency division multiple access
FFT	fast Fourier transform
FSK	frequency shift keying
ICI	intercarrier interference
IEEE	institute of electrical and electronics engineers
IFFT	inverse fast Fourier transform
ISI	intersymbol interference
LOS	line of sight
MAI	multiple access interference
MCM	multicarrier modulation
MIMO	multiple-input multiple-output
MISO	multiple-input single-output

MMSE	minimum mean square error
MS	mobile station
OFDM	orthogonal frequency division multiplexing
PER	packet error rate
QAM	quadrature amplitude modulation
QoS	quality of service
QPSK	quaternary phase shift keying
RF	radio frequency
SD	spatial diversity
SDR	software defined radio
SIMO	single-input multiple-output
SISO	single-input single-output
SM	spatial multiplexing
SNR	signal-to-noise ratio
STC	space-time coding
SVD	singular value decomposition
TDD	time division duplex
TDMA	time division multiple access
V-BLAST	vertical Bell laboratory layered space-time
WF	water filling
ZF	zero forcing
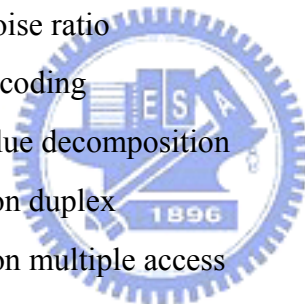
Notations

C channel capacity

E_b bit energy

E_s symbol energy

$h_{i,j}$ channel gain between the jth transmit and ith receive antenna

M modulation order

N_c number of subcarriers

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_{budget} power budget

q antenna state

S set of signal constellation

$\mathbf{T}$ set of switching levels

T_s symbol duration

T_{sample} sampling period

$\mathbf{w}_k$ weighting vector for the kth layer

σ_n^2 noise power

ε_{error} target BER

γ instantaneous SNR

τ_{rms} root mean squared excess delay spread

ρ average SNR at each receive antenna

λ eigenvalue

Chapter 1

Introduction

The increasing popularity of enhanced communication services, such as wireless multimedia, telecommuting, fast Internet access, and video conferencing, has promoted the growing demand of high data rate, high mobility, and high quality of service (QoS) requirements for users. However, the limited available bandwidth drives the wireless communication technology towards the emerging issue of high spectral efficiency. Besides, in wireless channels, the time-selective and frequency-selective fading caused by multipath propagation, carrier frequency/phase shift, and Doppler shift limit the developments of high data rate and reliable communications. As a remedy, some efficient modulation and coding schemes such as coded multicarrier modulation, multiple-input multiple-output (MIMO) technology [1]-[12], and adaptive resource allocation [13]-[17] are proposed to enhance the spectral efficiency and quality of wireless communication links.

MIMO techniques, i.e., a radio communication system equipped with multiple antenna elements at both the transmitting end and receiving end, have been demonstrated to effectively increase the transmission capacity and support high data rate. The signals on the transmit antennas at one end and on the receive antennas at the other end are "co-processed" in such a way that the quality or the data rate of the communication link can be improved. The core idea of MIMO 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. Furthermore, the key feature of MIMO systems is to efficiently exploit the multipaths, rather than mitigate

them, 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.

The high data rate wireless transmission over the multipath fading channels is mainly limited by intersymbol interference (ISI). Orthogonal frequency division multiplexing (OFDM) [18]-[19] has been considered as a reliable technology to deal with the ISI problem. The principle of the OFDM technology is to split a high data rate stream into a number of low data rate streams which are simultaneously transmitted on a number of orthogonal subcarriers. By adding a cyclic prefix (CP) to each OFDM symbol, both intersymbol and intercarrier interference can be removed and the channel also appears to be circular if the CP length is longer than the channel length. The multicarrier property of OFDM systems can not only improve the immunity to fast fading channels, but also make multiple access possible because the subcarriers are independent of each other.

OFDM combining antenna arrays at both the transmitter and receiver, which leads to a MIMO-OFDM configuration, can significantly increase the diversity gain or enhance the system capacity over time-variant and frequency-selective channels. Typical MIMO-OFDM systems can be categorized into two types: those based on spatial multiplexing (SM) [1], [20] and those based on spatial diversity (SD) schemes. The former system is a layered spatial transmission scheme, in which different data streams are transmitted from different transmit antennas simultaneously and received through nulling and canceling to mitigate the co-channel interference (CCI). The latter one uses space-time coding (STC) techniques to improve the transmission reliability. STC can provide diversity gain and increase the effective transmission rate without sacrificing the bandwidth.

In the multiuser MIMO-OFDM system, each of the multiple users' signals may undergo independent fading due to different locations of users. Therefore, the subcarriers in deep fade for one user may not be in the same fade for other users. In fact, it is quite unlikely that a subcarrier will be in deep fade for all users. Hence, for a specific subcarrier, the user with the best channel quality can use the subcarrier to transmit data yielding multiuser diversity effects [21]. Recently, methods for

dynamically assigning subcarriers to each user have been widely investigated [23]-[24]. These dynamic subcarrier allocation algorithms can be geared to decrease the power consumption for a given achievable data rate or to increase the data rate when the available power is limited. In this thesis, a dynamic subcarrier allocation algorithm suited to the multiuser OFDM system is developed to take both user-specific data rate and quality of service (QoS) requirements into account and allocate to each user the most appropriate subcarriers.

As the instantaneous channel state information (CSI) is determined beforehand, the multiuser MIMO-OFDM system incorporating the adaptive modulation technique can provide a significant performance improvement. Adaptive modulation can dynamically adjust transmission parameters to alleviate the effects of channel impairments. Subcarriers with good channel qualities can employ higher modulation order to carry more bits per OFDM symbol, while subcarriers in deep fade may employ lower modulation order or even no transmission. In addition, the adaptive modulation technique must take into account the additional signaling dimensions explored in future broadband wireless networks [16]. More specifically, the growing popularity of both MIMO and multiuser OFDM systems creates the demand for link adaptation solutions to integrate temporal, spectral, and spatial components together. In this thesis, an adaptive wireless transceiver is developed to effectively exploit the available degrees of freedom in wireless communication systems.

This thesis is organized as follows. In Chapter 2, the general system model and channel capacity of a MIMO communication link are described. In Chapter 3, the multiple access concepts of the multiuser OFDM system are introduced and a dynamic subcarrier allocation algorithm is proposed which allows users to choose the most appropriate subcarriers according to their requirements and the channel qualities. In Chapter 4, the adaptive modulation concepts are introduced and an adaptive bit loading algorithm suited to the multiuser MIMO-OFDM system is proposed to further increase the total transmission rate and still meet the target bit error rate (BER). Finally, Chapter 5 gives concluding remarks of this thesis and leads the way to some potential future works.

Chapter 2

Overview of MIMO Systems

Digital communication using multiple-input multiple-output (MIMO) has recently emerged as one of the most significant technical breakthroughs in wireless communications. The technology figures prominently on the list of recent technical advances with a chance of resolving the bottleneck of traffic capacity in future Internet intensive wireless networks. In this chapter, the basic ideas and key features of the MIMO system will be introduced.
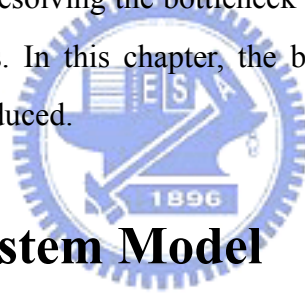
2.1   MIMO System Model

MIMO system architectures provide substantially better spectral efficiency than traditional systems. With this transmission scheme, there is a linear increase in spectral efficiency compared to a logarithmic increase in traditional systems utilizing receive diversity. The increase in spectral efficiency is based on the utilization of antenna or space diversity both at the transmitter and receiver sides in MIMO systems.

A key feature of MIMO systems is the ability to turn multipath propagation, traditionally a pitfall of wireless transmission, into a benefit for the user. MIMO systems effectively take advantage of random fading and multipath delay spread for multiplying transfer rates. In a rich scattering environment, the signals transmitted from each transmit antenna appear highly uncorrelated at each receive antenna. When the transmitter transmits signals to the receiver through uncorrelated channels, the signals from each transmit antenna obtain different spatial signatures. Then the receiver can exploit these different spatial signatures to separate the signals transmitted from

different transmit antennas but at the same frequency band simultaneously.

MIMO systems can be defined simply. Given an arbitrary wireless communication system, a link is considered for which the transmitting end is equipped with N_t transmit antennas and receiving end is equipped with N_r receive antennas. Such a setup is illustrated in Fig. 2.1.

Consider a communication link comprising N_t transmit antennas and N_r receive antennas that contributes to a MIMO channel, as shown in Fig. 2.1. Some important assumptions are made firstly:

1. The channel is constant during the transmission of a packet. It means the communication is carried out in packets that are of shorter time-span than the coherence time of the channel.

2. The channel is memoryless. It means that an independent realization of channel is drawn for each use of the channel. It also means that the channel capacity can be computed as the maximum of the mutual information as defined in Equation (2.6).

3. The channel is frequency-flat fading. It means that constant fading over the bandwidth is deserved when the case of narrowband transmission is dealt with. It also indicates that the channel gain can be represented as a complex number.

4. Only a single user transmits signals at any given time, so the received signal is corrupted by AWGN only.

With these assumptions, it is common to represent the input/output relations of a narrowband, single user MIMO link by the complex baseband vector notation

$$\mathbf{y} = \mathbf{Hx} + \mathbf{n} \tag{2.1}$$

where $\mathbf{x}$ is the $N_t \times 1$ transmit vector, $\mathbf{y}$ is the $N_r \times 1$ receive vector, $\mathbf{H}$ is the $N_r \times N_t$ channel matrix, and $\mathbf{n}$ is the $N_r \times 1$ additive white Gaussian noise (AWGN) vector at a given instant in time. All of the coefficients h_{ij} comprise the channel matrix $\mathbf{H}$. The coefficient h_{ij} represents the complex gain of the channel between the jth transmit antenna and the ith receive antenna.

The $N_r \times N_t$ channel matrix $\mathbf{H}$ can be written as

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1N_t} \\ h_{21} & h_{22} & \cdots & h_{2N_t} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_r 1} & h_{N_r 2} & \cdots & h_{N_r N_t} \end{bmatrix} \qquad (2.2)$$

where

$$\begin{aligned} h_{ij} &= \alpha_{ij} + j\beta_{ij} \\ &= \sqrt{\alpha_{ij}^2 + \beta_{ij}^2} \cdot e^{-j\tan^{-1}\frac{\beta_{ij}}{\alpha_{ij}}} \\ &= \left| h_{ij} \right| \cdot e^{j\phi_{ij}} \end{aligned} \qquad (2.3)$$

The coefficients $\{h_{ij}\}$ stand for unknown transmission properties of the medium, usually Rayleigh distributed in a rich scattering environment with no line-of-sight (LOS). If α and β are independent and Gaussian distributed random variables, then $\left| h_{ij} \right|$ is a Rayleigh distributed random variable. In fact, the coefficients $\{h_{ij}\}$ are likely to be subject to varying degrees of fading and change over time. Therefore, the determination of the channel matrix $\mathbf{H}$ is an important and necessary aspect of MIMO processing. If all these coefficients are known, there will be sufficient information for the receiver to eliminate interference from other transmitters operating at the same frequency band. At the beginning, the channel matrix is estimated using a well-designed preamble training sequences sent ahead of the payload. Then the channel matrix is refined dynamically using pilot tones that are sent in conjunction with the payload. Moreover, when comparing the systems equipped with different N_t and N_r antennas, the channel matrix $\mathbf{H}$ has to be normalized. The channel matrix is normalized such that $\left\| \mathbf{H} \right\|_F^2 = N_t \cdot N_r$, where $\left\| \cdot \cdot \right\|_F^2$ represents the Frobenius norm. This normalization step keeps the spatial characteristics but removes the influence of the time and frequency variation. There are some MIMO channel estimation methods proposed, which can be found in [10], [24].

2.2　Channel Capacity

A measure of how much information that can be transmitted and received with a negligible probability of error is called the channel capacity. To determine this measure of channel potential, a channel encoder receiving a source symbol every T_s second is assumed. If S represents the set of all source symbols and the entropy rate of the source is written as $H(s)$, the channel encoder will receive $H(s)/T_s$ information bits per second on average. A channel codeword leaving the channel encoder every T_c second is also assumed. In order to be able to transmit all the information from the source, there must be R information bits per channel symbol.

$$R = \frac{H(s)T_c}{T_s} \qquad (2.4)$$

The number R is called the information rate of the channel encoder. The maximum information rate that can be used causing negligible probability of errors at the output is called the capacity of the channel. By transmitting information with the rate R, the channel is used every T_c seconds. The channel capacity is then measured in bits per channel use. Assuming that the channel has bandwidth W, the input and output can be represented by samples taken $T_s = 1/2W$ seconds apart. With a band-limited channel, the capacity is measured in information bits per second. It is common to represent the channel capacity within a unit bandwidth of the channel, which means that the channel capacity is measured in bits/sec/Hz.

It is desirable to design transmission schemes that exploit the channel capacity as much as possible. Representing the input and output of a memoryless wireless channel with the random variables X and Y respectively, the channel capacity is defined as

$$C = \max_{p(x)} I(X;Y) \quad \text{bits/sec/Hz} \qquad (2.5)$$

where $I(X;Y)$ represents the mutual information between X and Y. Equation (2.5) states that the mutual information is maximized when considering all possible transmitter statistical distributions $p(x)$. Mutual information is a measure of the amount of information that one random variable contains about another one. The mutual information between X and Y can also be written as

$$I(X;Y) = H(Y) - H(Y|X) \tag{2.6}$$

where $H(Y|X)$ represents the conditional entropy between the random variables X and Y. The entropy of a random variable can be described as a measure of the uncertainty of the random variable. It can also be described as a measure of the amount of information required on average to describe the random variable. Due to Equation (2.6), mutual information can be described as the reduction in the uncertainty of one random variable due to the knowledge of the other. Note that the mutual information between X and Y depends on the properties of X (through the probability distribution of X) and the properties of channel (through a channel matrix **H**). In the following, four different kinds of channel capacities are introduced (single-input single-output (SISO), single-input multiple-output (SIMO), multiple-input single-output (MISO), and MIMO) to get the further concepts about the properties of the channel capacity.

2.2.1 SISO Channel Capacity

The ergodic (mean) capacity of a random channel ($N_t = N_r = 1$) with the average transmit power constraint P_T can be expressed as

$$C = E_H \left\{ \max_{p(x):P \le P_T} I(X;Y) \right\} \text{ bits/sec/Hz} \tag{2.7}$$

where E_H denotes the expectation over all channel realizations and P is the average power of a single channel codeword transmitted over the channel. Compared to the definition in Equation (2.5), the capacity of the channel is now defined as the maximum of the mutual information between the input and output over all statistical distributions on the input that satisfy the power constraint. If each channel symbol at the transmitter is denoted by s, the average power constraint can be expressed as

$$P = E\left[|s|^2 \right] \le P_T \tag{2.8}$$

Using Equation (2.7), the ergodic (mean) capacity of a SISO system ($N_t = N_r = 1$) with a random complex channel gain h_{11} is given by

$$C = E_H \left\{ \log_2 \left(1 + \rho \cdot |h_{11}|^2 \right) \right\} \text{ bits/sec/Hz} \tag{2.9}$$

where ρ is the average signal-to-noise (SNR) ratio at the receive branch. If $|h_{11}|$ is Rayleigh, $|h_{11}|^2$ follows a chi-squared distribution with two degrees of freedom. Equation (2.7) can then be written as

$$C = E_H \left\{ \log_2 \left(1 + \rho \cdot \chi_2^2 \right) \right\} \quad \text{bits/sec/Hz} \tag{2.10}$$

where χ_2^2 is a chi-square distributed random variable with two degrees of freedom.

2.2.2 SIMO and MISO Channel Capacity

As more antennas deployed at the receiving end, the statistics of capacity improve. Then the ergodic (mean) capacity of a SIMO system with N_r receive antennas is given by

$$C = \log_2 (1 + \rho \sum_{i=1}^{N_r} |h_{i1}|^2) \quad \text{bits/sec/Hz} \tag{2.11}$$

where h_{i1} represents the gain for the receive antenna i. Note that the crucial feature in Equation (2.11) is that increasing the number of receive antennas N_r only results in a logarithmic increase in the ergodic (mean) capacity. Similarly, if transmit diversity is opted, in the common case, where the transmitter doesn't have the channel knowledge, the ergodic (mean) capacity of a MISO system with N_t transmit antennas is given by

$$C = \log_2 (1 + \frac{\rho}{N_t} \sum_{i=1}^{N_t} |h_{1i}|^2) \quad \text{bits/sec/Hz} \tag{2.12}$$

where the normalization by N_t ensures a fixed total transmit power and shows the absence of array gain in that case (compared to the case in Equation (2.11), where the channel energy can be combined coherently). Again, the capacity has a logarithmic relationship with the number of transmit antennas N_t.

2.2.3 MIMO Channel Capacity

The capacity of a random MIMO channel with the power constraint P_T can be expressed as

$$C = E_H \left\{ \max_{p(\mathbf{x}):tr(\Phi) \le P_T} I(\mathbf{x};\mathbf{y}) \right\} \quad \text{bits/sec/Hz} \tag{2.13}$$

where $\Phi = E\{xx^H\}^2$ is the covariance matrix of the transmit signal vector $\mathbf{x}$. By the relationship between mutual information and entropy and using Equation (2.1), Equation (2.13) can be expanded as follows for a given channel matrix $\mathbf{H}$.

$$
\begin{aligned}
I(\mathbf{x};\mathbf{y}) &= h(\mathbf{y}) - h(\mathbf{y}\,|\,\mathbf{x}) \\
&= h(\mathbf{y}) - h(\mathbf{Hx} + \mathbf{n}\,|\,\mathbf{x}) \\
&= h(\mathbf{y}) - h(\mathbf{n}\,|\,\mathbf{x}) \\
&= h(\mathbf{y}) - h(\mathbf{n})
\end{aligned} \tag{2.14}
$$

where $h(\cdot)$ denotes the differential entropy of a continuous random variable. It is assumed that the transmit vector $\mathbf{x}$ and the noise vector $\mathbf{n}$ are independent.

When $\mathbf{y}$ is Gaussian, Equation (2.14) is maximized. Since the normal distribution maximizes the entropy for a given variance. The differential entropy of a complex Gaussian vector $\mathbf{y} \in \mathbf{C}^n$, the differential entropy is less than or equal to $\log_2 \det(\pi e \mathbf{K})$, with equality if and only if $\mathbf{y}$ is a circularly symmetric complex Gaussian with $E\{yy^H\} = \mathbf{K}$. For a real Gaussian vector $\mathbf{y} \in \mathbf{R}^n$ with zero mean and covariance matrix, $\mathbf{K}$ is equal to $\log_2 \left((2\pi e)^n \det \mathbf{K} \right) / 2$. Assuming the optimal Gaussian distribution for the transmit vector $\mathbf{x}$, the covariance matrix of the received complex vector $\mathbf{y}$ is given by

$$
\begin{aligned}
E\{\mathbf{yy}^H\} &= E\left\{ (\mathbf{Hx} + \mathbf{n})(\mathbf{Hx} + \mathbf{n})^H \right\} \\
&= E\left\{ \mathbf{Hxx}^H\mathbf{H}^H \right\} + E\left\{ \mathbf{nn}^H \right\} \\
&= \mathbf{H\Phi H}^H + \mathbf{K}^n \\
&= \mathbf{K}^d + \mathbf{K}^n
\end{aligned} \tag{2.15}
$$

The desired part and the noise part of Equation (2.15) denotes respectively by the superscript d and n. The maximum mutual information of a random MIMO channel is then given by

$$I = h(\mathbf{y}) - h(\mathbf{n})$$

$$= \log_2 \left[\det\left(\pi e \left(\mathbf{K}^d + \mathbf{K}^n\right)\right)\right] - \log_2 \left[\det\left(\pi e \mathbf{K}^n\right)\right]$$

$$= \log_2 \left[\det\left(\mathbf{K}^d + \mathbf{K}^n\right)\right] - \log_2 \left[\det\left(\mathbf{K}^n\right)\right]$$

$$= \log_2 \left[\det\left(\left(\mathbf{K}^d + \mathbf{K}^n\right)\left(\mathbf{K}^n\right)^{-1}\right)\right] \qquad (2.16)$$

$$= \log_2 \left[\det\left(\mathbf{K}^d \left(\mathbf{K}^n\right)^{-1} + \mathbf{I}_M\right)\right]$$

$$= \log_2 \left[\det\left(\mathbf{H}\boldsymbol{\Phi}\mathbf{H}^H \left(\mathbf{K}^n\right)^{-1} + \mathbf{I}_M\right)\right]$$

When the transmitter has no knowledge about the channel, it is optimal to use a uniform power distribution. The transmit covariance matrix is then given by $\boldsymbol{\Phi} = P_T \mathbf{I}_{N_t} / N_t$. It is also common to assume uncorrelated noise in each receive antenna described by the covariance matrix $\mathbf{K}^n = \sigma^2 \mathbf{I}_{N_r}$. The ergodic (mean) capacity for a complex additive white Gaussian noise (AWGN) MIMO channel can then be expressed as

$$C = E_H \left\{ \log_2 \left[\det\left(\mathbf{I}_{N_r} + \frac{P_T}{\sigma^2 N_t} \mathbf{H}\mathbf{H}^H \right)\right]\right\} \quad \text{bits/sec/Hz} \qquad (2.17)$$

Equation (2.17) can also be written as

$$C = E_H \left\{ \log_2 \left[\det\left(\mathbf{I}_{N_r} + \frac{\rho}{N_t} \mathbf{H}\mathbf{H}^H \right)\right]\right\} \quad \text{bits/sec/Hz} \qquad (2.18)$$

where $\rho = P_T / \sigma^2$ is the average signal-to-noise (SNR) at each receive antenna. By the law of large numbers, the term $\mathbf{H}\mathbf{H}^H / N_t \rightarrow \mathbf{I}_{N_r}$ as N_r is fixed and N_t gets large. Hence the capacity in the limit of large transmit antennas N_t can be written as

$$C = E_H \left\{ N_r \cdot \log_2 \left(1 + \rho\right)\right\} \quad \text{bits/sec/Hz} \qquad (2.19)$$

Further analysis of the MIMO channel capacity given in Equation (2.18) is possible by diagonalizing the product matrix $\mathbf{H}\mathbf{H}^H$ either by eigenvalue decomposition or singular value decomposition (SVD). By using SVD, the matrix product is written as

$$\mathbf{H} = \mathbf{U\Sigma V}^H \tag{2.20}$$

where **U** and **V** are unitary matrices of left and right singular vectors respectively, and Σ is a triangular matrix with singular values on the main diagonal. All elements on the diagonal are zero except for the first k elements. The number of non-zero singular values k of Σ equals the rank of the channel matrix. Substituting Equation (2.20) into Equation (2.18), the MIMO channel capacity can be written as

$$C = E_H \left\{ \log_2 \left[\det \left(I_{N_r} + \frac{\rho}{N_t} \mathbf{U\Sigma\Sigma}^H \mathbf{U}^H \right) \right] \right\} \quad \text{bits/sec/Hz} \tag{2.21}$$

The matrix product $\mathbf{HH}^H$ can also be described by using eigenvalue decomposition on the channel matrix **H** written as

$$\mathbf{HH}^H = \mathbf{E\Lambda E}^H \tag{2.22}$$

where **E** is the eigenvector matrix with orthonormal columns and $\mathbf{\Lambda}$ is a diagonal matrix with the eigenvalues on the main diagonal. Using this notation, Equation (2.18) can be written as

$$C = E_H \left\{ \log_2 \left[\det \left(\mathbf{I}_{N_r} + \frac{\rho}{N_t} \mathbf{E\Lambda E}^H \right) \right] \right\} \quad \text{bits/sec/Hz} \tag{2.23}$$

After diagonalizing the product matrix $\mathbf{HH}^H$, the capacity formulas of the MIMO channel now includes unitary and diagonal matrices only. It is then easier to see that the total capacity of a MIMO channel is made up by the sum of parallel AWGN SISO subchannels. The number of parallel subchannels is determined by the rank of the channel matrix. In general, the rank of the channel matrix is given by

$$rank(\mathbf{H}) = k \le \min(N_t, N_r) \tag{2.24}$$

Using the fact that the determinant of a unitary matrix is equal to 1 and Equation (2.24), Equations (2.21) and (2.23) can be expressed respectively as

$$C = E_H \left\{ \sum_{i=1}^{k} \log_2 \left(1 + \frac{\rho}{N_t} \sigma_i^2 \right) \right\} \quad \text{bits/sec/Hz} \tag{2.25}$$

$$= E_H \left\{ \sum_{i=1}^{k} \log_2 \left(1 + \frac{\rho}{N_t} \lambda_i \right) \right\} \quad \text{bits/sec/Hz} \tag{2.26}$$

where σ_i^2 are the squared singular values of the diagonal matrix Σ and λ_i are the eigenvalues of the diagonal matrix Λ. The maximum capacity of a MIMO channel is achieved in the unrealistic situation when each of the N_t transmitted signals is received by the same set of N_r antennas without interference. It can also be described as if each transmitted signal is received by a separate set of receive antennas, giving a total number of $N_t \cdot N_r$ receive antennas.

With optimal combining at the receiver and receive diversity only ($N_r = 1$), the channel capacity can be expressed as

$$C = E_H \left\{ \log_2 \left(1 + \rho \cdot \chi_{2N_r}^2 \right) \right\} \quad \text{bits/sec/Hz} \tag{2.27}$$

where $\chi_{2N_r}^2$ is a chi-distributed random variable with $2N_r$ degrees of freedom. If there are N_t transmit antennas and optimal combining between N_r antennas at the receiver, the capacity can be written as

$$C = E_H \left\{ N_t \cdot \log_2 \left(1 + \frac{\rho}{N_t} \cdot \chi_{2N_r}^2 \right) \right\} \quad \text{bits/sec/Hz} \tag{2.28}$$

Equation (2.28) represents the upper bound of a Rayleigh fading MIMO channel.

When the channel is known at the transmitter, the maximum capacity of a MIMO channel can be achieved by using the water-filling (WF) principle on the transmit covariance matrix. The capacity is then given by

$$\begin{aligned} C &= E_H \left\{ \sum_{i=1}^{k} \log_2 \left(1 + \varepsilon_i \frac{\rho}{N_t} \lambda_i \right) \right\} \\ &= E_H \left\{ \sum_{i=1}^{k} \log_2 \left(1 + \varepsilon_i \frac{\rho}{N_t} \sigma_i^2 \right) \right\} \\ &= \sum_{i=1}^{k} \log_2 (u \lambda_i)^+ \quad \text{bits/sec/Hz} \end{aligned} \tag{2.29}$$

where "+" denotes taking only those terms which are positive and u is a scalar, representing the portion of the available transmit power going into the ith subchannel which is chosen to satisfy

$$\rho = \sum_{i=1}^{k}(u - \lambda_i^{-1})^+ \tag{2.30}$$

Since u is a complicated nonlinear function of $\lambda_1, \lambda_2, \ldots, \lambda_k$, the distribution of the channel capacity appears intractable, even in the Wishart case when the joint distribution of $\lambda_1, \lambda_2, \ldots, \lambda_k$ is known. Nevertheless, the channel capacity can be simulated using Equations (2.29) and (2.30) for any given $\mathbf{HH}^H$ so that the optimal capacity can be computed numerically for any channel [12].

2.3   Spatial Multiplexing

The use of multiple antennas at both ends of a wireless link has recently been shown to have the potential of achieving extraordinary data rate. The corresponding technology is known as spatial multiplexing [1]-[3], [11]. It allows a data rate enhancement in a wireless radio link without additional power or bandwidth consumption. In spatial multiplexing systems, different data streams are transmitted from different transmit antennas simultaneously or sequentially and these data streams are separated and demultiplexed to yield the original transmitted signals according to their unique spatial signatures at the receiver. An illustration of the spatial multiplexing system is shown in Fig. 2.2. The separation step is made possible by the fact that the rich scattering multiplath contributes to lower correlation between MIMO channel coefficients, and creates a desirable full rank and low condition number coefficient matrix condition to resolve N_t unknowns from a linear system of N_r equations. In the following, two typical spatial multiplexing schemes, D-BLAST [1] and V-BLAST [3], [11], are introduced.

2.3.1 Diagonal Bell Lab's Layered Space-Time (D-BLAST)

Space-time coding (STC) performs channel coding across space and time to exploit the spatial diversity offered by MIMO systems to increase system capacity. However, the decoding complexity of the space-time codes is exponentially increased with the number of transmit antennas, which makes it hard to implement real-time decoding as the number of antennas grows. To reduce the complexity of space-time based MIMO systems, D-BLAST architecture has been proposed in [1]. Rather than try to achieve optimal channel coding scheme, in D-BLAST architecture, the input data stream is divided into several substreams. Each substream is encoded independently using an elegant diagonally-layered coding structure in which code blocks are dispersed across diagonals in space-time and the association of corresponding output stream with transmit antenna is periodically cycled to explore spatial diversity. To decode each layer, channel parameters are used to cancel interference from undetected signals to make the desired signal as "clean" as possible.

Fig. 2.3 shows the typical encoding steps in D-BLAST. Considering a system with N_t transmit and N_r receive antennas, the high rate information data stream is first demultiplexed into N_t subsequences. Each subsequence is encoded by a conventional 1-D constituent code with low decoding complexity. The encoders apply these coded symbols to generate a semi-infinite matrix C of N_t rows to be transmitted. The element in the pth row and tth column of C, c_t^p, is transmitted by the pth transmit antenna at time t. As illustrated in Fig. 2.2, $c_1^1, c_2^2, c_3^3, c_4^1, c_5^2, c_6^3$ are encoded by encoder α, $c_2^1, c_3^2, c_4^3, c_5^1, c_6^2, c_7^3$ are encoded by encoder β, and $c_3^1, c_4^2, c_5^3, c_6^1, c_7^2, c_8^3$ are encoded by encoder γ.

Fig. 2.4 shows the typical decoding steps of interference suppression, symbol detection, decoding, and interference cancellation performed in D-BLAST. The receiver generates decisions for the first diagonal of C. Based on these decisions, the diagonal is decoded and fed back to remove the contribution of this diagonal from the received data. The receiver continues to decode the next diagonal and so on. The encoded substreams share a balanced presence over all paths to the receiver, so none of
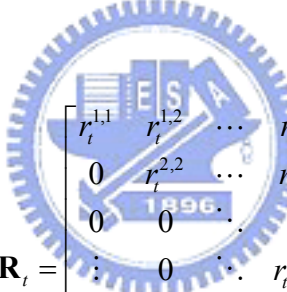
the individual substreams is subject to the worst path. Therefore, the data received at time t by the qth receive antenna is r_t^q, which contains a superposition of c_t^p, $p = 1, 2, \ldots, N_t$, and an AWGN noise component. Then, the received data vector can be expressed as $\mathbf{r}_t = \mathbf{H}_t \mathbf{c}_t + \xi_t$ at any time instance t. The D-BLAST method uses a repeated process of interference suppression, symbol detection, and interference cancellation to decode all symbols, $c_t^{N_t}, c_t^{N_t-1}, \ldots, c_t^1$. This decoding process can be expressed in a general form described in the following.

Let $\mathbf{Q}_t \mathbf{R}_t$ be the QR decomposition of $\mathbf{H}_t$, where $\mathbf{Q}_t$ is an $N_r \times N_r$ unitary matrix and $\mathbf{R}_t$ is an $N_r \times N_t$ upper triangular matrix. Multiplying the received signal by $\mathbf{Q}_t^H$, we can get

$$\mathbf{y}_t = \mathbf{Q}_t^H \mathbf{r}_t = \mathbf{Q}_t^H \mathbf{H}_t \mathbf{c}_t + \mathbf{Q}_t^H \xi_t = \underbrace{\mathbf{Q}_t^H \mathbf{Q}_t}_{\mathbf{I}_{N_r}} \mathbf{R}_t \mathbf{c}_t + \underbrace{\mathbf{Q}_t^H \xi_t}_{\tilde{\xi}_t} \tag{2.31}$$

$$= \mathbf{R}_t \mathbf{c}_t + \tilde{\xi}_t$$

where

$$\mathbf{y}_t = \begin{bmatrix} y_t^1 \\ y_t^2 \\ \vdots \\ y_t^{N_r} \end{bmatrix}, \quad \mathbf{R}_t = \begin{bmatrix} r_t^{1,1} & r_t^{1,2} & \cdots & r_t^{1,N_t} \\ 0 & r_t^{2,2} & \cdots & r_t^{2,N_t} \\ 0 & 0 & \ddots & \vdots \\ \vdots & 0 & \ddots & r_t^{N_t,N_t} \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \quad \tilde{\xi}_t = \begin{bmatrix} \tilde{\xi}_t^1 \\ \tilde{\xi}_t^2 \\ \vdots \\ \tilde{\xi}_t^{N_r} \end{bmatrix} \tag{2.32}$$

Since $\mathbf{R}_t$ is a upper triangular matrix, the elements in the vector $\mathbf{y}_t$ can be expressed as

$$y_t^p = r_t^{p,p} c_t^p + \tilde{\xi}_t^p + \left\{ \text{contribution from } c_t^{p+1}, c_t^{p+2}, \ldots, c_t^{N_t} \right\} \tag{2.33}$$

Hence, the interference from c_t^q, $q < p \le N_t$, is first suppressed in y_t^k and the residual interference terms in Equation (2.33) can be cancelled by the available decisions $\hat{c}_\tau^{p+1}, \hat{c}_\tau^{p+2}, \ldots, \hat{c}_\tau^{N_t}$. Assuming all these decisions are correct, then the present decision variable is

$$\tilde{c}_t^p = r_t^{p,p} c_t^p + \tilde{\xi}_t^p, \quad p = 1, 2, \ldots, N_t \tag{2.34}$$

The relation between c^p and $\tilde{c}^p$ in Equation (2.34) can be interpreted as the input and output of a SISO channel with the channel power gain $\left| r^{p,p} \right|^2$ and AWGN. The channel power gain $\left| r^{p,p} \right|^2$ is independently chi-squared distributed with $2 \times (N_r - p + 1)$ degrees of freedom. Moreover, if there are no decision feedback errors, the pth row of the C matrix can be treated as transmitted over a $(N_t, N_r) = (1, N_r - p + 1)$ system without interference from the other rows and all fades are i.i.d.

2.3.2 Vertical Bell Lab's Layered Space-Time (V-BLAST)

The D-BLAST algorithm has been proposed by Foschini for achieving a substantial part of the MIMO capacity. In an independent Rayleigh scattering environment, this processing structure leads to theoretical rates which grow linearly with the number of antennas (assuming equal number of transmit and receive antennas) with these rates approaching ninety percents of Shannon capacity. However, the diagonal approach suffers from certain implementation complexities which make it inappropriate for practical implementation. Therefore, a simplified version of the BLAST algorithm is known as V-BLAST (vertical BLAST) [3], [11]. It is capable of achieving high spectral efficiency while being relatively simple to implement. The essential difference between D-BLAST and V-BLAST lies in the vector encoding process. In D-BLAST, redundancy between the substreams is introduced through the use of specialized intersubstream block coding. In V-BLAST, however, the vector encoding process is simply a demultiplex operation followed by independent bit-to-symbol mapping of each substream. No intersubstream coding, or coding of any kind, is required, though conventional coding of the individual substreams will certainly be applied.

Fig. 2.5 shows the typical encoding steps in V-BLAST. The coding procedure can be viewed as there is an encoder on each transmit antenna. The output coded symbols of one encoder are transmitted from the corresponding transmit antenna. The output

coded symbol of the *p*th encoder is used to fill the *p*th row of C.

Fig. 2.6 shows the typical decoding steps in V-BLAST. The detection procedure is to extract the strongest substream from the signals received by the all receive antennas simultaneously. Then, the procedure proceeds with the remaining weaker signals, which are easier to recover when the strongest signals have been removed as a source of interference. Following the data model in D-BLAST, let $\tilde{\mathbf{H}}^{l=1} = \mathbf{H}_t$ and $\tilde{\mathbf{r}}^{l=1} = \mathbf{r}_t$ at the first decoding step at a given time instant t. In each step l, the pseudo-inverse of $\tilde{\mathbf{H}}^l$ is calculated to be the nulling matrix $\mathbf{G}^l$.

$$
\begin{aligned}
\mathbf{G}^l &= (\tilde{\mathbf{H}}^l)^+ \\
&= \left((\tilde{\mathbf{H}}^l)^H \tilde{\mathbf{H}}^l \right)^{-1} (\tilde{\mathbf{H}}^l)^H
\end{aligned}
\tag{2.35}
$$

Each row of $\mathbf{G}^l$ can be used to null all but the *l*th desired signal. The layer shows the biggest post-processing SNR suggested to be detected first to reduce the error propagation effect efficiently [3]. At this step, the row of $\mathbf{G}^l$ with the minimum norm is chosen and the corresponding row is defined as the nulling vector $\mathbf{w}_{k_l}^T$.

$$
k_l = \underset{j \notin \{k_1, \dots, k_{l-1}\}}{\arg\min} \| (\mathbf{G}^l)_j \|^2
\tag{2.36}
$$

$$
\mathbf{w}_{k_l} = (\mathbf{G}^l)_{k_l}^T
\tag{2.37}
$$

The post-processing SNR for the k_lth detected component of **c** can be defined as

$$
\rho_{k_l} = \frac{< | c^{k_l} |^2 >}{\sigma^2 \| \mathbf{w}_{k_l} \|^2}
\tag{2.38}
$$

Then, using $\mathbf{w}_{k_l}$ to suppresses all layers but the one transmitted from antenna k_l and a soft decision value is obtained

$$
\overline{c}_t^{k_l} = \mathbf{w}_{k_l}^T \tilde{\mathbf{r}}_t^l
\tag{2.39}
$$

Therefore, the k_l th layer can be detected within the constellation set S

$$
\hat{c}_t^{k_l} = \underset{\tilde{x} \in S}{\arg\min} \| \tilde{c} - \overline{c}_t^{k_l} \|^2
\tag{2.40}
$$

As soon as one layer is detected, the part of the detected signal can be subtracted from the received vector to improve the detection performance for the later layers.

$$\tilde{\mathbf{r}}_t^{l+1} = \tilde{\mathbf{r}}_t^l - \hat{c}_t^{k_l}\left(\tilde{\mathbf{H}}^l\right)^{k_l} \tag{2.41}$$

where $\left(\tilde{\mathbf{H}}^l\right)^{k_l}$ denotes the k_lth column of $\tilde{\mathbf{H}}^l$. Then, the channel matrix is deflated to account for its removal.

$$\tilde{\mathbf{H}}^{l+1} = \left(\tilde{\mathbf{H}}^l\right)^{\overline{k_l}} \tag{2.42}$$

where the notation $\left(\tilde{\mathbf{H}}^l\right)^{\overline{k_l}}$ denotes the matrix obtained by zeroing columns $k_1, k_2, \ldots, k_l$ of $\tilde{\mathbf{H}}^l$. Therefore, the diversity gain is increased by one at each step when we decrease the number of layers to be nulled out in the next step by one.

The Zero-Forcing (ZF) V-BLAST detection algorithm can be summarized as follows:

Initialization:

$$i \leftarrow 1$$
$$\mathbf{G}^1 = \left(\tilde{\mathbf{H}}^1\right)^+$$
$$k_1 = \arg\min_j \| (\mathbf{G}^1)_j \|^2$$
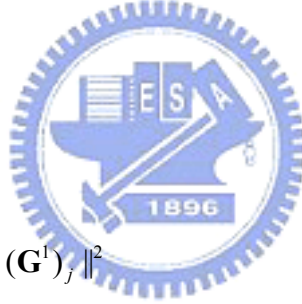
Recursion:

$$\mathbf{w}_{k_l} = (\mathbf{G}^l)_{k_l}^T$$
$$\overline{c}^{k_l} = \mathbf{w}_{k_l}^T \tilde{\mathbf{r}}^l$$
$$\overline{c}^{k_l} = \mathbf{w}_{k_l}^T \tilde{\mathbf{r}}^l$$
$$\hat{c}^{k_l} = Q(\overline{c}^{k_l})$$
$$\mathbf{r}^{l+1} = \mathbf{r}^l - \hat{c}^{k_l}(\tilde{\mathbf{H}}^l)^{k_l}$$
$$\tilde{\mathbf{H}}^{l+1} = \left(\tilde{\mathbf{H}}^l\right)^{\overline{k_l}}$$
$$\mathbf{G}^{l+1} = \left(\tilde{\mathbf{H}}^{l+1}\right)^+$$
$$k_{l+1} = \arg\min_{j \notin \{k_1, \cdots k_l\}} \| (\mathbf{G}^{l+1})_j \|^2$$
$$i \leftarrow i+1$$

By using the minimum mean square error (MMSE) nulling matrix instead of the

ZF one, it can improve the detection performance especially for the mid-range SNR values [3].

$$\mathbf{G}^l = \left((\tilde{\mathbf{H}}^l)^H \tilde{\mathbf{H}}^l + \frac{1}{SNR} \mathbf{I} \right)^{-1} (\tilde{\mathbf{H}}^l)^H \tag{2.43}$$

In the MMSE detection case, the noise level on the channel is taken into account besides nulling out the interference. Thus, the SNR has to be estimated at the receiver.

2.4 Computer Simulations

First, the capacity of a SISO channel compared with different kinds of Rayleigh fading channel scenarios such as SIMO, MISO, and MIMO channels are simulated. Second, the ZF V-BLAST BER performances for the ideal detection and cancellation case and with error propagation case are simulated. In simulations, the relationship between SNR and E_b / N_0 can be defined as

$$\text{SNR} = \frac{\text{signal power}}{\text{noise power}} = \frac{\dfrac{E_s}{T_s}}{N_0 B} = \frac{\dfrac{E_b \cdot N_t \cdot M}{T_s}}{N_0 \dfrac{1}{T_s}} = \frac{E_b}{N_0} \cdot (N_t \cdot M) \tag{2.44}$$

When the system transmit power is normalized to one, then the noise power σ^2 corresponding to a specific E_b / N_0 can be generated by

$$\sigma^2 = \frac{N_0}{E_b \cdot N_t \cdot M} \tag{2.45}$$

where E_s is the symbol energy, T_s is the symbol duration, B is the system bandwidth, and M is the modulation order.

In the first simulation, the capacity of a SISO channel compared to the ergodic capacity of Rayleigh fading SIMO channels with different numbers of receive antennas is examined. The results shown in Fig. 2.7 indicate that the channel capacity improves as more receive antennas are deployed. The ergodic channel capacity has a logarithmic increase when increasing the number of antennas at the receiving end.

In the second simulation, the capacity of a SISO channel compared to the ergodic capacity of Rayleigh fading MISO channels with different numbers of transmit

antennas is examined. The results shown in Fig. 2.8 have the same trends as those shown in Fig. 2.7. When more transmit antennas are deployed, the channel capacity also has a logarithmic increasing improvement. However, the ergodic capacity of MISO channels is smaller than that of SIMO channels because the channel energy can not be combined coherently in the MISO channel case. In other words, the array gain is absent in the MISO channel case.

In the third simulation, the capacity of a SISO channel compared to the ergodic capacity of Rayleigh fading MIMO channels with different numbers of transmit and receive antennas is examined. The results shown in Fig. 2.9 indicate that the channel capacity increases linearly with the minimum number of transmit and receive antennas rather than logarithmically. Owing to the use of diversity at both transmitter and receiver, the capacity of MIMO channels is larger than those of different channel scenarios such as SIMO, MISO, and SISO channels.

In the forth simulation, the ZF V-BLAST BER performance with ideal detection and cancellation is examined. The results shown in Fig. 2.10 indicate that the diversity gain increases as the number of effective transmit antennas decreases.

In the fifth simulation, the ZF V-BLAST BER performance with the error propagation is examined. The results shown in Fig. 2.11 indicate that the BER performance will suffer from error propagation and the diversity gain degrades. The error propagation phenomenon will limit the performance because there are no preliminary decisions available to increase the reliability of the symbols detected and used for canceling in the interference cancellation step.

2.5   Summary

Information theory shows that multiple-input multiple-output (MIMO) communication systems can significantly increase the capacity of band-limited wireless channels by a factor of the minimum number of transmit and receive antennas, provided that a rich multipath scattering environment is utilized. In Sections 2.1 and 2.2, MIMO system model and MIMO channel capacity are introduced.

In order to achieve extraordinary data rate in MIMO systems, spatial multiplexing technique is presented in Section 2.3. Spatial multiplexing allows a data rate enhancement in a wireless radio link without additional power or bandwidth consumption. It is realized by transmitting independent data signals from the individual transmit antennas. Two typical spatial multiplexing schemes, D-BLAST and V-BLAST, are introduced in Sections 2.3.1 and 2.3.2.
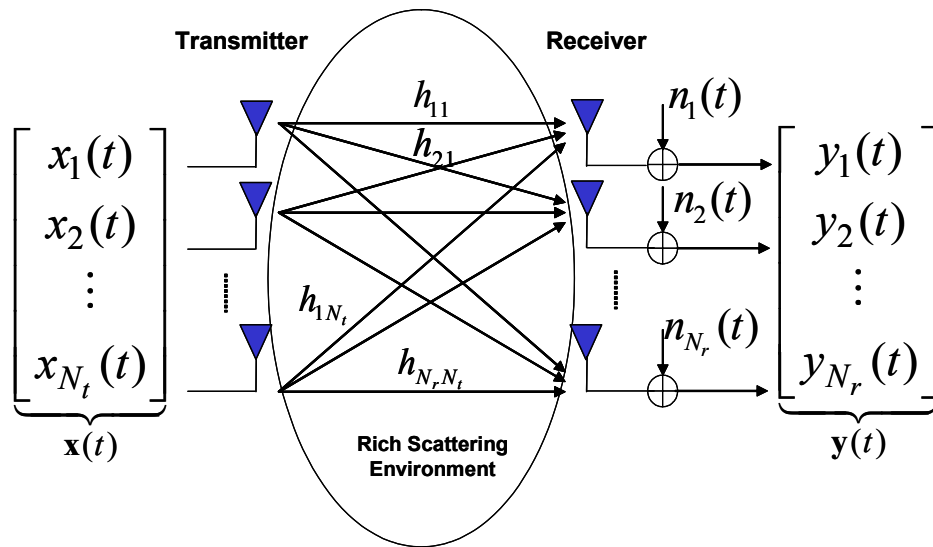
$$\mathbf{y}(t) = \mathbf{H}\mathbf{x}(t) + \mathbf{n}(t)$$

Figure 2.1: MIMO wireless transmission system model.
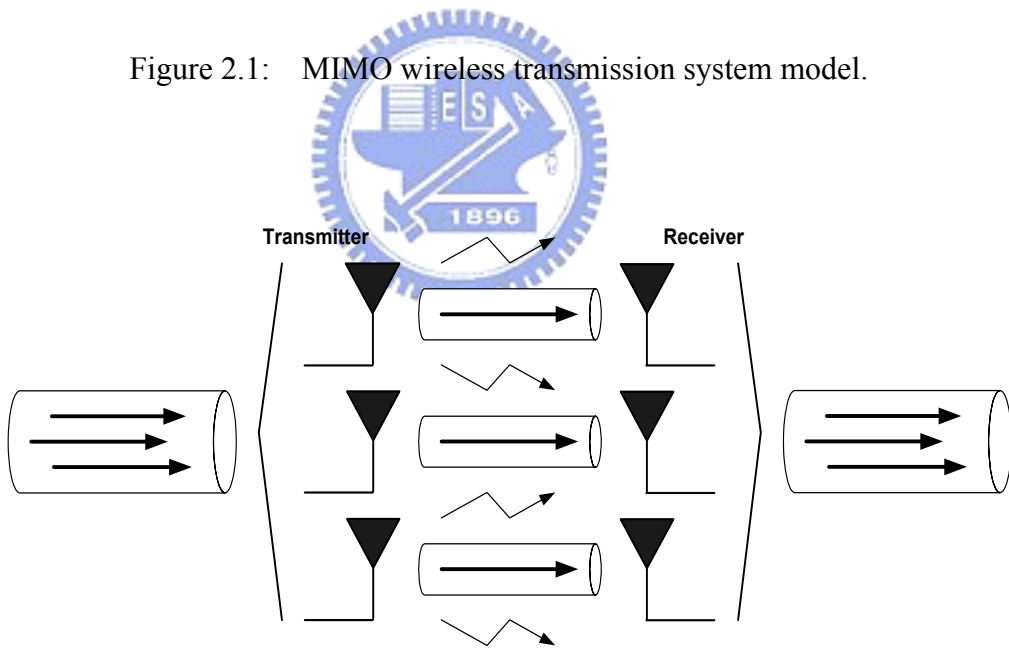
Figure 2.2: An illustration of a spatial multiplexing system.

$$\mathbf{C} = \begin{bmatrix} c_1^1 & c_2^1 & c_3^1 & c_4^1 & c_5^1 & c_6^1 & & & \cdots \\ 0 & c_2^2 & c_3^2 & c_4^2 & c_5^2 & c_6^2 & c_7^2 & & \cdots \\ 0 & 0 & c_3^3 & c_4^3 & c_5^3 & c_6^3 & c_7^3 & c_8^3 & \cdots \end{bmatrix}$$

Data Rate $\propto \mathbf{N}$
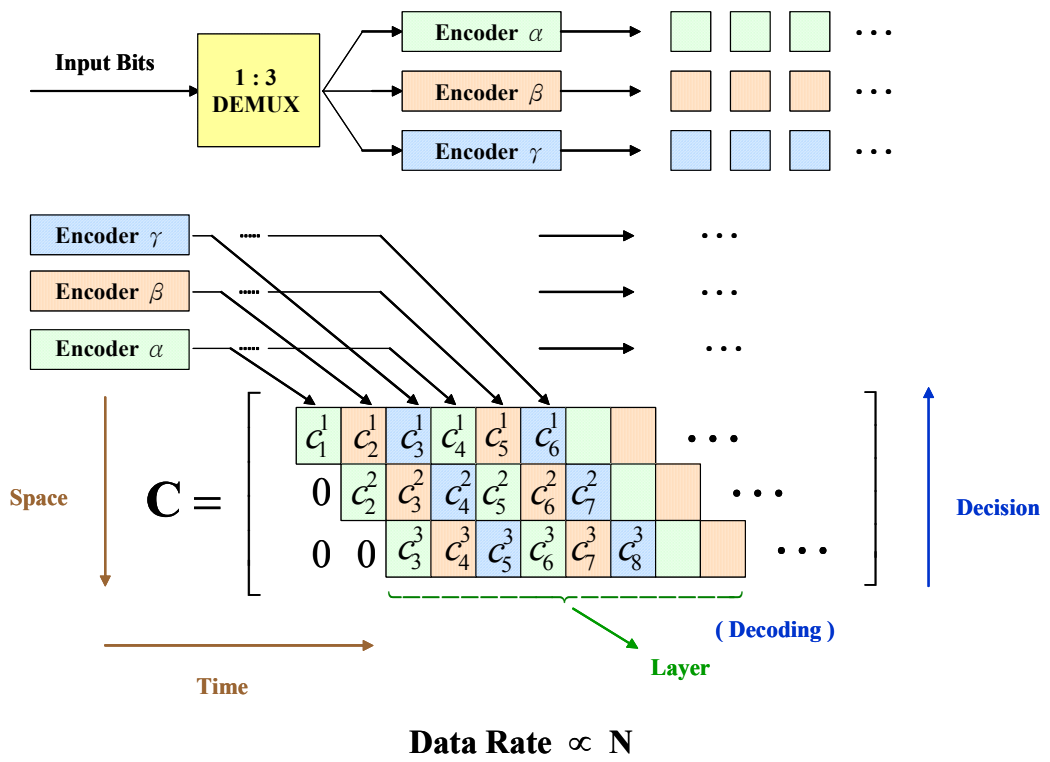
Figure 2.3:   Diagonal Bell Labs' Layered Space-Time encoding procedure.
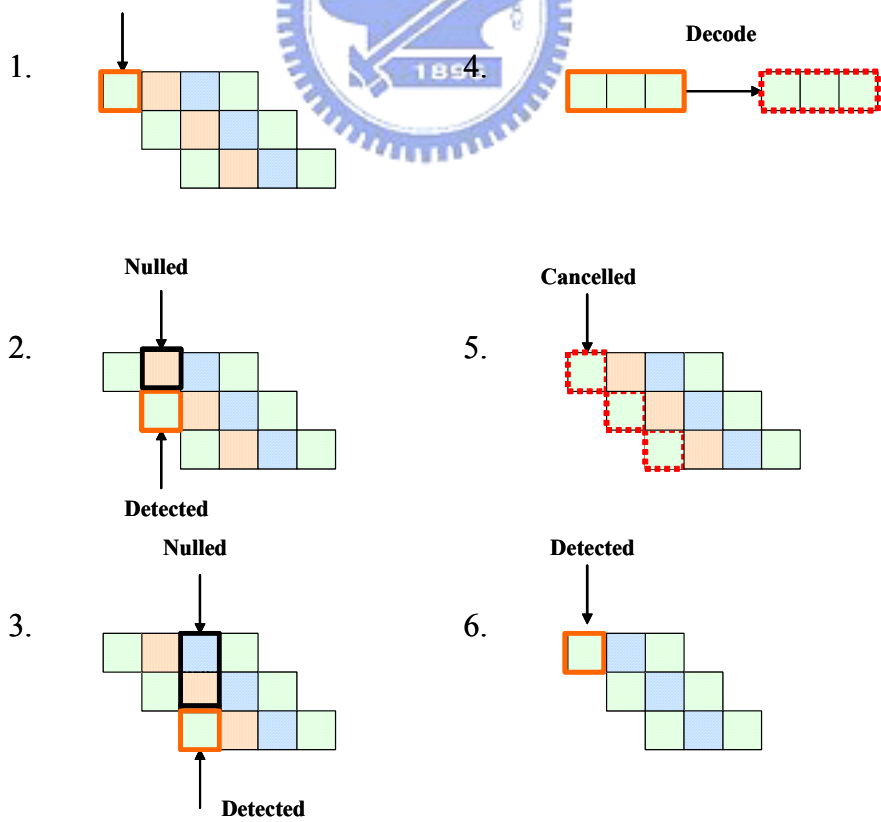
Figure 2.4:   Diagonal Bell Labs' Layered Space-Time decoding procedure.

$$C = \begin{bmatrix} c_1^1 & c_2^1 & c_3^1 & c_4^1 & \cdots \\ c_1^2 & c_2^2 & c_3^2 & c_4^2 & \cdots \\ c_1^3 & c_2^3 & c_3^3 & c_4^3 & \cdots \end{bmatrix}$$

Data Rate $\propto$ N
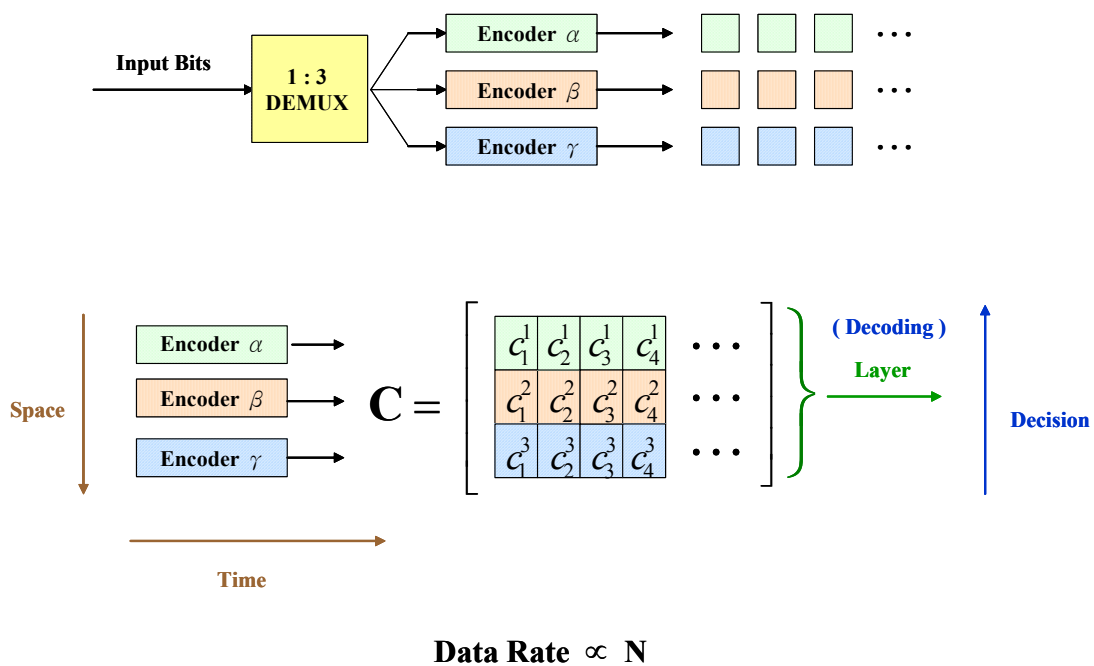
Figure 2.5:   Vertical Bell Labs' Layered Space-Time encoding procedure.
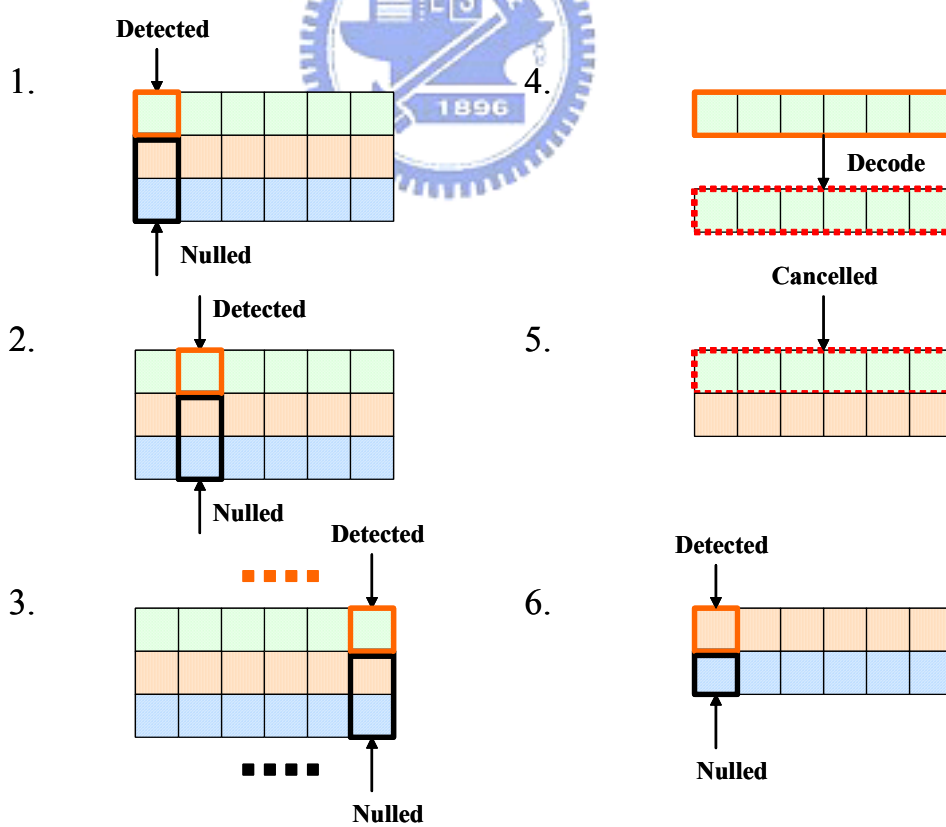
Figure 2.6:   Vertical Bell Labs' Layered Space-Time decoding procedure.
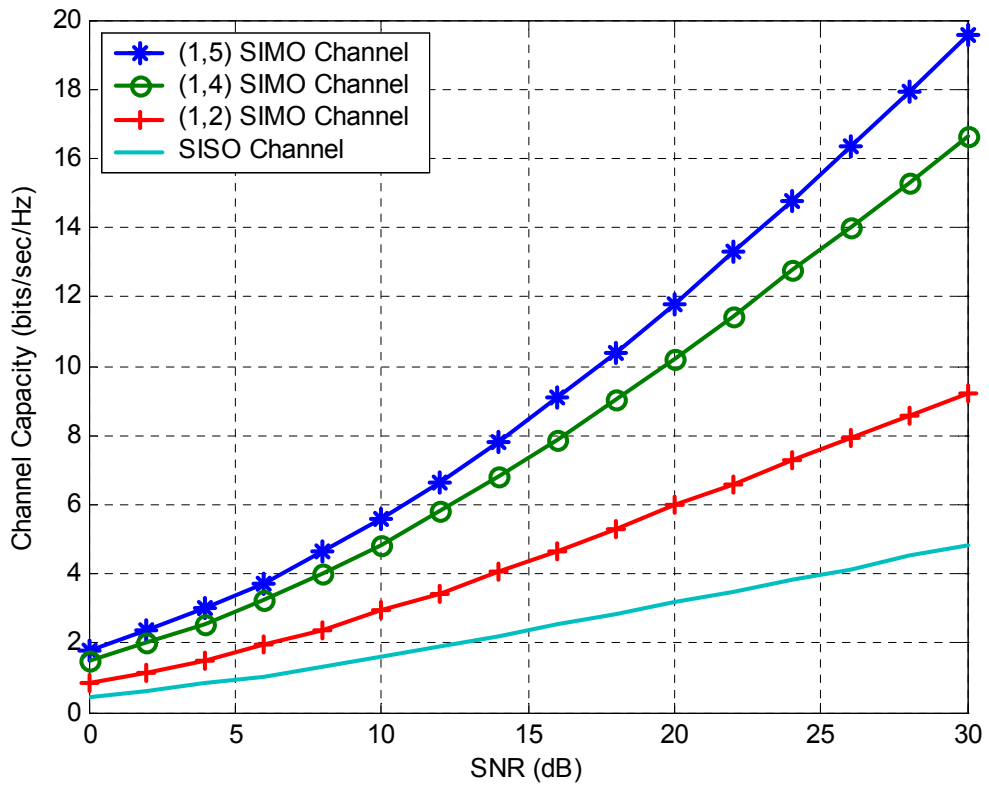
Figure 2.7:   Capacity of a SISO channel compared to the ergodic capacity of Rayleigh
fading SIMO channels with $(N_t, N_r) = (1, 2)$, $(1, 4)$, and $(1, 5)$.
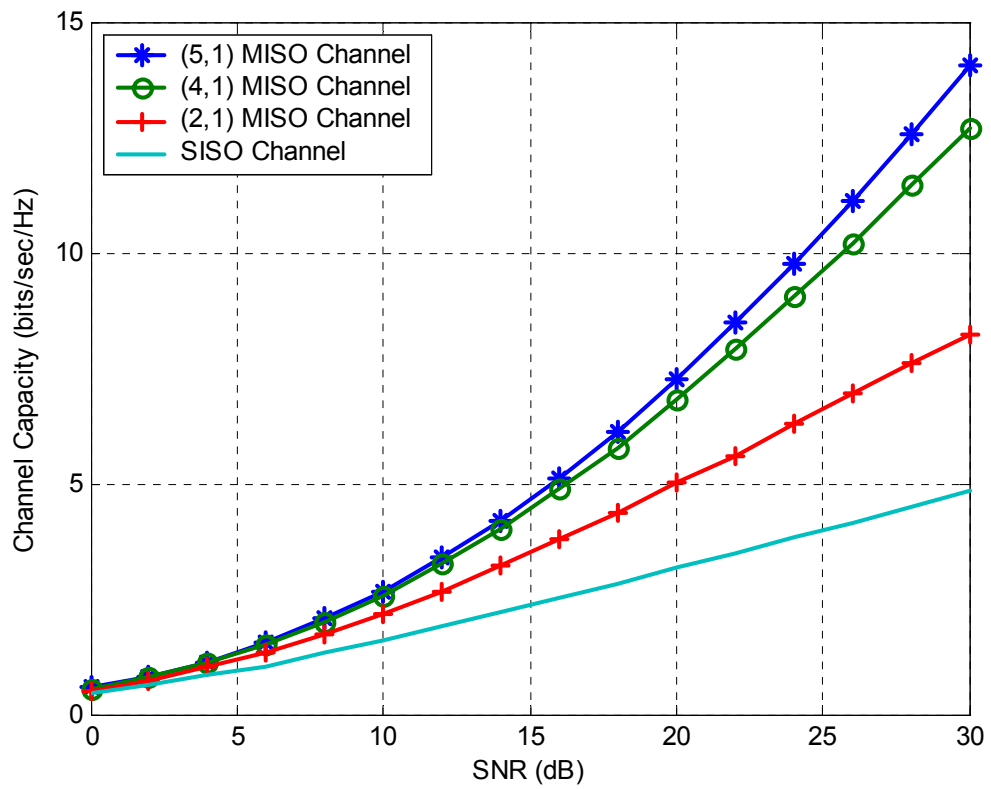
Figure 2.8:   Capacity of a SISO channel compared to the ergodic capacity of Rayleigh fading MISO channels with $(N_t, N_r) = (2, 1), (4, 1)$, and $(5, 1)$.
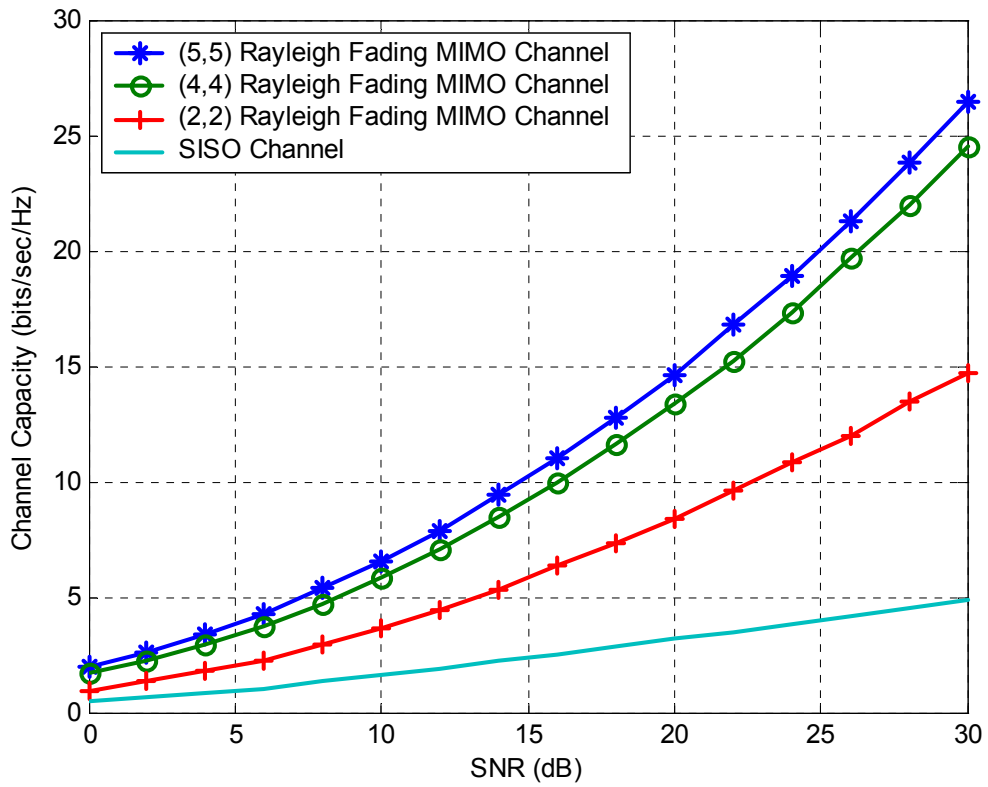
Figure 2.9:   Capacity of a SISO channel compared to the ergodic capacity of Rayleigh fading MIMO channels with (N_t, N_r) = (2, 2), (4, 4), and (5, 5).
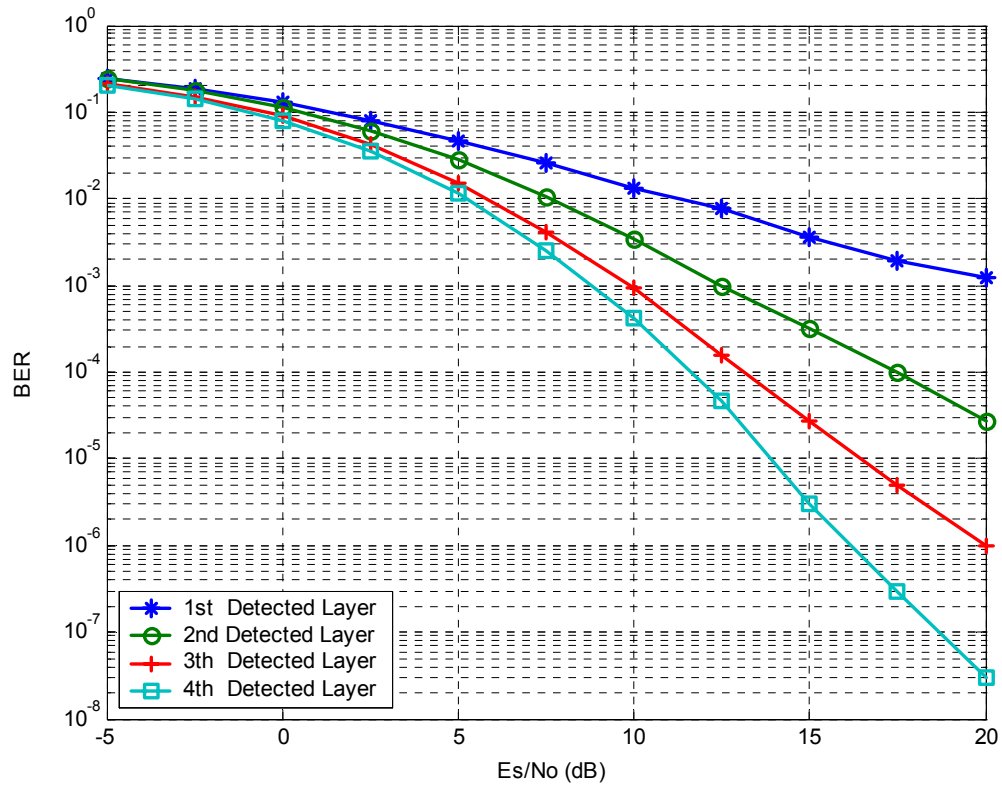
Figure 2.10: ZF V-BLAST BER performance with ideal detection and cancellation. QPSK modulation is used and $(N_t, N_r) = (4, 4)$.
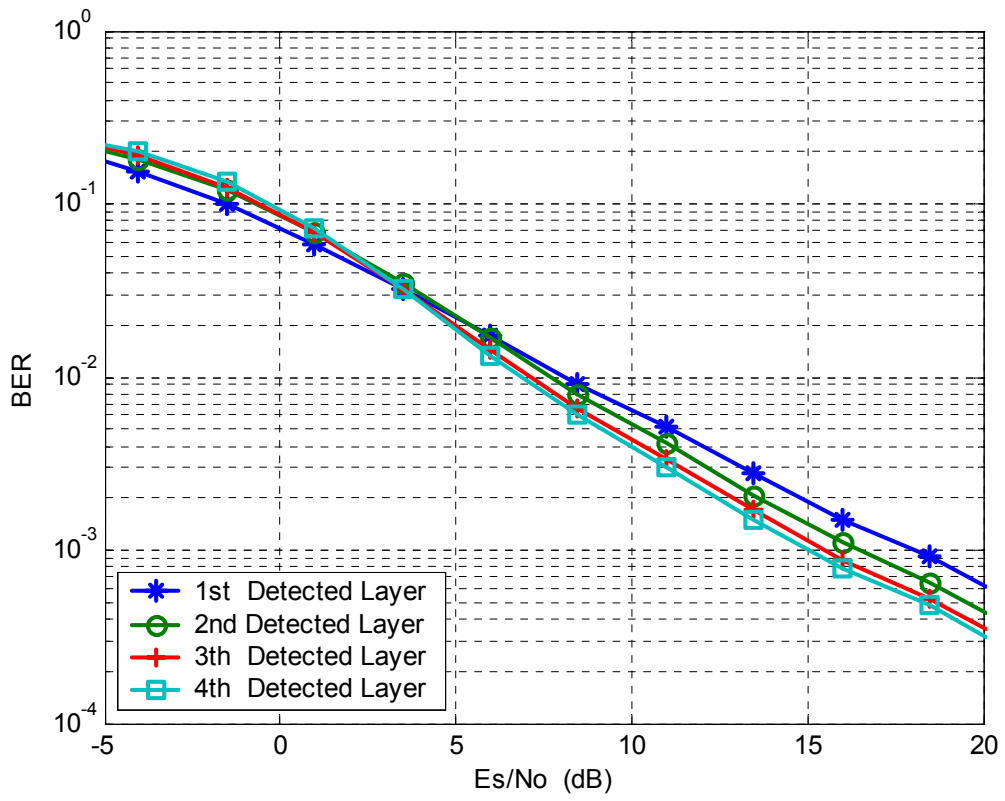
Figure 2.11: ZF V-BLAST BER performance with error propagation. QPSK modulation is used and $(N_t, N_r) = (4, 4)$.

Chapter 3

Multiuser OFDM Systems and Subcarrier Allocation Schemes

In recent years, there has been substantial research interest in applying orthogonal frequency division multiplexing (OFDM) to high speed wireless communications due to its advantage in mitigating the severe effects of frequency-selective fading [18]-[19].

By allowing a number of users to share an OFDM symbol, the multiple access concepts are possible in OFDM systems. Owing to the multiple access concepts, there are two kinds of resource allocation schemes. One is fixed resource allocation scheme [25] and the other is dynamic resource allocation scheme [26]-[32]. Fixed resource allocation scheme including time division multiple access (TDMA) and frequency division multiple access (FDMA) assigns an independent dimension, such as time slot or subchannel, to each user regardless of the current channel quality. On the other hand, dynamic resource allocation scheme allocates a dimension adaptively to each user based on users' channel qualities. Due to the time-varying nature of wireless channels, dynamic resource allocation scheme makes full use of the multiuser diversity to achieve better performance. In this chapter, the basic ideas and resource allocation schemes suited to the multiuser OFDM systems will be introduced.

3.1 Review of OFDM

OFDM is a special case of multicarrier transmission, where a single data stream is transmitted over a number of low data rate subcarriers. OFDM can be thought of as a

hybrid of multicarrier modulation (MCM) and frequency shift keying (FSK) modulation scheme. The principle of MCM is to transmit data by dividing the data stream into several parallel data streams and modulate each of these data streams onto individual subcarriers. FSK modulation is a technique whereby data is transmitted on one subcarrier from a set of orthogonal subcarriers in symbol duration. Orthogonality between these subcarriers is achieved by separating these subcarriers by an integer multiples of the inverse of symbol duration of the parallel data streams. With the OFDM technique used, all orthogonal subcarriers are transmitted simultaneously. In other words, the entire allocated channel is occupied through the aggregated sum of the narrow orthogonal subbands.

The main reason to use OFDM systems is to increase the robustness against frequency-selective fading or narrowband interference. In a single carrier system, a single fade or interference can cause the entire link fail, but in a multicarrier system, only a small amount of subcarriers will be affected. Then the error correction coding techniques can be used to correct errors. The equivalent complex baseband OFDM signal can be expressed as

$$
x(t) = \begin{cases} \displaystyle\sum_{k=-\frac{N_c}{2}}^{\frac{N_c}{2}-1} d_k \phi_k(t) & 0 \leq t \leq T \\ 0 & \text{Otherwise} \end{cases} = \left[\sum_{k=-\frac{N_c}{2}}^{\frac{N_c}{2}-1} d_k \phi_k(t) \right] u_T(t) \tag{3.1}
$$

where N_c is the number of subcarriers, T is the symbol duration, d_k is the transmitted subsymbol (*M*-PSK or *M*-QAM), $\phi_k(t) = e^{j2\pi f_k t}/\sqrt{T}$ is the *k*th subcarrier with the frequency $f_k = k/T$, and $u_T(t)$ is the time windowing function. Using the correlator-based OFDM demodulator, the output of the *j*th branch can be presented as

$$
y_j = \int_0^T x(t) \phi_j^*(t) dt = \frac{1}{T} \sum_{k=-\frac{N_c}{2}}^{\frac{N_c}{2}-1} d_k \int_0^T e^{j2\pi \frac{k-j}{T} t} dt \tag{3.2}
$$

$$
= d_j
$$

By sampling $x(t)$ with the sampling period $T_d = T/N_c$, the discrete time signal x_n can be expressed as

$$x_n = x(t)\Big|_{t=nT_d} = \begin{cases} \dfrac{1}{\sqrt{N_c}} \displaystyle\sum_{k=-\frac{N_c}{2}}^{\frac{N_c}{2}-1} d_k e^{j2\pi\frac{k}{N_c}n} & 0 \leq n \leq N_c - 1 \\ 0 & \text{Otherwise} \end{cases} = \text{IFFT}\{d_k\} \quad (3.3)$$

Note that x_n is the inverse fast Fourier transform (IFFT) output of the N input data subsymbols. Similarly, the output of the jth branch can also be presented in the digital form

$$y_j = \text{FFT}\{x_n\} = \frac{1}{\sqrt{N_c}} \sum_{n=0}^{N_c-1} x_k e^{-j2\pi\frac{j}{N_c}n} = \sum_{k=-\frac{N_c}{2}}^{\frac{N_c}{2}-1} x_k \delta[k-j] = d_j \quad (3.4)$$

In theory, the orthogonality of subcarriers in OFDM systems can be maintained and individual subcarriers can be completely separated by the fast Fourier transform (FFT) at the receiver when there are no intersymbol interference (ISI) and intercarrier interference (ICI) introduced by transmission channel distortions. However, it is impossible to obtain these conditions in practice. In order to eliminate ISI completely, a guard interval is imposed into each OFDM symbol. The guard interval is chosen larger than the expected delay spread, such that the multipath from one symbol cannot interfere with the next symbol. The guard interval can consist of no signals at all. However, the effect of ICI would arise in that case due to the loss of orthogonality between subcarriers. To eliminate ICI, the OFDM symbol is cyclically extended in the guard interval to introduce cyclic prefix (CP) as shown in Figs. 3.1-2. This ensures that delayed replicas of the OFDM symbol always have an integer number of cycles within the FFT interval, as long as the delay is smaller than the guard interval. As a result, the delayed multipath signals which are smaller than the guard interval will not cause ICI. The complete OFDM signal with CP is given by

$$\tilde{x}_n = \begin{cases} \dfrac{1}{\sqrt{N_c}} \displaystyle\sum_{k=-\frac{N_c}{2}}^{\frac{N_c}{2}-1} d_k e^{j2\pi\frac{k}{N_c}(n-N_{cp})} & 0 \leq n \leq N_c + N_{cp} - 1 \\ 0 & \text{Otherwise} \end{cases} \quad (3.5)$$

where N_{cp} is the number of samples in CP. Due to CP, the transmitted OFDM symbol

becomes periodic, and the linear convolution process of the transmitted OFDM symbols with the channel impulse responses will become a circular convolution one. Assuming the value of N_{cp} is larger than the channel length, the received data vector can be expressed as

$$\mathbf{y} = \mathbf{Hx} + \mathbf{\eta}$$

$$
\underbrace{\begin{bmatrix} y_{N_c-1} \\ \vdots \\ y_0 \end{bmatrix}}_{\mathbf{y}} = \underbrace{\begin{bmatrix} h_0 & h_1 & \cdots & h_{N_{cp}} & 0 & \cdots & 0 \\ 0 & h_0 & h_1 & \cdots & h_{N_{cp}} & 0 & \vdots \\ 0 & 0 & \ddots & \ddots & \cdots & \ddots & 0 \\ 0 & \cdots & 0 & h_0 & h_1 & \cdots & h_{N_{cp}} \end{bmatrix}}_{\mathbf{H}} \underbrace{\begin{bmatrix} x_{N_c-1} \\ \vdots \\ x_0 \\ x_{-1} \\ \vdots \\ x_{-N_{cp}} \end{bmatrix}}_{\mathbf{x}} + \underbrace{\begin{bmatrix} \eta_{N_c-1} \\ \vdots \\ \eta_0 \end{bmatrix}}_{\mathbf{\eta}}
$$
(3.6)

Applying SVD on the channel response, we have

$$\mathbf{H} = \mathbf{U\Sigma V}^H$$
(3.7)

where $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices, and $\mathbf{\Sigma}$ is a diagonal matrix. Substituting Equation (3.7) and the equalities of $\mathbf{x} = \mathbf{VX}$ and $\mathbf{Y} = \mathbf{U}^H \mathbf{y}$ into Equation (3.6), the received data vector can be written as

$$\mathbf{Y} = \mathbf{U}^H \mathbf{y} = \mathbf{U}^H (\mathbf{Hx} + \mathbf{\eta}) = \mathbf{U}^H \mathbf{HVX} + \underbrace{\mathbf{N}}_{\mathbf{U}^H \mathbf{\eta}} = \mathbf{\Sigma X} + \mathbf{N}$$
(3.8)

This means that the output $\mathbf{Y}$ can be expressed in terms of the product of $\mathbf{\Sigma}$ and $\mathbf{X}$ plus noise. When $x_{-i} = x_{N-i}$ for $i = 1, ..., N_{cp}$, a more compact matrix form of the guard interval can be written as

$$
\begin{bmatrix} y_{N_c-1} \\ \vdots \\ y_0 \end{bmatrix} = \begin{bmatrix} h_0 & h_1 & \cdots & h_{N_{cp}} & 0 & \cdots & 0 \\ 0 & h_0 & h_1 & \cdots & h_{N_{cp}} & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & h_0 & h_1 & \cdots & h_{N_{cp}} \\ h_{N_{cp}} & 0 & \cdots & 0 & h_0 & \cdots & h_{N_{cp}-1} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ h_1 & \cdots & h_{N_{cp}} & 0 & \cdots & 0 & h_0 \end{bmatrix} \begin{bmatrix} x_{N_c-1} \\ \vdots \\ x_0 \end{bmatrix} + \begin{bmatrix} \eta_{N_c-1} \\ \vdots \\ \eta_0 \end{bmatrix}
$$
(3.9)

where $\mathbf{H}$ becomes a circulant matrix ($\mathbf{H} = \mathbf{Q}^H \mathbf{\Lambda Q}$) and $\mathbf{Q}$ is a discrete Fourier

transform (DFT) matrix with the *l*th entry as

$$\mathbf{Q}_l = \frac{1}{\sqrt{N_c}} e^{-j2\pi \frac{l}{N_c}} \tag{3.10}$$

As in Equation (3.8), the received data **y** can be transformed into **Y**

$$\mathbf{Y} = \mathbf{Q}^H \mathbf{y} = \mathbf{Q}^H (\mathbf{Hx} + \boldsymbol{\eta}) = \underbrace{\mathbf{Q}^H \mathbf{HQ}^H}_{\Sigma} \mathbf{X} + \underbrace{\mathbf{Q}^H \boldsymbol{\eta}}_{\mathbf{N}} \tag{3.11}$$
$$= \boldsymbol{\Sigma}\mathbf{X} + \mathbf{N}$$

According to Equation (3.11), by adding CP to the OFDM symbol, the modulation in OFDM is equivalent to multiplying the frequency domain signals of the OFDM symbol with the channel's frequency response $\boldsymbol{\Sigma}$.

The block diagrams of the OFDM transceiver is shown in Fig. 3.3, where the upper path is the transmitter chain and lower path corresponds to the receiver chain. In the center, IFFT modulates a block of input values onto a number of subcarriers. In the receiver, the subcarriers are demodulated by the FFT, which performs the reverse operation of the IFFT. In fact, the IFFT can be made using the FFT by conjugating input and output of the FFT and dividing the output by the FFT size. This makes it possible to use the same hardware for both transmitter and receiver. This complexity saving is only possible when the transceiver doesn't have to transmit and receive simultaneously. The functions before the IFFT can be discussed as follows. Binary input data is first encoded by a forward error correction code. The encoded data is then interleaved and mapped onto QAM values. In the receiver path, after passing the radio frequency (RF) part and the analog-to-digital conversion (ADC), the digital signal processing starts with a training sequence to determine symbol timing and frequency offset. The FFT is used to demodulate all subcarriers. The FFT outputs are mapped onto binary values and decoded to produce binary output data. In order to successfully map the QAM values onto binary values, the reference phases and amplitudes of all subcarriers have to be acquired first.

In conclusion, OFDM is a powerful modulation technique that simplifies the removal of distortion due to the multipath channel and increases bandwidth efficiency. The key advantages of OFDM transmission scheme can be summarized as follows:

1. OFDM is an efficient way to deal with multipath. For a given delay spread, the implementation complexity is significantly lower than that of a single carrier system with an equalizer.

2. In relatively slow time-varying channels, it is possible to significantly enhance the capacity by adapting the data rate per subcarrier according to the signal-to-noise ratio (SNR) of that particular subcarrier.

3. OFDM is robust against narrowband interference because such interference affects only a small amount of subcarriers.

4. OFDM makes single-frequency networks possible, which is especially attractive for broadcasting applications.

3.2 Multiple Access Techniques

Multiple access concepts can be defined as the sharing of a fixed communication resource by a group of users. For wireless communications, the communication resource can be thought of as a hyperplane in frequency and time dimensions. The goal of multiple access techniques is to allow users to share the communication resource without creating unmanageable interference with each other. In the following, two of the most basic multiple access techniques for wireless communications will be reviewed: one is frequency division multiple access (FDMA) and the other is time division multiple access (TDMA).

In FDMA systems, the frequency-time plane is partitioned into non-overlapping frequency bands. Each of them serves a single user. Every user is therefore equipped with a transmitter for a given, predetermined, frequency band, and a receiver for each band which can be implemented as a single receiver for the entire range with a bank of bandpass filters for the individual bands. The graph in Fig. 3.4 (a) illustrates the FDMA concept. The spectral regions between adjacent channels are called guard bands, which help reduce the interference between channels. FDMA is used exclusively for analog cellular systems, even though FDMA can be used for digital systems in theory. Essentially, FDMA splits the allocated spectrum into many subchannels. In current analog cell systems, each subchannel is 30 kHz. The main advantage of FDMA is its simplicity. It doesn't require any coordination or synchronization among users since

each user can use its own frequency band without interference. However, this is also the cause of waste especially when the load is momentarily uneven since when one user is idle his share of bandwidth can't be used by other users. It should be noted that if the users have uneven long term demands, it is possible to divide the frequency range unevenly, i.e., proportional to the demands. However, FDMA is not flexible when adds a new user to the network requiring equipment modification, such as additional filters, in every other user. As a result, FDMA is considered the least efficient way of mobile communications. It is being replaced by the new digital networks such as TDMA.

In TDMA systems, sharing of the communication resource is accomplished by dividing the frequency-time plane into non-overlapping time slots which are transmitted in periodic bursts. Every user is allowed to transmit freely during the time slot assigned to him, that is, the entire system resources are devoted to that user during the assigned time slot. The graph in Fig. 3.4 (b) illustrates the TDMA concept. Time is segmented into intervals called frames. Each frame is further partitioned into user assignable time slots. An integer number of time slots constitutes a burst time or burst. Guard times are allocated between bursts to prevent overlapping of bursts. Each burst is comprised of a preamble and the message portion. The preamble is the initial portion of a burst used for carrier and clock recovery, station identification, and other housekeeping tasks. The message portion of a burst contains the coded information sequence. In some systems, a training sequence is inserted in the middle of the coded information sequence. The advantage of this scheme is that it can aid the receiver in mitigating the effects of the channel and interference. The disadvantage is that it lowers the frame efficiency; that is, the ratio of the bit available for messages to the total frames length.

The TDMA systems are designed for use in a range of environments and situations, form hand portable using in a downtown office to a mobile user traveling at high speed on the freeway. It also supports a variety of services for the end user, such as voice, data, fax, short message services, and broadcast messages. TDMA offers a flexible air interface, providing high performances in capacity, coverage, mobility, and capability to handle different types of user needs. While TDMA is a good digital system, it is still somewhat inefficient since it has no flexibility for varying data rates and has no

accommodations for silence in telephone conservation. TDMA also requires strict signaling and timeslot synchronization. A point worth noting is that both FDMA and TDMA system performances degrade in the presence of the multipath fading. More specifically, due to the high data rates of TDMA systems, the time dispersive channel (a consequence of delay spread phenomenon) causes intersymbol interference (ISI). This is a serious problem in TDMA systems thus requiring adaptive techniques to maintain system performance. The key advantages of TDMA can be summarized as follows:

1. Sharing single carrier frequency with multiple users.

2. Non-continuous transmission makes handoff simpler. It means that mobile assisted handoff is possible.

3. Less stringent power control due to the reduced interuser interference.

4. Slots can be assigned on demand (concatenation and reassignment). It means that bandwidth can be supplied on demand.

The disadvantages of TDMA can also be summarized as follows:

1. High synchronization overhead is needed.

2. Equalization is necessary for high data rates.

3. Power envelop will pulsate. It is caused by interfering with other devices.

4. High frequency/slot allocation complexity is needed.

3.3   Multiple Access in OFDM Systems

In one OFDM symbol, the data is modulated using N_c subcarriers and when there are multiple users, each user may be allocated a set of subcarriers, i.e., the data transmission to a particular user can take place using the allotted set of subcarriers. This is illustrated using the OFDM time and frequency grid shown in Fig. 3.5. Each rectangle in the grid can be considered as a resource and it can be allocated to any particular user in an OFDM symbol time. The user can employ a specific type of modulation depending on the quality of service (QoS) requirements and the channel quality allocated to him. Therefore, OFDM systems can offer flexibility in modulation order and multiple access. In addition, the base station must send the signaling information to each user indicating the subcarriers and time slots allotted to the user.

In general, it can be said that uncertainties concerning OFDM as a multiple access

concept concerns the system uplink. The main issue is to keep the mobile synchronized to the base station in time and frequency grid. This means that the mobiles must transmit the information with some timing advance due to the different propagation delay of the radio channels. The mobiles need to be synchronized to preserve the system orthogonality and avoid intercarrier interference (ICI). On the contrary, the downlink follows the same paths as the broadcast concepts and has already proven functional. All users are always orthogonal to each other because they are multiplexed in the base station. If a mobile loses the synchronization on the downlink, the only thing that happens is that the connection is lost. No other user will suffer from this type of failure. In the following, the OFDM transmission technique combining with multiple access techniques, such as FDMA and TDMA will be introduced.

3.3.1 Frequency Division Multiple Access (FDMA)

In OFDM-FDMA systems, each user is allocated a predetermined band of subcarriers and each user's data is transmitted using only the subcarriers allotted to the user. This scheme can also be defined as OFDMA systems. In each allocated subcarrier, adaptive bit loading can be performed depending on the subchannel SNR to achieve the user's requirements. By allocating distinct sets of subcarriers to different users, the available bandwidth can be flexibly shared between different users while avoiding any multiple access interferences (MAI) between them. There are variations possible in OFDM-FDMA systems, such as block FDMA and interleaved FDMA, which are stated below.

In the block FDMA scheme, each user is allocated a group of adjacent subcarriers as illustrated in Fig. 3.6. The different shades in Fig. 3.6 represent different users and the subcarriers allocated to them. The allocation of blocks to users can be accomplished by using the greedy algorithm. Assuming that each user is allocated only one block, the first block is allocated to the user with the best SNR of the block in the first step. Then, this particular user and the allotted block are removed from the search and the procedure is continued for the next block until all blocks are allocated. In most cases, the adjacent subcarrier gains are highly correlated and consequently, the bit loading of the block can be considered together. It means that the same modulation mode can be

used for all subcarriers in the block. The main advantage of the block FDMA scheme is easy allocation of subcarriers with less computational complexity but it lacks in robustness. There is a high probability that all subcarriers allocated to a user will fade at the same time. As a remedy, an improvement called interleaved FDMA will be introduced in the following.

In the interleaved FDMA scheme, the subcarriers assigned to a particular user are interlaced with other users' subcarriers in the frequency domain. If a deep fade occurs, only a single subcarrier of a particular user is affected and the data can be recovered by using coding techniques across many OFDM symbols. Fig. 3.7 illustrates the interleaved FDMA concept.

The key advantages of OFDM-FDMA can be summarized as follows:
1. No multiple access interferences (MAI).
2. Incoherent or coherent modulation.
3. Adaptation to channel characteristics
4. Robustness against estimation errors.

3.3.2 Time Division Multiple Access (TDMA)

In OFDM-TDMA systems, each user is assigned a time slot during which all the subcarriers can be used for the particular user. This is illustrated in Fig. 3.8. The duration of each slot is assumed equal to the OFDM symbol. Due to the variations of the subcarrier channel gains for users, channel gains for some subcarriers may be quite low, while they are quite high for other subcarriers. Owing to the channel gain variations regarding one specific subcarrier for different users, the subcarrier with a low channel gain for one user may experience a high channel gain for some other users. However, OFDM-TDMA doesn't provide the flexibility to adapt fully to these channel gain variations.

The key advantages of OFDM-TDMA can be summarized as follows:
1. Incoherent or coherent modulation.
2. Adaptation to channel characteristics.
3. High coding gain introduced by the diversity.
4. Robustness against estimation errors.

5. No multiple access interferences (MAI) in case of synchronization errors
6. Easy implementation.

3.4 Dynamic Subcarrier Allocation Algorithms for Multiuser OFDM Systems

In multiuser systems using static FDMA or TDMA as multiple access schemes, each user is allocated a predetermined frequency band or time slot to apply OFDM. If there are unused subcarriers within the allocated frequency band or time slot of a user, these subcarriers will be wasted and will not be used by other users. However, the subcarriers which appear in deep fade to one user may not be in the same fade for other users. As the fading parameters for different users are mutually independent, it is quit unlikely that a subcarrier will be in deep fade for all users. Therefore, methods for adaptively assigning subcarriers to users have been widely investigated. These adaptive subcarrier assignment algorithms can be geared to decrease the power consumption for a given achievable data rate or to increase the data rate while the available power is limited. However, complexity is an important issue in computing an optimal assignment for a given description of all subcarrier channel qualities regarding all users. Recent work has shown a tradeoff between achieved results regarding the chosen metric and complexity of the assignment algorithm [22]-[24].

In order to minimize the transmit power under data rate constraints, an algorithm based on Lagrange optimization was derived by Wong et al. [22]. This algorithm nearly reaches the optimal solution. However, due to its complexity and its slow convergence, it is computationally very expensive. As a remedy, Kivanc et al. presented suboptimal heuristics for the same problem [23]. While these heuristics cannot achieve the optimal lower bound on the consumed power, they are much cheaper in terms of complexity.

The alternative problem, maximizing the data rate of multiple users while the transmit power is limited, can be decomposed into two steps: first determining the number of subcarriers each user is assigned and then assigning the concrete subcarriers to users. Yin et al. suggest a scheme where the number of subcarriers assigned to each user is determined first by considering each user's average channel gains and its

required data rate [24]. Then, the subcarrier assignment is done. In the following, three different dynamic algorithms for assigning subcarriers to each user will be introduced.

3.4.1 Basic Subcarrier Allocation Algorithm

First, the system model will be presented. U users locating in a cell with radius R are assumed. The downlink data transmission is considered. The channel is assumed to be linear and time invariant. For data transmission, an OFDM symbol is employed, providing a total of N_c subcarriers. With the specific channel gain and noise level of the subcarrier v for the user u given by $H_{u,v}$ and $n_{u,v}$ respectively, the channel gain and noise levels of all subcarriers can be expressed as vector forms:

$$\mathbf{H_u} = [H_{u,1}, H_{u,2}, \cdots, H_{u,N_c}] \tag{3.12}$$

$$\mathbf{n_u} = [n_{u,1}, n_{u,2}, \cdots, n_{u,N_c}] \tag{3.13}$$

According to Equations (3.12) and (3.13), the channel gain-to-noise ratio (CNR) for the user u is given by

$$\mathbf{T_u} = [\frac{|H_{u,1}|^2}{\Gamma_u |n_{u,1}|^2}, \frac{|H_{u,2}|^2}{\Gamma_u |n_{u,2}|^2}, \cdots, \frac{|H_{u,N_c}|^2}{\Gamma_u |n_{u,N_c}|^2}] \tag{3.14}$$

such that the SNR gap Γ_u is given by

$$\Gamma_\mathbf{u} = \frac{1}{3} \cdot [Q^{-1}(\frac{p_{s,u}}{4})]^2 \tag{3.15}$$

where $Q^{-1}[\cdot]$ is the inverse Q-function and $P_{s,u}$ is the symbol error probability. According to Equation (3.15), different symbol error rates and different QoS classes can be defined for each other. A coding gain and a margin can be considered in Γ_u, thus allowing different users to employ different coding schemes. At the beginning, the base station knows about the CNRs for all users, given by $\mathbf{T}$:

$$\mathbf{T} = [T_1, T_2, \cdots, T_u] \tag{3.16}$$

The CNR value for the subcarrier v and user u, $T_{u,v}$, can be transformed into an

signal-to -noise ratio (SNR) value by multiplying it by the assigned power value p_v for the subcarrier v, yielding the SNR value denoted by $x_{u,v} = H_{u,v} \cdot p_v$. Afterwards, the resulting SNR value yields an achievable data rate for this subcarrier by a function $F(\cdot)$, which is given by the considered modulation type and the maximum tolerable symbol error probability. In the first step, the base station allocates the number of subcarriers to be assigned to each user. Let the number of allocated subcarriers for the uth user is given by k_u. The goal is to assign specific subcarriers to each user such that the total data rate is maximized. The assignment of a particular subcarrier v to the user u is given by $c_{u,v} = 1$. Otherwise, the value is set to zero. Hence, the optimization problem can be formulated as

$$\max \sum_{\forall u,v} c_{u,v} \cdot F(p_v \cdot T_{u,v}) \tag{3.17}$$

subject to:

$$\forall u : \sum_{\forall v} c_{u,v} = k_u$$
$$\forall v : \sum_{\forall u} c_{u,v} = 1$$

The subcarrier assignment is assumed to be perfectly signaled to users to inform them which modulation types and subcarriers they are supposed to use.

Then, the basic allocation algorithm will be introduced below [33]. In the first step, the algorithm assigns priorities 1 to U to each user, where priority 1 is the highest one. In the second step, the algorithm iterates over all users, sorted by their current priority. In each iteration step, the considered user is assigned the k_u best subcarriers available in terms of CNR values. In the third step, these assigned subcarriers are removed form the list of available subcarriers. The disadvantage of this algorithm is that the user with a lower priority will never be assigned the subcarriers with a better quality than the user with a higher priority. To balance the unfairness, the priorities are switched after each transmission packet. While the channel gains and noise levels change from transmission packet to transmission packet, they are assumed to be fixed during one transmission packet. The priority of each user is decreased by one for each transmission packet. The user with the highest priority (priority 1) receives for the next transmission

packet the lowest priority (priority U). The initial setting of priorities is generated randomly.

Sorting a set of n elements requires an algorithmic complexity of $O(n\log(n))$. Therefore, the complexity of the basic allocation algorithm is given by $O(U \cdot N_c \log(N_c))$. Then, the basic allocation algorithm can be summarized as follows:

Initialization:

Initialize $(N_c, \mathbf{T})$;

Recursion:

Iterate over User u
sorted by priority p
{
  SORT $(N_c, \mathbf{T_u})$;
  ASSIGN _ BEST (k_u, N_c);
}

3.4.2 Advanced Subcarrier Allocation Algorithm

In the advanced allocation algorithm [33], a scheme will be introduced where additional information, the weight, about subcarriers is used. When considering a particular user for subcarrier assignment, the weight of a subcarrier expresses how well this subcarrier may be used by all other users with a lower priority than the currently target one. A weight $w_{u,v}$ of the subcarrier v for the user u is given by the sum of all CNR values of this subcarrier regarding all users with a lower priority than the user u. The weight $w_{u,v}$ can be written as

$$w_{u,v} = \sum_u T_{u,v} \qquad \forall u \text{ with lower priority than } u \tag{3.18}$$

Then, the weight ratio of the subcarrier v with respect to the user u can be defined as

$$r_{u,v} = \frac{T_{u,v}}{w_{u,v}} \tag{3.19}$$

The advanced allocation algorithm extends the basic allocation algorithm by always

selecting the subcarrier with the highest possible weight ratio between CNR value and weight. By choosing the subcarriers with the highest possible weight ratios, this algorithm assigns the user with a higher priority with the particular ones which have better qualities and cannot be used well by all other users with a lower priority than the currently target one. As in the basic allocation algorithm, subcarriers are assigned to users according to their priority. The user u is assigned k_u subcarriers with the largest weight ratios. The disadvantage of this algorithm is that weights for all unassigned subcarriers need to be recomputed and sorted with respect to the next user to be assigned after an assignment of subcarriers to one user.

The advanced allocation algorithm has the same complexity of $O(U \cdot N_c \log(N_c))$ as that of the basic allocation algorithm. For each assignment step, the algorithm has to sort the remaining subcarriers by the weight ratios, which have to be computed first. Therefore, the execution time of the advanced allocation algorithm is higher than that of the basic allocation algorithm. The advanced allocation algorithm can be summarized as follows:

> *Initialization:*
>
> *Initialize* $(N_c, \mathbf{T})$;
>
> *Recursion:*
>
> *Iterate over User u*
> *sorted by priority p*
>
> $\{$
> $\mathbf{w_u} = CALC\_WEIGHT\ (N_c, u)$;
> $\mathbf{r_u} = CALC\_RATIO\ (\mathbf{T_u}, \mathbf{w_u})$;
> $SORT\ (N_c, \mathbf{r_u})$;
> $ASSIGN\_BEST\ (k_u, N_c)$;
> $\}$

3.4.3 Two-Stage Subcarrier Allocation Algorithm

In the above allocation algorithms, the number of allocated subcarriers for each user is assumed already known. In fact, it has to be estimated due to the limit of

available subcarriers. In the following, the two-stage allocation algorithm [32] will be introduced including the estimation of the number of allocated subcarriers for each user and the determination of which subcarriers are given to which user. The two-stage allocation algorithm involves two stages based on the following reasons:

1. The resources for one user, i.e., the number of subcarriers and the transmit power, mainly depend on the mean CNR of its channel.

2. Which subcarrier is assigned to one user depends on the CNR according to Equations (3.14) and (3.16).

In the first stage, the number of allocated subcarriers for each user has to be estimated. It takes into account the users' mean CNRs, the desired minimum data rate $B_{min}(u)$, and the total transmit power P_t. The mean CNR of the user u can be defined as

$$\overline{T}_u = \frac{1}{N_c} \sum_{v=1}^{N_c} T_{u,v} \qquad (3.20)$$

where $T_{u,v} = |H_{u,v}|^2 / \Gamma_u |n_{u,v}|^2$ is the CNR value of the user u and subcarrier v. Each user is assigned a number of subcarriers k_u such that the desired data rate $B_{min}(u)$ can be achieved. The corresponding energy for the user u with the number of allocated subcarriers k_u can be defined as

$$E_{tot}(u) = k_u \overline{T}_u^{-1} (2^{B_{min}(u)/k_u} - 1) \qquad (3.21)$$

Besides, the sum of total required energy for all users must be smaller than the total transmit power. At the beginning, k_u can be calculated as if the maximum number of bits per symbol b_{max} could be applied to all subcarriers:

$$k_u = [\frac{B_{min}(u)}{b_{max}}] \qquad (3.22)$$

Normally, in this step much more subcarriers are assigned than available. Then, a subcarrier will be removed form the user who has to increase the smallest amount of transmit power without this subcarrier. This procedure is repeated until exactly N_c subcarriers are granted. On the contrary, if there are much less subcarriers assigned than available, new subcarriers will be assigned for each user until the total required

energy does not exceed the total transmit power P_t. This procedure repeats until no subcarriers are left. Fig. 3.9 presents the flow chart of the estimation of the number of allocated subcarriers for each user. It can be stated as follows:

Initialization:

$$k_u = [\frac{B_{\min}(u)}{b_{\max}}]$$

$$E_{tot}(u) = k_u \overline{T}_u^{-1}(2^{B_{\min}(u)/k_u} - 1), \ \forall u$$

Recursion:

while $\sum_u k_u < N_c$
  for $u = 1, \ldots, U$
    while $\sum_u E_{tot}(u) < P_t$
    $k_u = k_u + 1$
    $E_{tot}(u) = k_u \overline{T}_u^{-1}(2^{B_{\min}(u)/k_u} - 1)$
  end
end
while $\sum_u k_u > N_c$
  $E_{new}(u) = (k_u - 1)\overline{T}_u^{-1}(2^{B_{\min}(u)/(k_u-1)} - 1), \ \forall u$
  $u' = \arg \min \{E_{new}(u) - E_{tot}(u)\}$
  $k_{u'} = k_{u'} - 1; \ E_{tot}(u') = E_{new}(u')$
end

In the second stage, which subcarriers are given to which user has to be determined. The idea for the subcarrier assignment is that the users choose alternatively the subcarrier with the best CNR. Therefore, the order in which the users choose their subcarriers is important. A procedure based on priorities controls the order. The reference priority can be defined as the number of subcarriers of the user u over the total number of subcarriers.

$$p_0(u) = \frac{k_u}{N_c} \tag{3.23}$$

After the user u has chosen one subcarrier, k_u is subtracted by one. Afterwards, k_u stands for the number of subcarriers that are still to be assigned. Then, the actual priority of the user u can be defined as

$$p(u) = \frac{k_u}{\displaystyle\sum_{u=1}^{U} k_u}, \quad u = 1, \ldots, U \tag{3.24}$$

The user with the most subcarriers begins the subcarrier assignment, then after each step the user with the greatest difference between the reference and actual priority is picked out for the next turn. The subcarrier allocation matrix can be defined as $\mathbf{c} = (c_{u,v})$, with $c_{u,v} = 1$ if the subcarrier v is assigned to the user u, and zero otherwise. Fig. 3.10 presents the flow chart of subcarrier assignment which can be stated as follows:

Initialization:

$$\mathbf{c} = o, \; p_0(u) = \frac{k_u}{N_c}, \quad \forall u$$

$$U = \left\{ u \middle| u = \arg\max_{u'=1,\ldots,U} \{k_{u'}\} \right\}$$

for $u \in U$

$$v_1 = \arg\min_{v \in M} \{T_{u,v}^{-1}\}, \; \text{with } M = \left\{ v \middle| \sum_{u'=1}^{U} c_{u',v} = 0 \right\}$$

$$k_u = k_u - 1$$

$$c_{u,v_1} = 1$$

end

Recursion:

while $\sum_u k_u > 0$

$$p(u) = \frac{k_u}{\displaystyle\sum_{u'} k_{u'}}, \quad \forall u$$

$$U = \left\{ u \middle| u = \arg\max \{p(u') - p_0(u')\} \right\}$$

for $u \in U$

$$v_1 = \arg\min_{v \in M} \{T_{u,v}^{-1}\}, \; \text{with } M = \left\{ v \middle| \sum_{u'=1}^{U} c_{u',v} = 0 \right\}$$

$$k_u = k_u - 1$$

$$c_{u,v_1} = 1$$

end

end

3.5   Computer Simulations

Computer simulations are conducted to evaluate the performance of above mentioned dynamic subcarrier allocation algorithms employed in multiuser OFDM systems, including the proposed one presented in Section 3.4.3. Two assumptions are made in the following simulations: perfect CSI is available at the transmitter side and the tap gains are constant over an OFDM symbol (quasi-static fading). The standard wireless exponential decay channel model with delay profile illustrated in Fig. 3.11 is employed in the simulations [40]. The corresponding channel impulse response can be expressed as:

$$\begin{cases} h_i = N(0, \frac{1}{2}\sigma_k^2) + jN(0, \frac{1}{2}\sigma_k^2) \\ \sigma_k^2 = \sigma_0^2 e^{-kT_{sample}/\tau_{rms}} \\ \sigma_0^2 = 1 - e^{-T_{sample}/\tau_{rms}} \end{cases} \tag{3.25}$$

where T_{sample} is the sampling period and τ_{rms} is the root mean square delay spread of the channel. The variance σ_0^2 is chosen to normalize the average receive power to one, i.e. $\sum_k \sigma_k^2 = 1$. The number of samples taken in the impulse response should ensure sufficient decay of the impulse response tail, such that $i_{max} = \tau_{rms}/T_{sample}$.

In the first simulation, the fading effect of a system with the subcarrier number varied is examined. Figs. 3.13-14 show the subcarrier channel gains and corresponding number of bits for two users and thirty users, respectively. Owing to the channel gain variations associated with one specific subcarrier for different users, the subcarrier with poor channel quality for one user may experience good channel qualities for other users. For instance, the 20th subcarrier has a low channel gain for the 20th user, but it has a high channel gain for the 30th user. It appears that the 30th user is more suitable to use the 20th subcarrier than the 20th user. In fact, it's quite improbable that a subcarrier is unsuitable to use for all users.

In the second simulation, the data rate is evaluated as a function of input SNR E_s/N_0. The corresponding results obtained with different numbers of users are shown in Fig. 3.15. The data rate increases when more users are simultaneously considered. This is because the multiuser diversity possesses a great chance for subcarriers to have

good channel qualities when more users are considered.

In the third simulation, the efficacy of three dynamic subcarrier allocation algorithms is investigated as a function of user number. The results shown in Fig. 3.16 indicate that the two-stage subcarrier allocation algorithm outperforms other two subcarrier allocation algorithms. This is because that the two-stage subcarrier allocation algorithm allocates one subcarrier to one user at a time according to the order, while the basic and advanced subcarrier allocation algorithms allocate a group of subcarriers to one user at a time. Therefore, the two-stage subcarrier allocation algorithm will allocate to each user the most appropriate subcarriers and possesses better performance as compared with other methods.

In the final simulation, the computational complexity in terms of execution time is examined as a function of user number. The results shown in Fig. 3.17 obviously indicate that the execution time increases when the user number increases. The basic subcarrier allocation algorithm requires the lowest execution time because it assigns the best k_u subcarriers to the current target user in each iteration step without computing coefficients such as $w_{u,v}$, $p_0(u)$, and $p(u)$ used in the advanced and two-stage subcarrier allocation algorithms. Besides, the advanced subcarrier allocation algorithm requires the most execution time because it needs to recompute the weights for all unassigned subcarriers and sort with respect to the next user to be assigned after an assignment of subcarriers to one user. Although, the advanced and two-stage subcarrier allocation algorithms require higher execution time, they are still computationally efficient algorithms.

3.6 Summary

Orthogonal frequency division multiplexing (OFDM) has gained wide acceptance in wireless communications as an appropriate broadband modulation scheme. OFDM systems have desirable immunity to intersymbol interference (ISI) caused by delay spread of wireless channels. Therefore, it is a promising technique for high data rate transmission over frequency-selective fading channels. In Section 3.1, OFDM systems are introduced.

Since multiple access technique plays an important role in multiuser systems, different multiple access schemes for OFDM systems are developed. Time division multiple access (TDMA) and frequency division multiple access (FDMA) assign an independent dimension, e.g. time slot or subchannel, to each user which are introduced in Sections 3.2 and 3.3. However, TDMA and FDMA are not optimal because they are fixed regardless of the current channel quality. Therefore, the two-stage dynamic subcarrier allocation algorithm is proposed in Section 3.4.3 to allocate subcariers adaptively to each user based on users' channel qualities and quality of service (QoS) requirements. The simulation results show that the two-stage dynamic subcarrier allocation algorithm can indeed enhance the overall data rate of multiuser OFDM systems.

| $X_{N-N_{cp}}$ | .... | X_{N-1} | X_0 | X_1 | .... | X_{N-1} |

$$\overset{\longleftarrow \text{Cyclic Prefix} \longrightarrow}{N_{cp}}$$ $$\overset{\longleftarrow \text{Useful Part} \longrightarrow}{N}$$

$$\overset{\longleftarrow \text{Complete OFDM Signal} \longrightarrow}{N+N_{cp}}$$
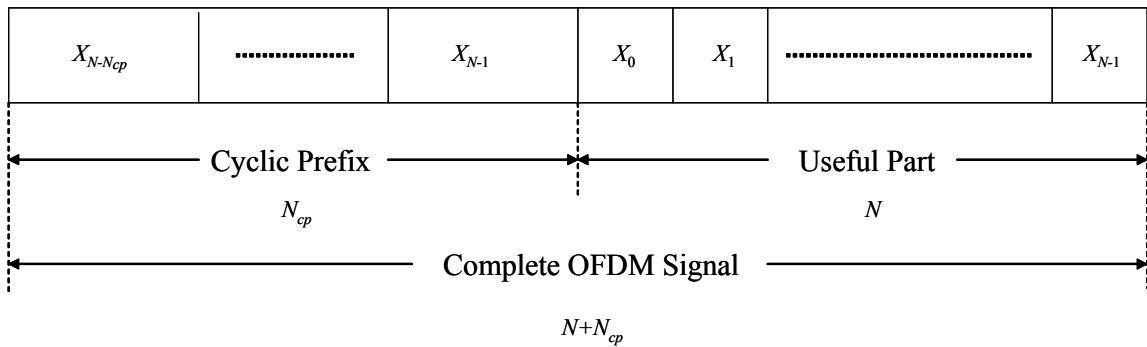
Figure 3.1: OFDM signal with cyclic prefix extension.
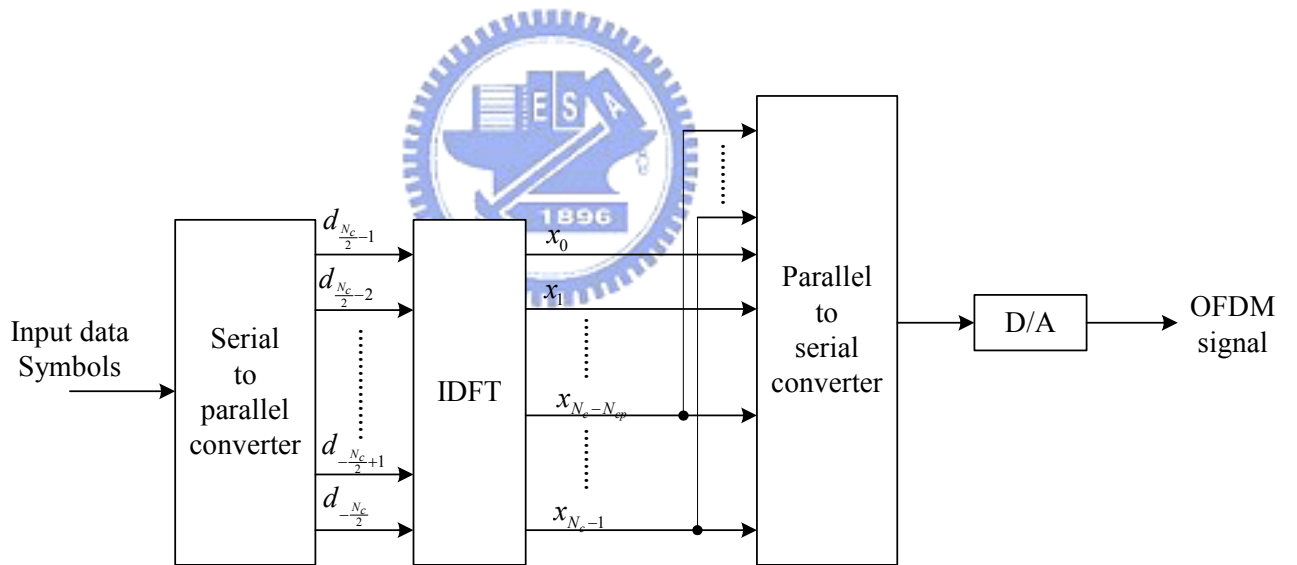
Figure 3.2: A digital implementation of appending cyclic prefix into OFDM signal in the transmitter.
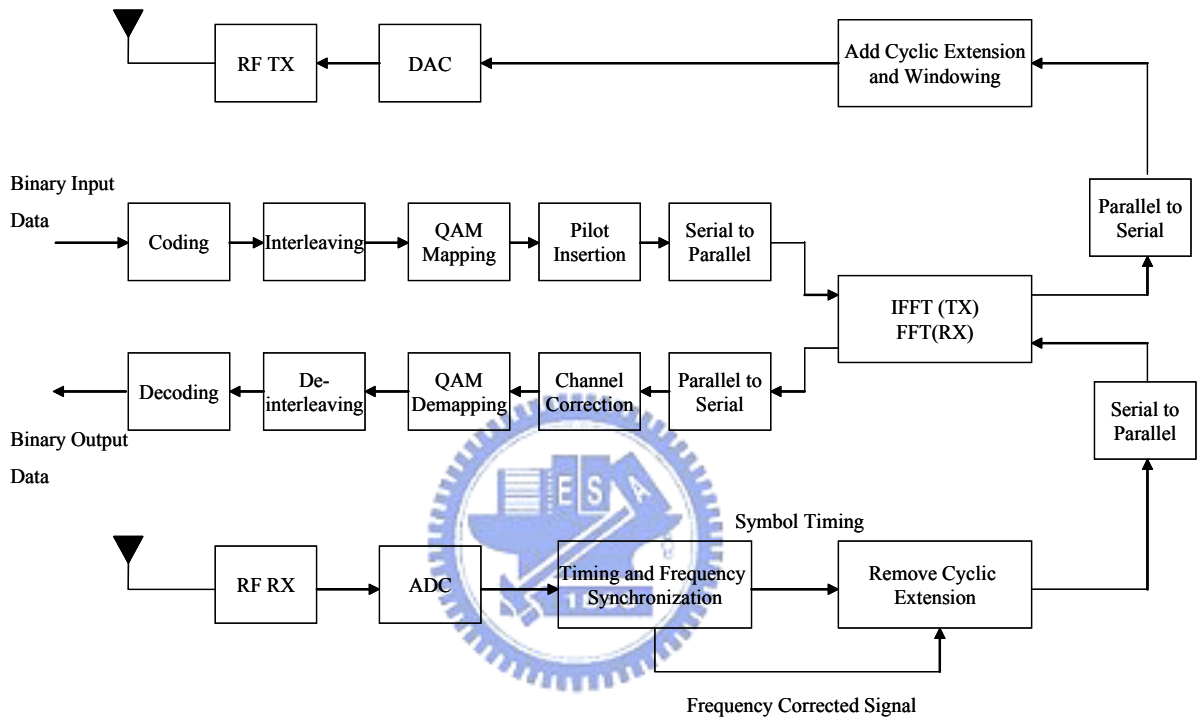
Figure 3.3: Black diagrams of an OFDM transceiver.

(a)

(b)
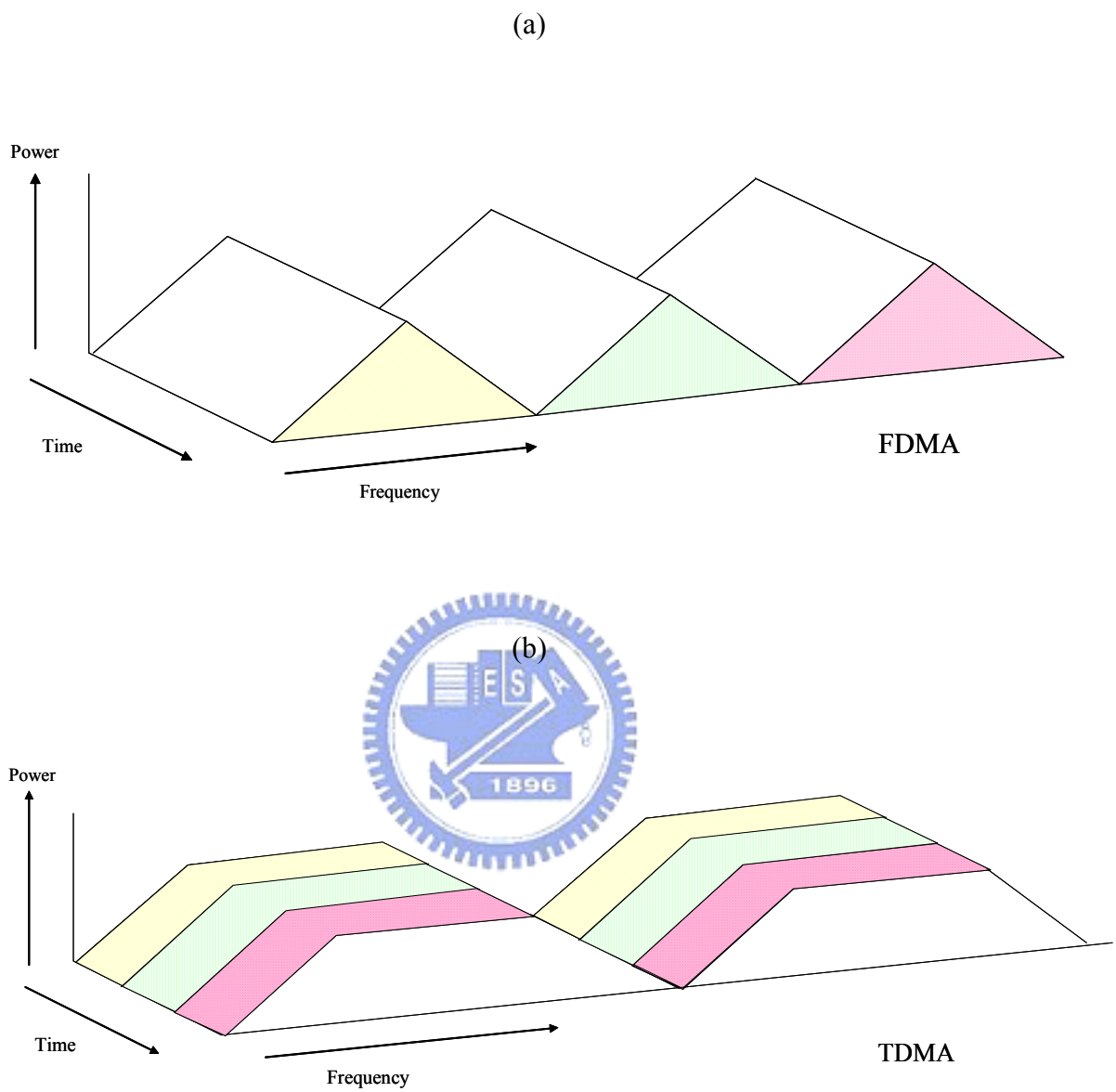
Figure 3.4: Illustration of different multiple access techniques (a) FDMA and
(b) TDMA.
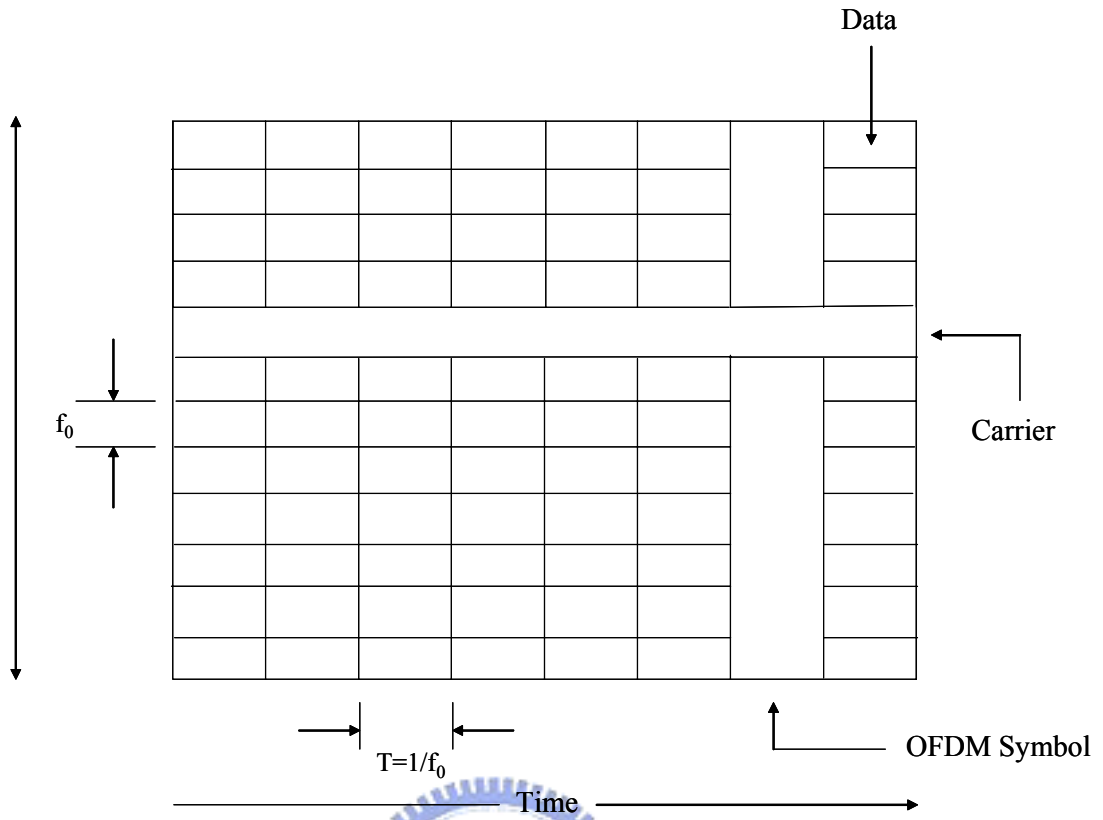
Figure 3.5: OFDM time-frequency grid.
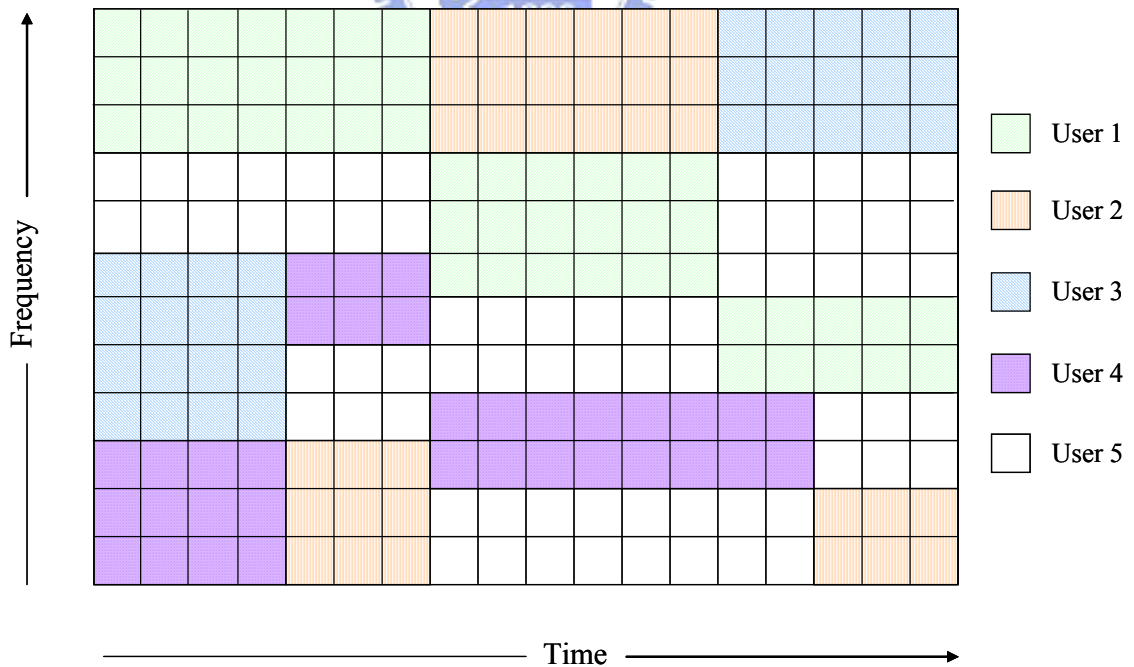
Figure 3.6: Illustration of block FDMA.

55
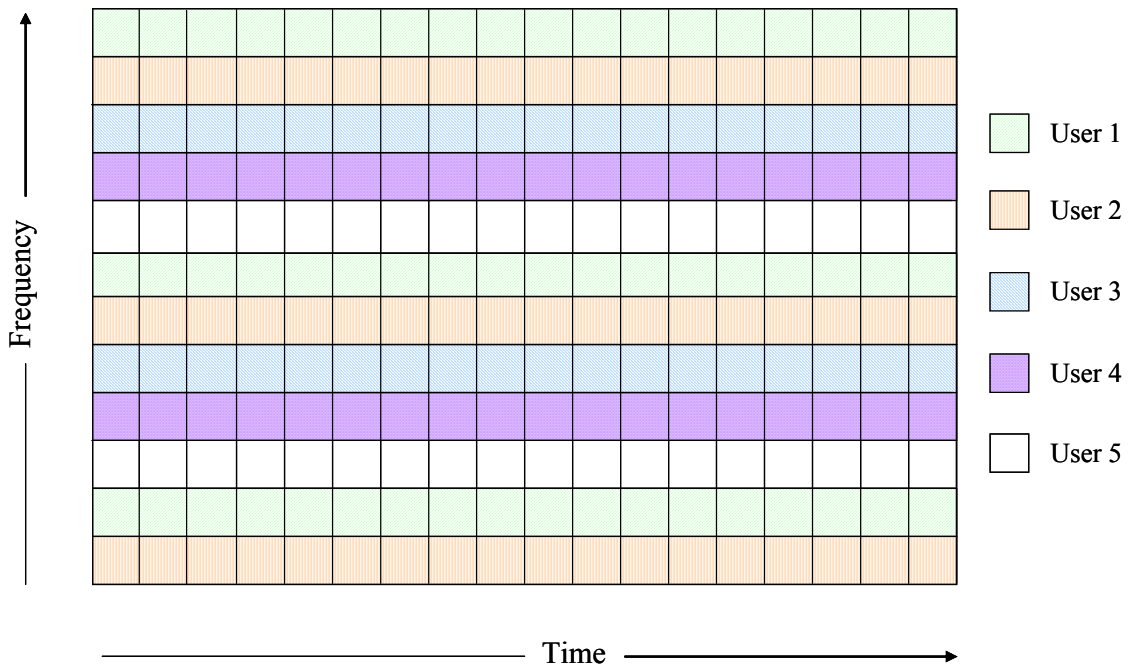
Figure 3.7: Illustration of interleaved FDMA.
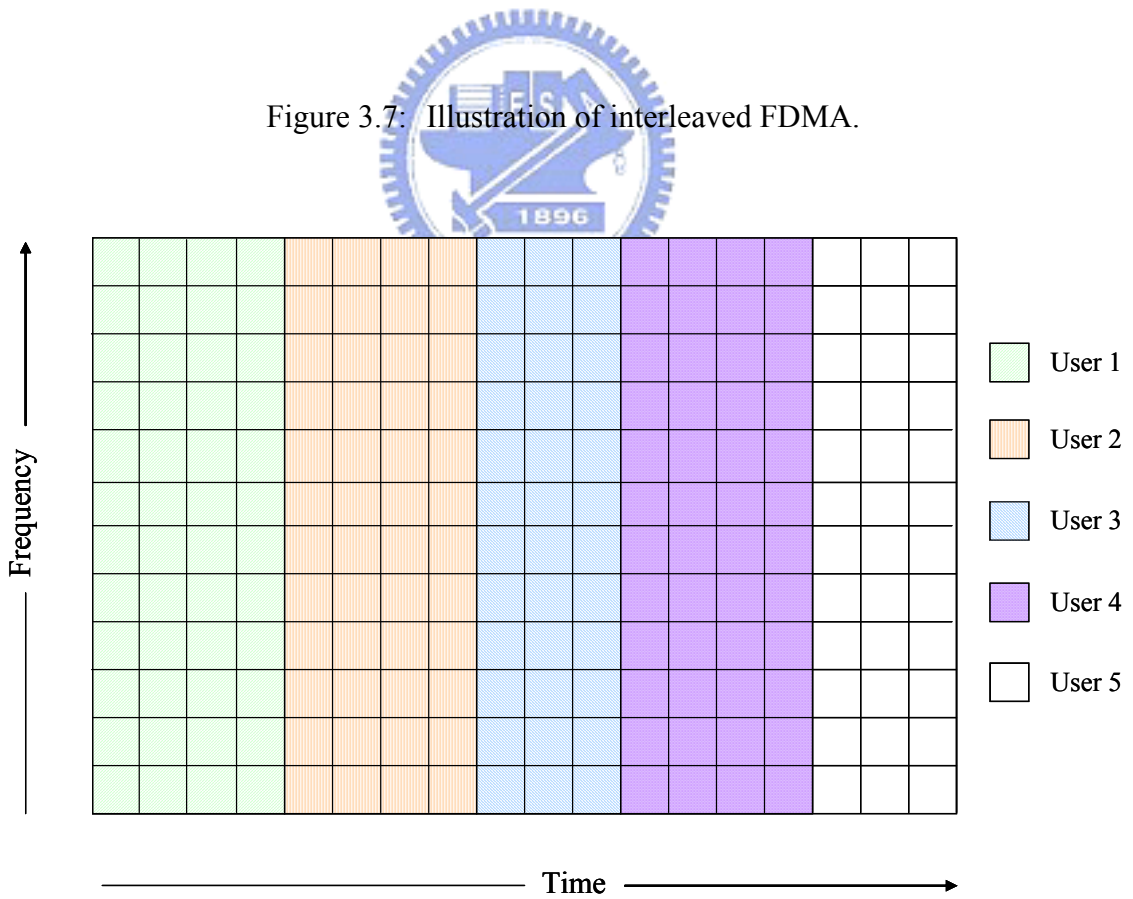
Figure 3.8: Illustration of OFDM-TDMA.
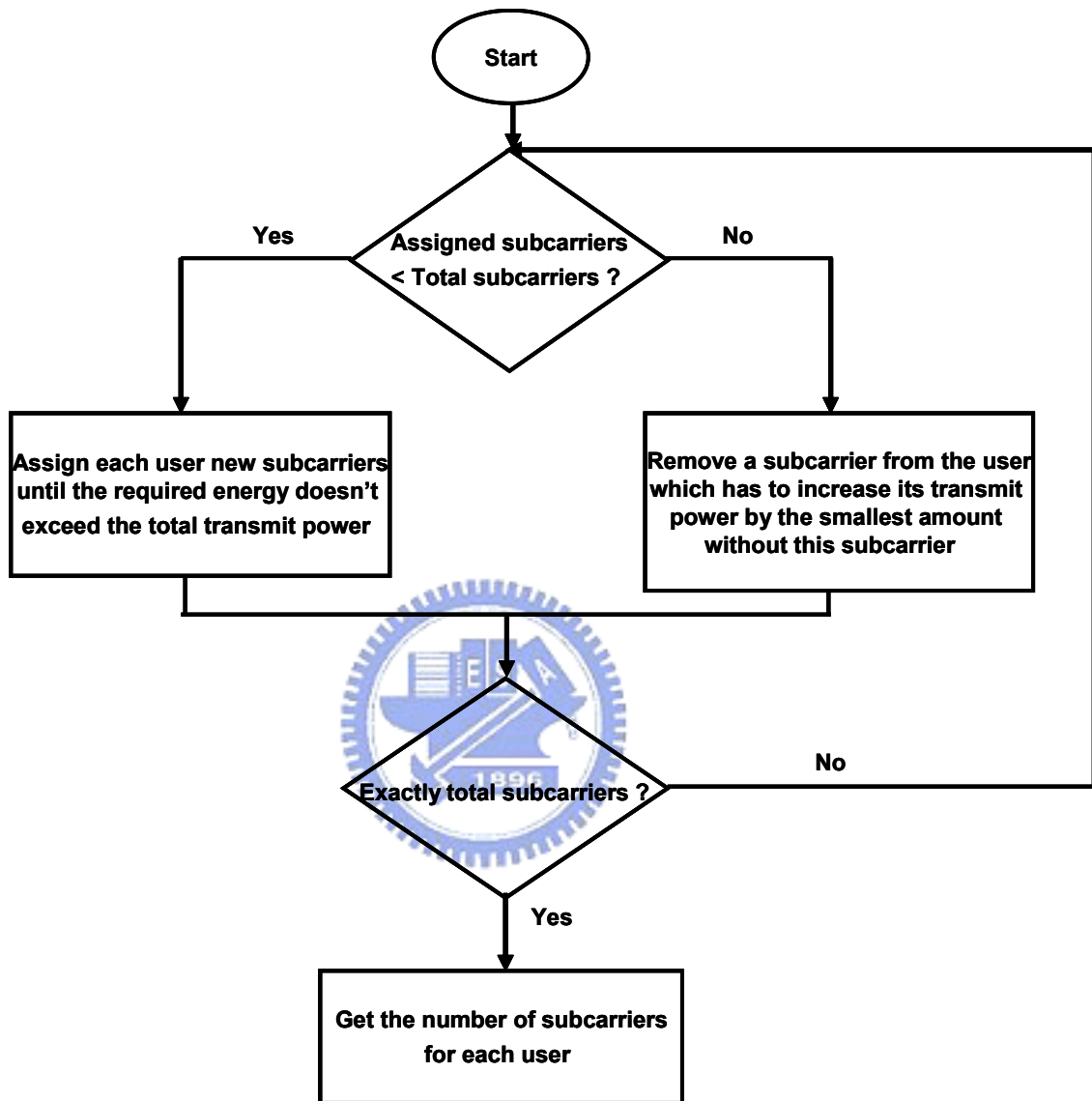
Figure 3.9: Flow chart of the estimation of the number of allocated subcarriers for each user.
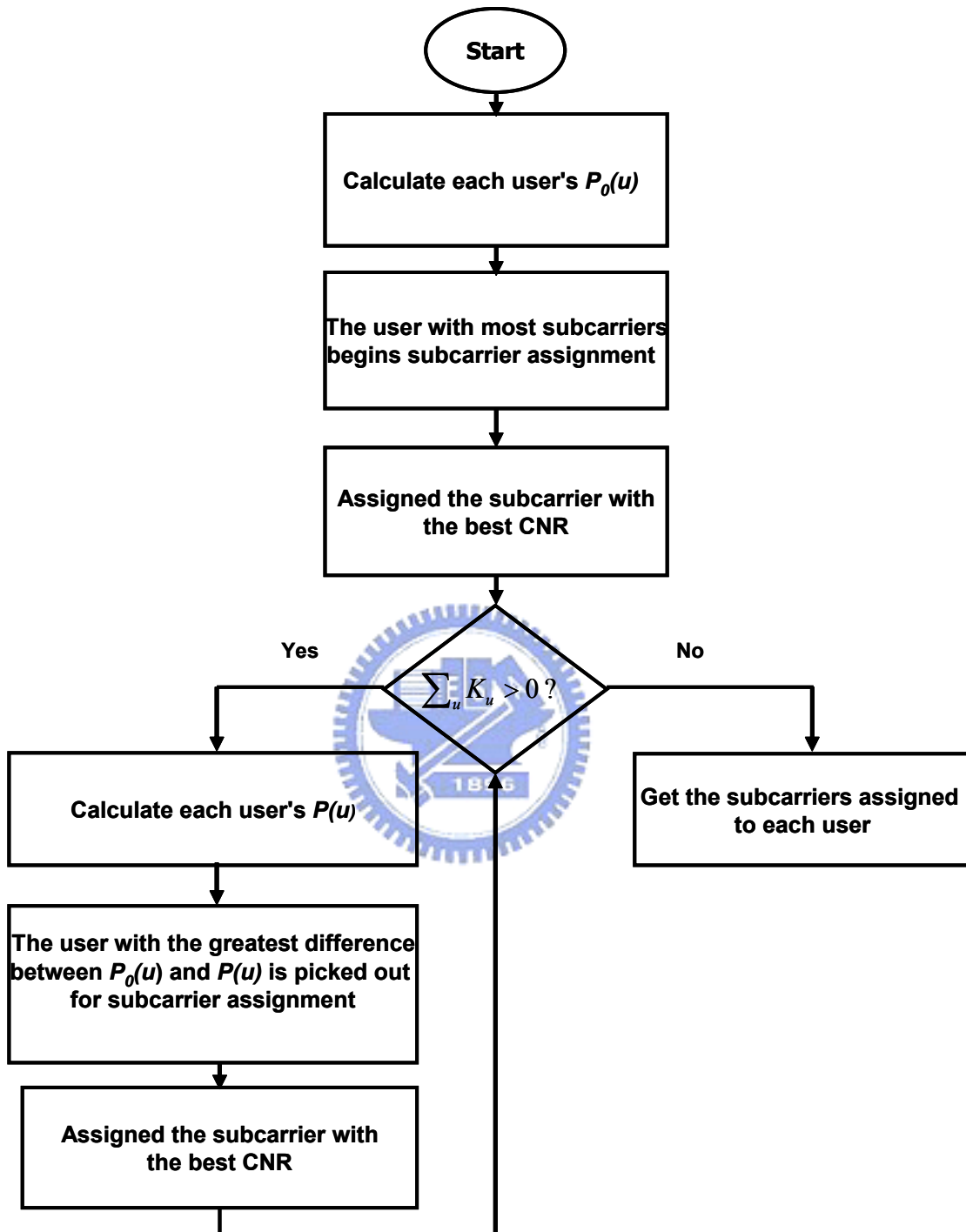
Figure 3.10: Flow chart of the subcarrier assignment.

Magnitude

0   T_s $2T_s$ $3T_s$ $4T_s$ $5T_s$ $6T_s$ $7T_s$ $8T_s$ $9T_s$ $10T_s$ Time
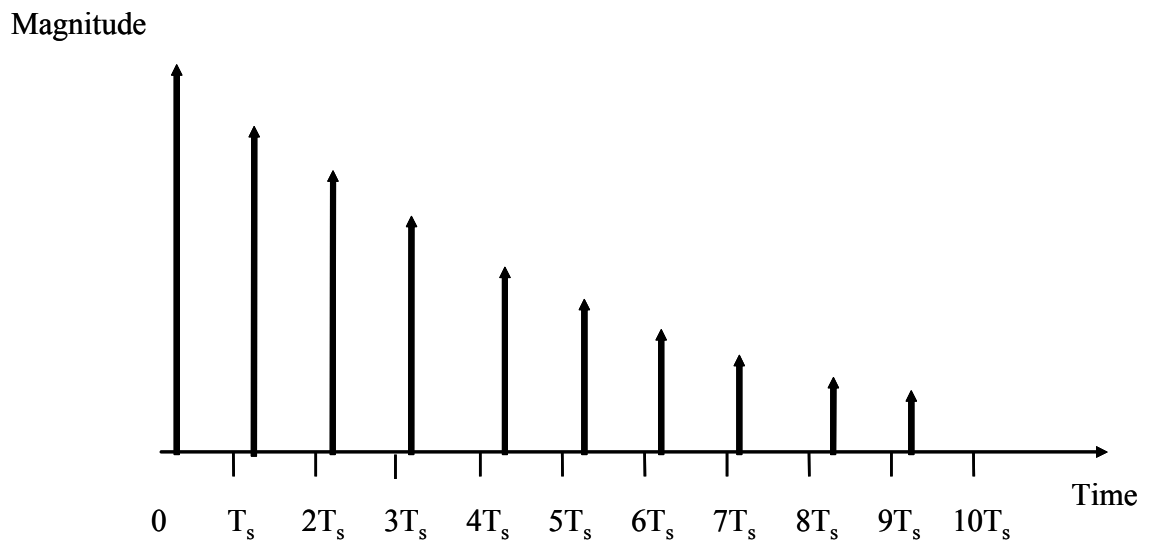
Figure 3.11:   Channel impulse response for IEEE 802.11a.
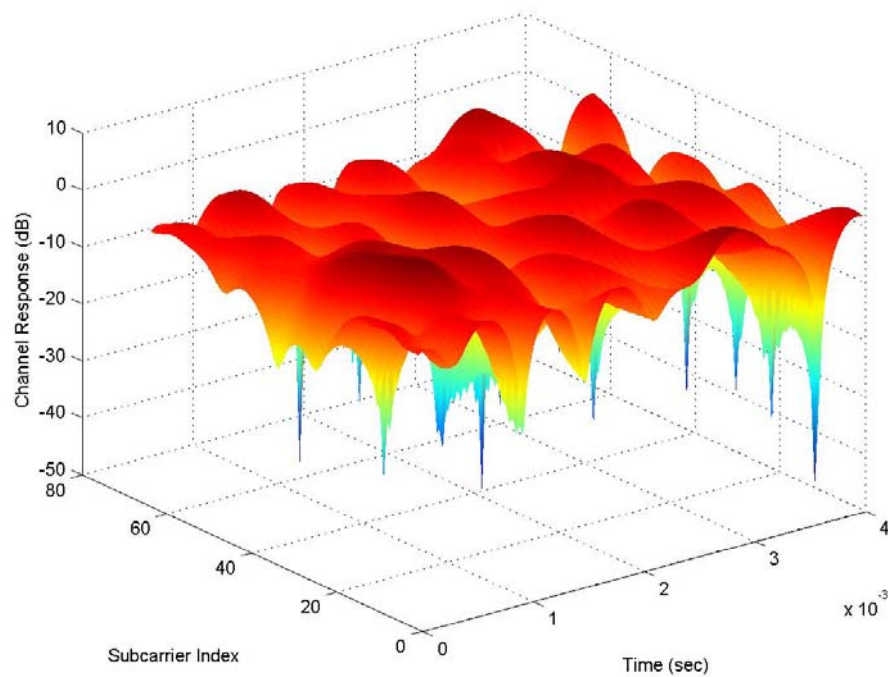
Fig. 3.12:   A typical time-selective and frequency-selective fading channel (assuming an exponential decay channel model with $\tau_{rms} = 50$ ns   and a speed of 60 m/s at 5 GHz).

59

(a)

(b)
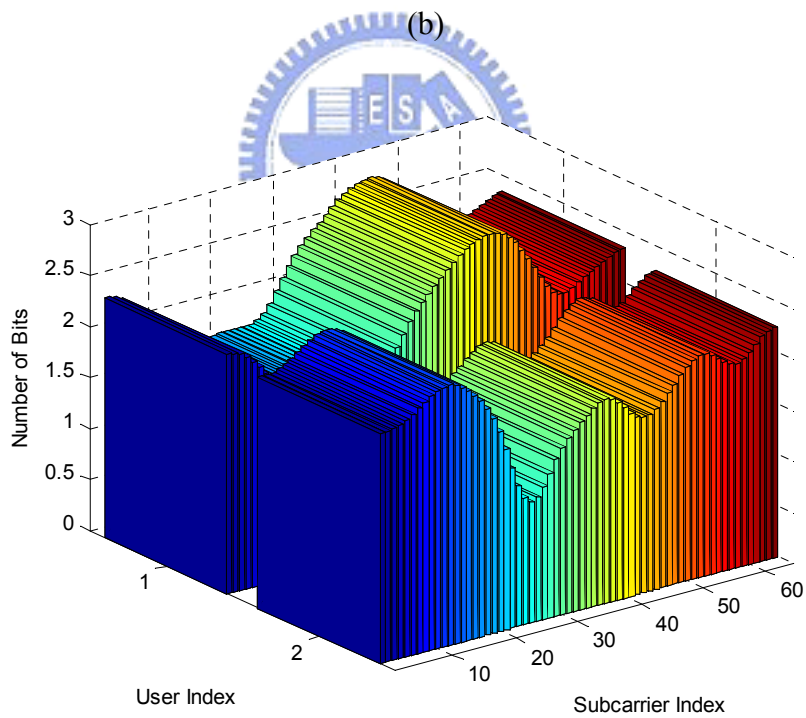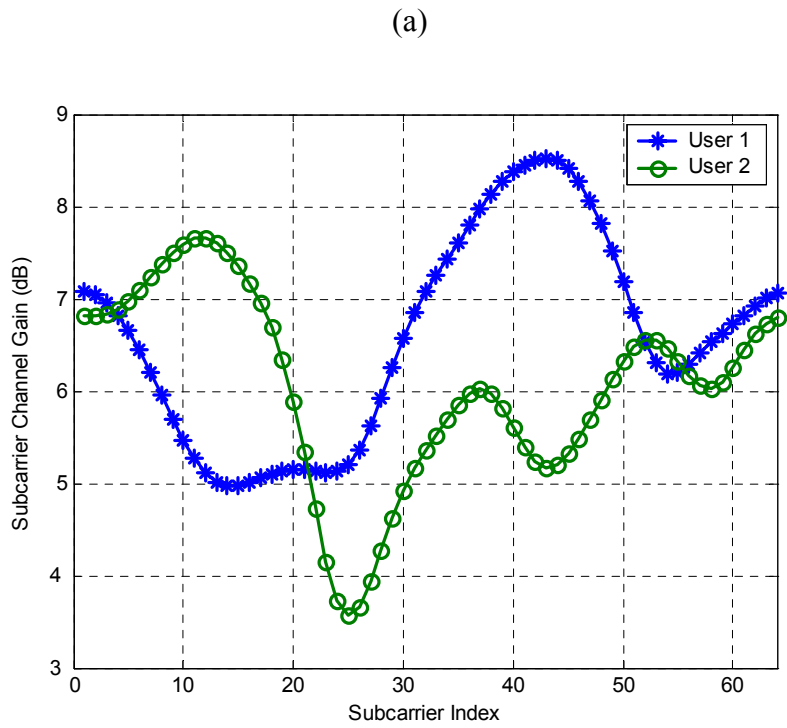
Figure 3.13: Subcarrier channel gains and corresponding number of bits for two uses under the exponential decay Rayleigh fading channel with τ_{rms} =50 ns, and $f_d = 0$ Hz. Other parameters are listed in Table 3.1.
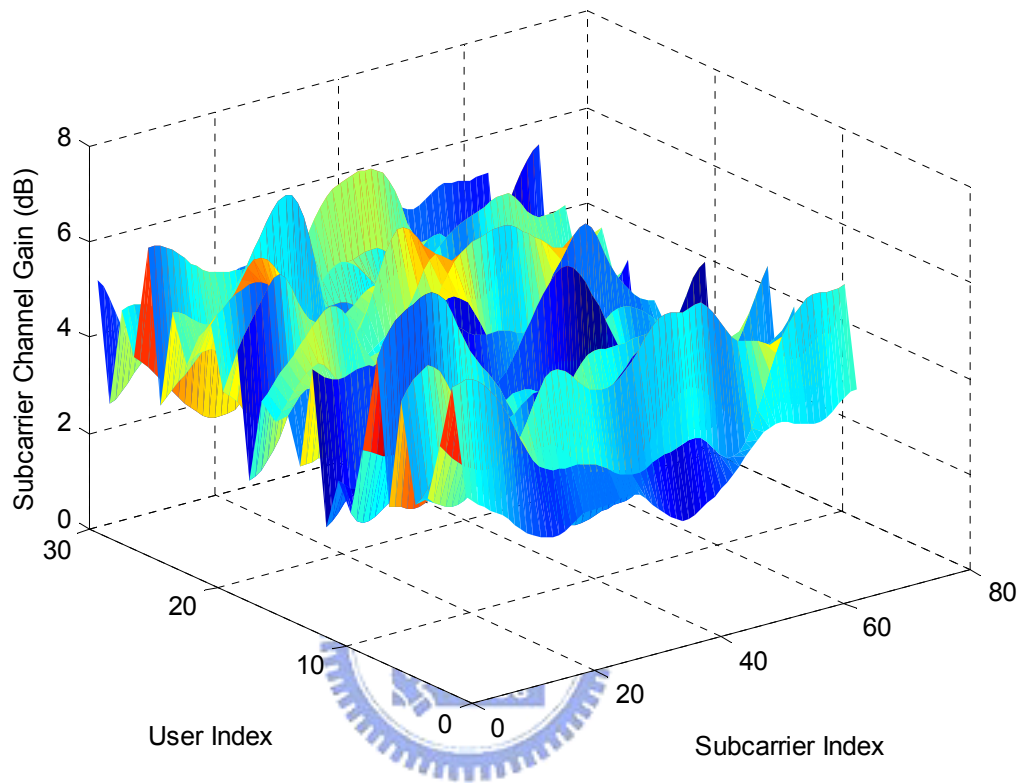
Figure 3.14:   Subcarrier channel gains for thirty users under the exponential decay
                Rayleigh fading channel with τ_{rms} =50 ns, and $f_d = 0$ Hz . Other
                parameters are listed in Table 3.1.
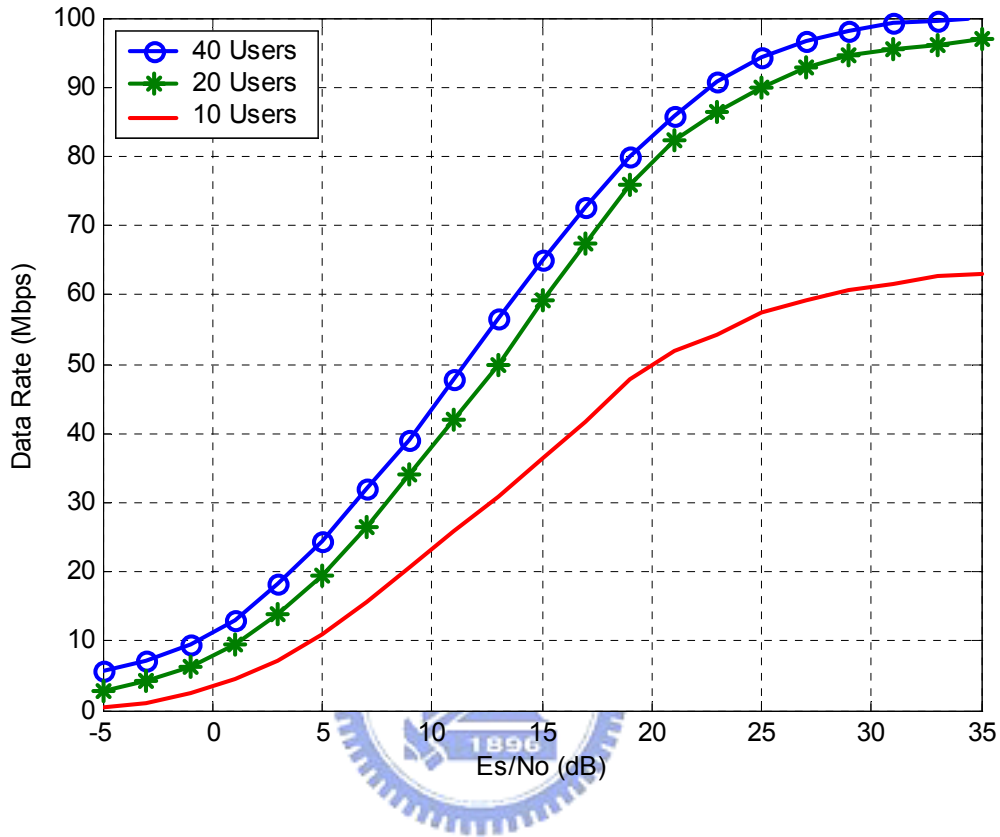
Figure 3.15: Data rate versus E_s/N_0 for the OFDM system with the two-stage subcarrier allocation algorithm under the exponential decay Rayleigh fading channel with τ_{rms} =50 ns, and $f_d = 0$ Hz. The number of users is 10, 20, and 40. Other parameters are listed in Table 3.1.
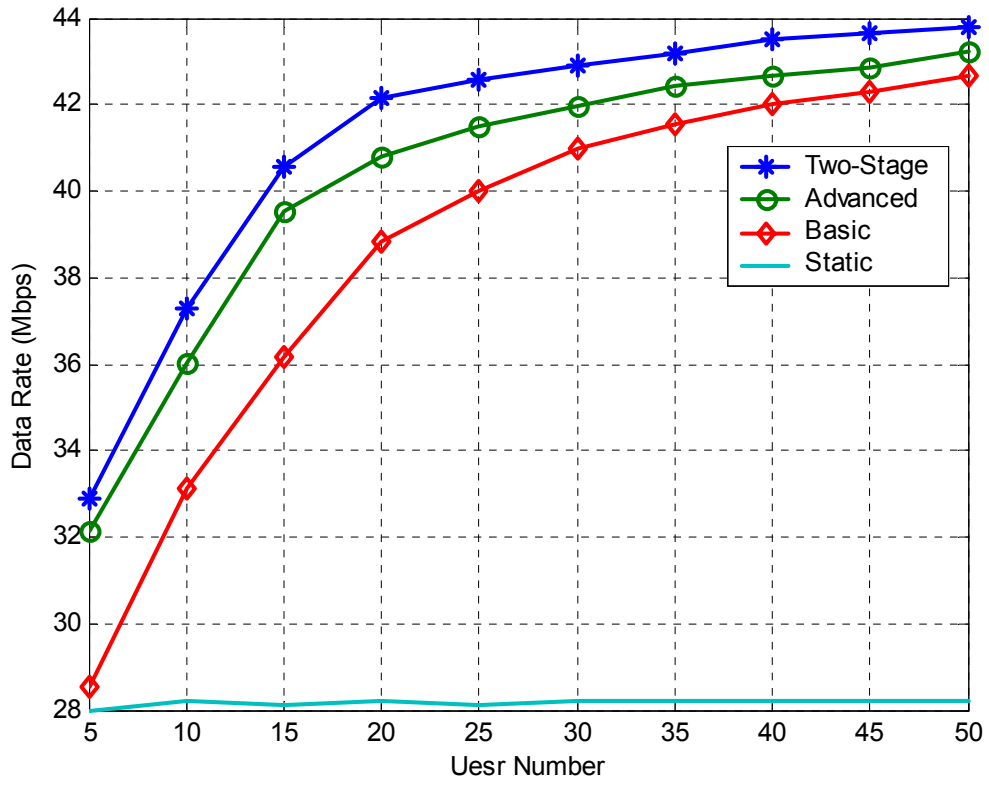
Figure 3.16: Data rate versus number of users for the OFDM system with different subcarrier allocation algorithms under the exponential decay Rayleigh fading channel with τ_{rms} =50 ns, and $f_d = 0$ Hz. Other parameters are listed in Table 3.1.
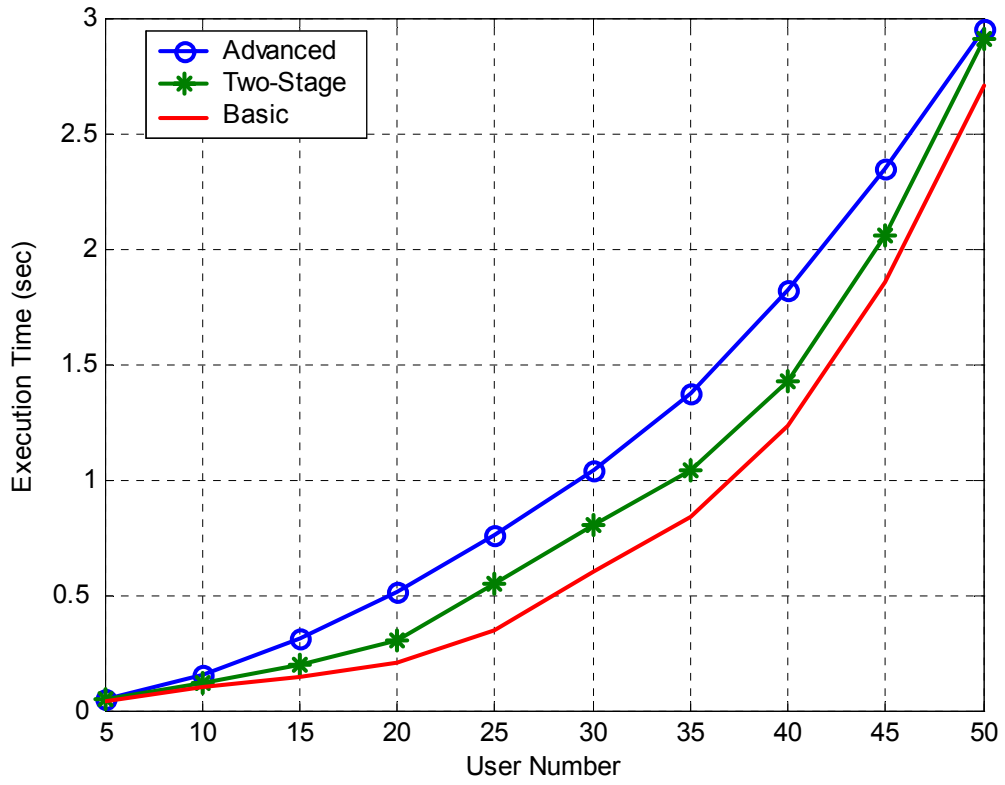
Figure 3.17:   Execution time versus number of users for the OFDM system with different subcarrier allocation algorithms under the exponential decay Rayleigh fading channel with τ_{rms} =50 ns, and $f_d = 0$ Hz . Other parameters are listed in Table 3.1.

Table 3.1: Simulation parameters for the OFDM system.

Number of transmit/receive antennas	1/1
Carrier frequency	5 GHz
Channel bandwidth	20 MHz
Number of carriers	64
OFDM symbol duration	$3.2\,\mu s$
Guard interval	$0.8\,\mu s$
M-QAM available	0, 1, 2, 3, 4, 5, 6
Number of OFDM symbols in a packet	100
Number of users	10, 20, 40
Doppler spread	0 Hz
Channel model	Exponential delay profile, $\tau_{rms} = 50\,n$s

Chapter 4

Multiuser Adaptive MIMO-OFDM Systems

In Chapter 3, different subcarrier allocation schemes are presented to make users choose the most appropriate subcarriers for them according to their QoS requirements and the channel condition. Besides choosing the subcarriers with large channel gains, adaptive modulation is also an important technique to increase data rates. Assuming the transmitter knows the instantaneous channel transfer functions of all users, significant performance improvement can be achieved if adaptive modulation is used in multiuser MIMO-OFDM systems.

In this chapter, a practical adaptive loading algorithm will be introduced in multiuser MIMO-OFDM systems. It uses the V-BLAST as both its channel quality indicator and detection algorithm. Under the total transmit power constraint, this algorithm aims to maximize the data rates and still maintain a target system performance.

4.1   V-BLAST Based OFDM Systems

The OFDM systems combined with the V-BLAST algorithm can dramatically increase the capacity of wireless radio links with no additional power and bandwidth consumption. It can achieve high spectral efficiency to ease the scarcity of radio spectrum. The core idea in such scheme is that with the aid of OFDM, the whole detection problem in MIMO-OFDM systems would be translated into N_c parallel

sub-problems. In addition, the V-BLAST algorithm implements a non-linear detection technique, which is somewhat analogous to the decision feedback equalization to decouple co-channel interference and makes the spatial multiplexing possible. In this scheme, flat fading in the channel bandwidth is required.

Fig. 4.1 illustrates the V-BLAST based MIMO-OFDM transmitter architecture. A traditional 1-D channel encoder is used to encode the information bits. Then these coded bits are mapped on the symbols of constellation adopted for each subcarrier. $N_c \times N_t$ bit streams $\{c_i[n,k]: k = 0,1,\ldots N_c\}$ for $i = 1,2\ldots, N_t$ are fed to the IFFT at the ith transmit antenna on the kth subcarrier to generate the nth transmitted OFDM symbols from the ith transmit antenna at a given time slot n.

Fig. 4.2 illustrates the V-BLAST based MIMO-OFDM receiver architecture. Receive antennas $1{\sim}N_r$ will receive the radiate signal from transmit antennas $1{\sim}N_t$, 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. The output for each frequency subcarrier can be expressed as

$$r_j[n,k] = \sum_{i=1}^{N_t} H_{j,i}[n,k]c_i[n,k] + \eta_j[n,k] \qquad \forall k = 1,2,\ldots,N_c \qquad (4.1)$$

where $H_{i,j}[n,k]$ is the flat fading coefficient representing the channel gain form 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-sided power spectral density $N_0/2$ per dimension and uncorrelated for different n's, k's, and j's.

The use of OFDM allows considering flat fading in the channel bandwidth which is true in each sub-band. Hence at time n, the N_r outputs for the frequency k, $\{r_j[n,k]: j = 1,2,\ldots,N_r\}$, are fed to a V-BLAST component to detect the N_t transmitted signals at the kth subcarrier. This detecting process is repeated for subcarriers $1{\sim}N_c$, and produces N_tN_c estimated symbols, $\{\hat{c}_i[n,k]: i = 1,2,\ldots,N_t; k = 1,2,\ldots,N_c\}$ in the end. These symbols are then multiplexed to the demodulator to complete the traditional 1-D channel decoding.

4.2 Adaptive Modulation for OFDM Systems

Adaptive modulation is a promising technique to increase data rate that can be reliably transmitted over fading channels. The basic idea behind adaptive modulation technique is to adapt the transmission parameters to take advantage of prevailing channel conditions. It aims to exploit the variations of the wireless channel (over time, frequency, and space) by dynamically adjusting certain transmission parameters to the changing environmental and interference conditions observed between the base station and the subscriber. Many parameters are provided that can be adjusted relative to the channel fading, including data rate, transmit power, instantaneous BER, symbol rate, and channel code rate. In practical implementations, the values for the transmission parameters are quantized and grouped together in what referred to as a set of modes. The goal of an adaptive modulation algorithm is to ensure that the most efficient mode is always used, over varying channel conditions, based on a mode selection criterion (minimum transmit power, maximum data rate, etc). Making modes available that enable communication even in poor channel conditions renders the system robust. Under good channel conditions, spectrally efficient modes are alternatively used to increase throughput. In contrast, systems with no adaptive modulation are constrained to use a single mode that is often designed to maintain acceptable performance when the channel quality is poor to get maximum coverage. In other words, these systems are effectively designed for the worst case channel conditions, resulting in insufficient utilization of the full channel capacity. In the following, the general steps of an adaptive modulation system can be presented to react to the change of channel conditions.

1. *Channel quality estimation*: If the channel is reciprocal and the communication between two stations is bi-directional, then the two stations can estimate the channel quality on the basis of received symbols, and adapt the transmission parameters of the local transmitter to this estimation in an open-loop manner. On the other hand, if the channel is not reciprocal, the receiver has to estimate the channel quality and signal explicitly this perceived channel quality to the transmitter via the uplink in a close-loop manner.

2. *Choosing the appropriate transmission parameters for the next transmission*:

According to the prediction of the channel quality for the next time slot, the transmitter has to choose the most appropriate modulation mode for each subchannel.

3. *Signaling of the used parameters*: The information that demodulator parameters to employ for the received packet can be either estimated by a blind detection mechanism at the receiver or conveyed by the transmitted signal itself.

The practical challenges of adaptive modulation technique can be described as follows:

1. *Adaptation rate*: In the slow varying channels, a low adaptation rate adaptive modulation algorithm is required to track the large variations influenced by user location within the cell, seasons, road traffic, and cell deployment. On the other hand, a high adaptation rate adaptive modulation algorithm applied in the fast varying channels is required to track the small variations influenced by the time-frequency selective fading channels. It is easily understand that faster adaptation leads to larger capacity gain, since the channel variations are exploited in a more accurate manner. However, fast adaptation has practical limitations, in both time and frequency. Fast adaptation increases the number of mode change messages that must be sent over the air, consuming bandwidth and time resources.

2. *Feedback*: The feedback load should be minimized since it consumes resources that would be otherwise used for data. In general, the mode selection is made at the receiver, and only the mode is fed back to the transmitter. It is more efficient since there are typically a small number of modes, thereby limiting the amount of feedback. Furthermore, the mode isn't fed back at each prediction, but rather only if the new predicted mode differs from the mode the transmitter is currently using.

3. *Determination of adaptation thresholds*: The thresholds can be obtained by either measurements or simulation. It is often a challenge to compute them accurately because the required sample size grows as the order of the statistical average increases. Therefore, the trade-off in determining the thresholds consists of picking the least amount statistical channel information while still describing the essence of the channel behavior.

4.3 Switching Levels in Adaptive Modulation for OFDM Systems

There are various metrics that may be used as CSI. Typically, BERs are available via packet error rates (PERs) at the link layer, which are normally extracted from the cyclic redundancy check (CRC) information. At the physical layer, SNR or SINR could be available by exploiting power measurement in slots without intended transmit data. These two types of CSI described above all come with their pros and cons. Error based CSI captures accurate performance of the modes on the observed link quality for each possible mode candidate instead of relying on theoretical BER curves. However, the limitation comes with the number of packets observed in any window. The method relies on the estimation of PER statistics, which can require up to several thousands of packets to be transmitted for a given mode in order to obtain a reliable estimate, thereby making the adaptation loop slow. SNR-based CSI offers the flexibility to adapt the modes on a very fast basis. However, it relies on the computation of adaptation/switching thresholds that may be inaccurate. The accuracy of the threshold mechanism increases by taking into account higher order statistics of the SNR than just the mean. A typical possible solution for adaptive modulation can be described as follows:

1. Measure the SNR at the receiver.

2. Convert the SNR information into BER information for each mode candidate.

3. Based on a target BER, select for each SNR measurement the mode that yields the largest throughput while remaining within the BER target bounds.

4. Feed back the select mode to the transmitter.

Step 1 corresponds to the assessment of the CSI. Step 2 refers to the computation of the switching thresholds. In this case, a threshold is defined as the minimum required SNR for a given mode to operate at a given target BER. Step 3 refers to the selection of the optimal mode, based on a set of thresholds and SNR measurement. Step 4 is concerned with the feedback of information to the transmitter. Fig 4.3 represents a set of BER curves for different modulations such as BPSK, QPSK, 8-QAM, 16-QAM, 32-QAM, and 64-QAM. The BER curves of each constituent modulation mode obtained from

simulations over an AWGN channel can be used to find the SNR value where each modulation mode satisfies the target BER requirement. By using this approach, the instantaneous BER always remains below a certain threshold BER ε_{error}. In order to satisfy this constraint, the first modulation mode should be "no transmission". Then, the set of switching levels **T** can be expressed as

$$\mathbf{T} = \left\{ t_0 = -\infty, \ t_k \mid \varepsilon_{m_k}(t_k) = \varepsilon_{error}, \ \forall k \geq 1 \right\} \tag{4.2}$$

where $\varepsilon_{m_k}(\gamma)$ is the BER of the m_k-ary constituent modulation mode over the AWGN channel when the instantaneous SNR is presented as γ. In the following discussions, this approach will be employed and the target BER is defined as 10^{-4} which is usually used for low-BER data transmission systems. According to the target BER, the SNR thresholds are 8.4, 11.4, 15.5, 18.5, 21.6, and 28.3 dB respectively which can be found in Table 4.1. They are independent of the average SNR and the underlying fading.

4.4   Adaptive MIMO-OFDM Systems

Transmitter designs adapted to the intended propagation channel are capable of improving both performance and rate of communication links. The resulting channel adaptive transmission adjust transmission parameters such as power levels, constellation sizes, coding schemes, and modulation types, depending on the CSI that is assumed available to the transmitter. The potential improvement increases considerably when multiple transmit and receive antennas are deployed. In multiuser adaptive MIMO-OFDM systems, the optimization problem can be defined as allocating the appropriate subcarriers to the users and determining the number of bits and the power level transmitted on each subchannel based on the instantaneous fading characteristics of all users to maximize data rate under the total transmit power constraint and a target BER requirement. Hence, the V-BLAST based MIMO-OFDM system is employed because it can dramatically increase the capacity of wireless radio links with no additional power and bandwidth consumption and a practical adaptive loading algorithm is presented to fully integrate temporal, spatial, and spectral

components together. Besides, the V-BLAST algorithm is modified and its weight vectors are used to express the channel quality of a MIMO system in terms of *post-processing* SNR as in Equation (2.38) which are differ form simple processing SNR levels measured at the receive antennas. The variation of post-processing SNR is monitored all the time over space, time, and frequency, thus enabling the adaptation procedure to fully exploit all these three dimensions. By treating each OFDM subcarrier as a narrow band MIMO, the adaptive loading algorithm can separate the joint space-frequency optimization problem into two separated stages. It is the main feature of this adaptive loading algorithm.

4.4.1 System Architecture

Fig 4.5 illustrates the system architecture of the V-BLAST based multiuser adaptive MIMO-OFDM system. It is assumed that F users in this system and the fth user has a data rate equal to R_f bits per OFDM symbol. In the transmitter, the serial data from the F users are fed into the subcarrier allocation block which assigns different subcarriers to different users according to their QoS requirements and the CSI information. Then the adaptive loading algorithm assigns the number of bits per OFDM symbol to be transmitted on each subchannel. Depending on the number of bits assigned to each subchannel, the adaptive modulator will use a corresponding modulation scheme, and the transmit power level will be adjusted according to the adaptive loading algorithm. Moreover, the subcarrier and bit allocation information is assumed to send to the receiver via a separate control channel. In the receiver, the bit allocation information is used to configure the demodulators while the subcarrier allocation information is used to extract the demodulated bits form the subcarriers assigned to the fth user.

Fig. 4.6 illustrates the transmitter architecture of the V-BLAST based adaptive MIMO-OFDM system. At a given time slot n, there are $N_c \times N_t$ bit streams, where N_c is the number of subcarriers and N_t is the number of spatial channels per subcarrier. These bit streams can be expressed as $\{\mathbf{b}_i[n,k], \ k = 0, 1, \ldots, N_c\}$ for $i = 1, 2, \ldots, N_t$ which are encoded separately in different puncturing rates by the encoder into $N_c \times N_t$ coded bit streams. The coded bit streams can be presented as

$\{\mathbf{c}_i[n,k], \quad k=0,1,\ldots,N_c\}$ for $i=1,2\ldots,N_t$, respectively and then the modulator modulates $\mathbf{c}_i[n,k]$ into QAM symbols that can be expressed as $\{t_i[n,k], \quad k=0,1,\ldots,N_c\}$ for $i=1,2\ldots,N_t$. Besides, the adaptive loading processor at the transmitter informs the length of $\mathbf{b}_i[n,k]$, coding rate of $\mathbf{c}_i[n,k]$, and the modulation level of $t_i[n,k]$ according to the receiving CSI information. The CSI information can be transmitted either from feedback signaling or the channel estimator. Then the nth OFDM symbols at the ith transmit antenna is generated by passing through an IFFT which can be written as $\{t_i[n,k], \quad k=0,1,\ldots,N_c\}$. A CP will also be added before transmission to ensure orthogonality between the subcarriers, provided that the maximum time dispersion is less than the CP.

Fig. 4.7 illustrates the receiver architecture of the V-BLAST based adaptive MIMO-OFDM system. At the receiver, receive antennas $1 \sim N_r$ will receive the radiated signal from transmit antennas $1 \sim N_t$. It is assumed that the number of receive antennas N_r must be larger than the number of transmit antennas N_t. Then the received data at each receive antenna will pass through a FFT with the removal of the CP. The CP is removed to eliminate the effects of ISI and ICI. The FFT output signal can be expressed as

$$r_j[n,k] = \sum_{i=1}^{N_t} H_{i,j}[n,k]t_i[n,k] + \eta_j[n,k] \qquad \forall j = 1,2,\ldots,N_r \tag{4.3}$$

where $H_{i,j}[n,k]$ is the channel frequency response corresponding to the ith transmit and jth receive antenna for the $\underline{k}$th subcarrier at the time slot n and $\eta_j[n,k]$ is the additive complex Gaussian noise at the jth receive antenna. The additive complex Gaussian noise assumes to be uncorrelated for different n's, k's, or j's and zero-mean with two-sided power spectral density $N_0/2$ per dimension. These spectrally mixed data at the receiver are further processed by the V-BLAST algorithm as presented in Section 4.1 except that the modulation order and coding rate used in each subchannel are not fixed. Besides, due to the use of adaptive modulation, the ordering strategy used in the original V-LABST algorithm is unnecessary because those ill channels are assigned more robust transmission modes to eliminate error bursts and the conventional interleaving technique is also unessential by the same reason.

4.4.2 Two-Stage Bit Loading Algorithm

The capacity of a MIMO system is maximized by using simple singular value decomposition (SVD) weights combined with optimal power distribution over the transmit antennas which is known as the water filling (WF) algorithm. However, the WF algorithm requires infinite length codebook, continuous modulation order, and continuous power level. Thus, it isn't possible to use the result directly in practice. A group of bit loading algorithms were proposed to approach the WF solution with the constraints of discrete modulation order. These bit loading algorithms distribute the available energy among a set of parallel AWGN channels as to maximize the overall bit rate for discrete loading problem in practice [34]-[37]. Therefore, on account of simplicity and capability, the Campllo's loading criteria [35] is promoted that could be somewhat modified and extended to the V-BLAST based multiuser adaptive MIMO-OFDM system with reasonable computation complexity.

The modified ZF V-BLAST detection algorithm can generate a set of weighting vectors $\mathbf{w}_{k_i}$ to perform spatial multiplexing and calculate the post-processing SNRs as in Equation (2.38). By simulation, the norm values of the weighting vectors are observed to change in an unpredictable manner when some transmit antennas turn off. This undesirable result due to the nonlinear operations in the V-BLAST algorithm means that once some transmit antennas are chosen to turn off, the norm values will be recalculated according to the changed channel matrix. By this reason, an exhaustive search over all possible combinations of transmit antennas is required to find the optimal one subject to the constraint in Equation (4.4). By properly exploiting causality, the exhaustive search can be significantly mitigated.

The joint space-frequency bit loading problem in the V-BLAST based MIMO-OFDM system can be described as

$$\max \sum_{i=1}^{N_t} \sum_{k=1}^{N_c} b_i[k] \quad \text{subject to} \quad \begin{cases} \sum_{i=1}^{N_t} \sum_{k=1}^{N_c} P_i[k] \le P_{budget} \\ \varepsilon_i[k] \le \varepsilon_{error} \quad \forall i,k \end{cases} \tag{4.4}$$

where $b_i[k]$, $P_i[k]$, and $\varepsilon_i[k]$ are the rate, allocated power, and error probability of the ith transmit antenna at the kth subcarrier respectively,, ε_{error} is the target BER, and

P_{budget} is the total power constraint. P_{budget} is normalized to 1 in the simulation to guarantee a fair comparison between systems equipped with different transmit antennas. Considering the above observation, the joint loading problem should be taken apart into two separated subproblems by the following reasons:

1. In order to avoid the unpredictable manner introduced by the V-BLAST detection algorithm, the active subchannels should be predetermined before a full search over all subchannels ($N_c \times N_t$).

2. The sorting complexity is significantly reduced by taking the joint bit loading problem apart into two smaller ones.

Therefore, the adaptive loading algorithm is applied to each subcarrier to obtain an optimal bit and power allocation over its N_t spatial channels at the first stage. At the second stage, the same loading algorithm is used over those active subchannels surviving from the first stage (at most $N_t \times N_c$). Through this two-stage processing, each subchannel's condition will be precisely monitored and the unpredictable phenomenon happened in the V-BLAST detection algorithm won't have to be worried about. The adaptive two-stage algorithm can be presented as follows:

Stage 1:

For each subband k containing N_t spatial channels, the allocation problem can be stated as follows:

$$\max \sum_{i=1}^{N_t} b_i[k] \quad \text{subject to} \quad \begin{cases} \sum_{i=1}^{N_t} P_i[k] \le P_{distributed}[k] \\ \varepsilon_i[k] \le \varepsilon_{error} \quad \forall i \end{cases} \tag{4.5}$$

where $P_{distributed}[k]$ is the distributed power at the kth subcarrier. It can be decided according to the ratio of the kth subcarrier channel gain and the sum of total subcarriers' channel gains which would be presented as

$$P_{distributed}[k] = \frac{|H_k|^2}{\sum_{k=1}^{N_c} |H_k|^2} \cdot P_{budget} \tag{4.6}$$

When considering the practicability, the rate *b* should be restricted to an integer number. Nevertheless, systems usually use a specified channel encoder along with different puncturing rate to make the rate *b* equivalent to some fraction numbers. For instance, if a convolution encoder is used to encode a sequence of source bits and puncture output bits to rate 2/3, and then use 16-QAM as modulation order, the information rate defined as *b* will be equivalent to 8/3. To assist the following statements, the case without channel coding will be described.

Initialization

Step 1:

Defined *q* as the state of the N_t transmit antennas according to their active modes. For instance, assuming that four antennas are available at the transmitter side, two of them are selected, e.g. the 1st and 3rd antennas to be active, the state *q* will become 10, which is the result of converting the corresponding active mode vector [1, 0, 1, 0] to a decimal number. Let $\text{Rate} = 0$ and $P_{residual\_final} = P_{distributed}[k]$ at first.

Step 2:

Load equal power on those active transmit antennas $N_{t,active}[k]$ which can be stated as

$$P_{i\_active}[k] = \frac{P_{distributed}[k]}{N_{t\_active}[k]} \quad \forall 1 \le i\_active \le N_{t,active} \tag{4.7}$$

Step 3:

Calculate the post-processing SNR of each active layer according to Equation (2.38)

$$\rho_{i\_active}[k] = \frac{P_{i\_active}[k]}{\sigma^2 \| \mathbf{w}_{i\_active}[k] \|^2} \tag{4.8}$$

Step 4:

Clip the power of each layer to reduce its SNR in order to fit the nearest threshold below it by consulting the threshold table in table 4.1, and collect the residual power

$$P_{i\_active}[k] \leftarrow \frac{P_{i\_active}[k] \times 10^{(t_k/10)}}{\rho_{i\_active}[k]} \tag{4.9}$$

$$P_{residual}[k] = P_{budget} - \sum_{i=1}^{N_{t,active}[k]} P_{i,active}[k] \tag{4.10}$$

The Power-Tighten Algorithm

A bit distribution is said to be power tighten if

$$0 \le P_{budget} - \sum_{i=1}^{N_t} P_i \le \min_{1 \le i \le N}[\Delta p_i(b_i + 1)] \tag{4.11}$$

The Power-Tighten algorithm can be described as

Step 1:

Since the power-tighten algorithm would be employed for all subcarriers, the index k will be dropped for simplicity in following expressions. Assumed that $\Delta p_i(b_i)$ is the power required for ith layer to increase rate from $b_i - 1$ to b_i. It can be defined as

$$\Delta p_i(b_i) = \begin{cases} P_i \times 10^{\frac{(t_{b_i} - t_{b_i-1})}{10}} - P_i, & b_i > 1 \\ 10^{\frac{t_{b_i}}{10}}, & b_i = 1 \end{cases} \tag{4.12}$$

A power increment table will be constructed to contain $\Delta p_i(b_i)$ and $\Delta p_i(b_i + 1)$ for $i = 1, 2 \ldots, N_{t,active}$ to record the least amount of power needed to step from current transmission mode into the next high rate mode.

Step 2:

Find the active transmit antenna m requiring the least amount of power to step into the next high rate mode which can be defined as

$$m = \arg\left\{\min_{1 \le i \le N_{t,active}}[\Delta p_i(b_i + 1)]\right\} \tag{4.13}$$

Step 3:

While $p_{residual} > \Delta p_m(b_m + 1)$, the following substeps would be executed:

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

The Power-Efficientizing Algorithm

A bit distribution is said to be power efficient if

$$\max_i[\Delta p_i(b_i)] \le \min_i[\Delta p_i(b_i+1)] \tag{4.14}$$

The Power-Efficientizing algorithm can be described as

Step 1:

Find the active transmit antenna m which requires the least amount of power to step into the next higher rate mode and the active transmit antenna n which releases the most amount of power to go back for the next lower rate mode. Both of them can be defined as

$$m = \arg\left\{\min_{1\le i\le N_{t,active}} [\Delta p_i(b_i+1)]\right\}, \text{ and } n = \arg\left\{\max_{1\le i\le N_{t,active}} [\Delta p_i(b_i)]\right\} \tag{4.15}$$

Step 2:

If $n = m$, the following substeps would be carry out:

(a) $l = \arg\left\{\min_{1\le i\ne m\le N_{t,active}} [\Delta p_i(b_i+1)]\right\}$, and $j = \arg\left\{\max_{1\le i\ne n\le N_{t,active}} [\Delta p_i(b_i)]\right\}$

(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 $n \leftarrow j$

Step 3:

While $\Delta p_n(b_n) > \Delta p_m(b_m+1)$, the following substeps would be performed:

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

Comparing and Recording

Step 1:

If all the predetermined active layers are assumed to remain surviving, the following process will be executed:

78

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

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

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

      else $q \leftarrow q-1$

    else $q \leftarrow q-1$

else $q \leftarrow q-1$

In this stage, a time-consuming exhaustive search is performed to determine which transmit antennas would be active to support the optimal bit and power allocation for every subcarrier. Fortunately, the needed effort can be reduced due to the causality between every possible combination. Fig. 4.8 shows the flow chart of the first stage adaptive bit loading algorithm.

Stage 2:

   In this stage, the Power-Tighten algorithm and the Power-Efficientizing algorithm are reused for those surviving sub-channels from the first stage including both frequency and spatial channels to further exhaust the total residual power to get a rate enhancement.

The Power-Tighten Algorithm

The Power-Tighten algorithm would be employed according to the principle:

$$0 \leq P_{budget} - \sum_{i=1}^{N_t} \sum_{k=1}^{N_c} P_i[k] \leq \min_{1 \leq i \leq N_t; 1 \leq k \leq N_c} [\Delta p_i[k](b_i[k]+1)] \tag{4.16}$$

The Power-Efficientizing Algorithm

The Power-Efficientizing algorithm would be employed according to the principle:

$$\max_{i;k}[p_i[k](b_i[k])] \leq \min_{i;k}[p_i[k](b_i[k]+1)] \tag{4.17}$$

4.5   Computer Simulations

Computer simulations are conducted to evaluate the performance of the proposed V-BLAST based adaptive MIMO-OFDM system in this section. Channel estimation and timing synchronization are assumed to be perfect at first. The discrete time signal processing in the baseband is only dealt with throughout the simulation. Hence, pulse-shaping and matched-filtering are removed from consideration for simulation simplicity. Table 4.2 lists all parameters used in the following simulation. The configuration here is a MIMO-OFDM system with a bandwidth of 20 MHz and 64 subcarriers. The set of QAM constellation used in the simulation is {0, 2, 4, 8, 16, 32, and 64}. Each link in MIMO is modeled as an exponential decay Rayleigh fading channel with $\tau_{rms} = 50\ ns$. The IEEE 802.11 Working Group suggests this channel profile as the baseline for predicting multipath in IEEE 802.11a (5 GHz). This model is ideal for software simulations in predicting performance results of a given implementation. The channel impulse response is illustrated in Fig. 3.11.

Before the simulation, a look-up table that contains the SNR threshold values of each modulation mode should be established at first. By consulting each BER curve to find the corresponding SNR vale that meets the BER requirement (10^{-4} in our simulations), these threshold values could be obtained from Fig. 4.3.

Fig. 4.9 shows the selection probability of each modulation mode at different E_s / N_0 without considering the mobility issue. From the figure, higher order modulation modes are preferred as E_s / N_0 increases. In the low E_s / N_0 scenarios, the adaptive loading algorithm gives up adopting higher order modulation modes and turns to choose lower order modulation modes to meet the target BER requirement. The algorithm would force some of transmit antennas to be blocked frequently to avoid inefficient or unreliable transmission.

Fig. 4.10 and Fig 4.11 shows BER versus E_s / N_0 with different detection criteria at different number of transmit and receive antennas. It is obvious that the ZF method always meets the BER requirement and the MMSE method exhibits performance degradation at low E_s / N_0 scenarios. By using the MMSE criterion, the actual post-processing SNR in Fig 4.3 doesn't hold due to the bias of the signal component in the soft decision value. This is the reason why the MMSE method exhibits performance

degradation at low E_s/N_0 scenarios. However, at high E_s/N_0 scenarios, the MMSE method will approximate the ZF one and reaches the same performance.

Fig 4.12 shows BER versus E_s/N_0 with ZF detection criteria at different number of transmit and receive antennas. The difference between Fig 4.10 and Fig 4.12 is that the system using residual power not only meets the BER requirement but also achieves a better BER performance. At high E_s/N_0 scenarios, most subchannels are fully loaded with little power consumption, hence there is extra power remaining. The residual power could be effectively used to achieve a significant BER improvement by uniformly assigning them to those active sub-channels.

In Fig 4.13 and Fig 4.14, the modulation mode selection probability and the transmission rate at different E_s/N_0 scenarios are compared for the two cases: space loading (without the second stage) and space-frequency loading (with the second stage). The first stage is performed to select the active subchannels while the second stage is done over the subchannels surviving from the first stage to make a further use of the residual power. From Fig 4.13, it is obvious that the space-frequency loading will gain a higher peak in each modulation type at the second stage and lead to a rate enhancement which can be shown in Fig 4.14.

Fig 4.15 and Fig 4.16 show the modulation mode selection probability and the transmission rate at different E_s/N_0 scenarios. Both of them are compared with different number of transmit and receive antennas. The diversity gain in V-BLAST algorithm increases from an order of $N_r - N_t + 1$ to N_r as the effective number of transmit signals decreases at each step of the SIC. Therefore, the systems with more receive antennas can always extract more diversity than those with fewer ones by assuming no error propagation problem. In the V-BLAST based adaptive MIMO-OFDM system, diversity gain can allow for the use of higher order modulation without degrading the BER performance. This feature proves why the transmission rate curve corresponding to the case of $(N_t, N_r) = (5, 5)$ is steeper than the other curves in Fig. 4.16.

Fig 4.17 shows the unutilized power ratio after the two-stage adaptive loading algorithm. From the figure, it is evident that much power remains unused in both the low and high E_s/N_0 scenarios. The reason is that most subchannels suffer from ill

channel conditions to force them turning off to save power at low E_s/N_0 scenarios. On the contrary, at high E_s/N_0 scenarios, most subchannels are employed higher order modulation modes with little power consumption, and there is extra power remaining. This residual power could be saved to reduce the total transmit power or effectively used to achieve a higher performance margin.

Fig 4.18 shows the transmission bit rate versus E_s/N_0 with different loading algorithms including the rounding off WF algorithm, the QoS based WF algorithm [38], and the Chow's algorithm. From the figure, it can be shown that the two-stage adaptive loading algorithm has the best performance than other loading algorithms. It is because that this algorithm uses the Power-Tighten algorithm and the Power- Efficientzing algorithm to achieve the maximum transmission rate efficiently.

Fig 4.19 shows BER versus E_s/N_0 with different detection criteria. In this simulation, a constant channel estimation error is assumed. The channel estimation error ΔH is defined to be equal to the noise power. From the figure, both methods exhibit significant performance degradation. This is because that the channel estimation error makes the chosen modulation mode not optimum with regards to the actual channel quality and hence degrades the BER performance.

In Fig 4.20, the proposed system is simulated in a realistic TDD system. In TDD system, the estimation of the channel quality priori to transmission is used to select the appropriate modulation mode for the next transmission. Hence a channel quality estimation delay is incurred in this scheme. During this delay, the fading channel quality varies according to the Doppler frequency and consequently, the channel quality estimates perceived priori to transmission may become antiquated. A delay time of one packet (0.4 ms) is assumed in this simulation. From the figure, it is obvious that the proposed system works well to meet the BER constraint. The proposed system only exhibits slight performance degradation even when the user moves with a high speed (120 m/s). The reason is that the estimation delay time is usually smaller than the channel coherence time. During the channel coherence time, the channel condition wouldn't change severely. Therefore the chosen modulation mode would reflect to the current channel quality and doesn't degrade the BER performance.

4.6   Summary

Due to the scarcity of radio spectrum, high spectral efficiency becomes a need 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. In Section 4.1, the V-BLAST based MIMO-OFDM system is introduced.

Adaptive modulation is also a promising technique to increase data rate that can be reliably transmitted over fading channels. This technique aims to exploit the variations of the wireless channel (over time, frequency, and space) by dynamically adjusting certain transmission parameters to the changing environmental and interference conditions observed between the base station and the subscriber which is introduced in Section 4.2 and Section 4.3.

From the analysis of MIMO channel capacity, the waterfilling distribution of power over channels with different SNR values achieves the optimal transmission scheme. 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. Therefore, a practical two-stage adaptive loading algorithm is presented 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. This two-stage adaptive loading algorithm is introduced in Section 4.4.
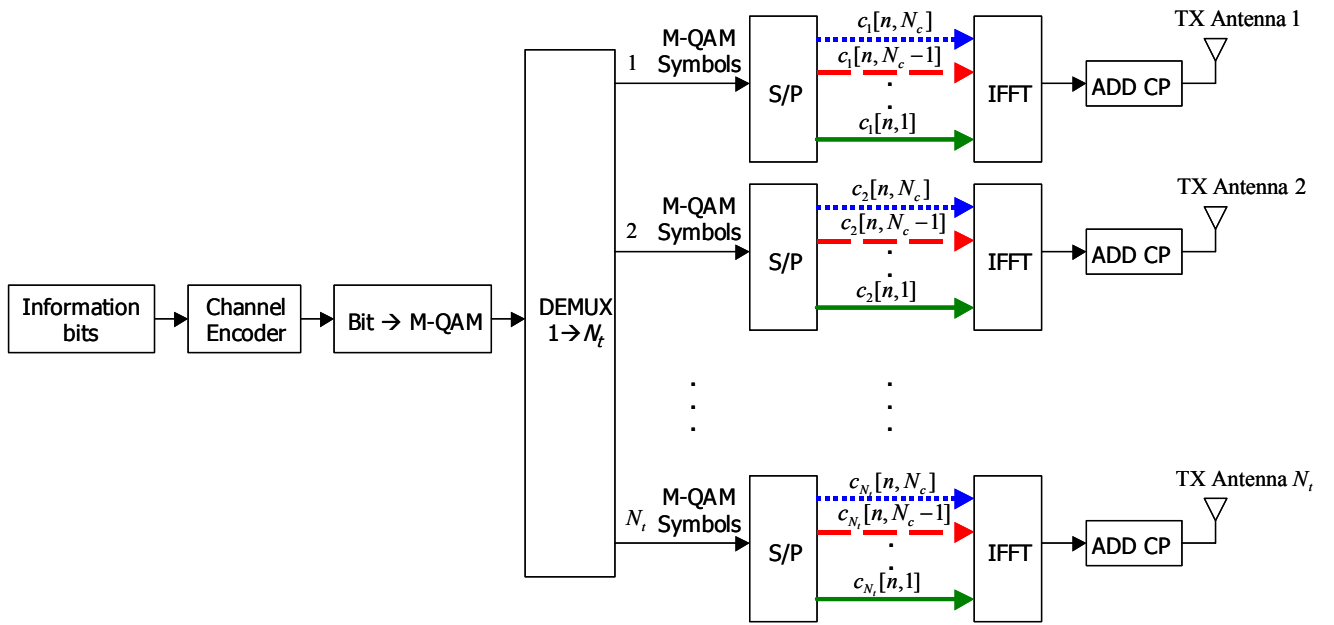
Figure 4.1: V-BLAST based MIMO-OFDM transmitter architecture.
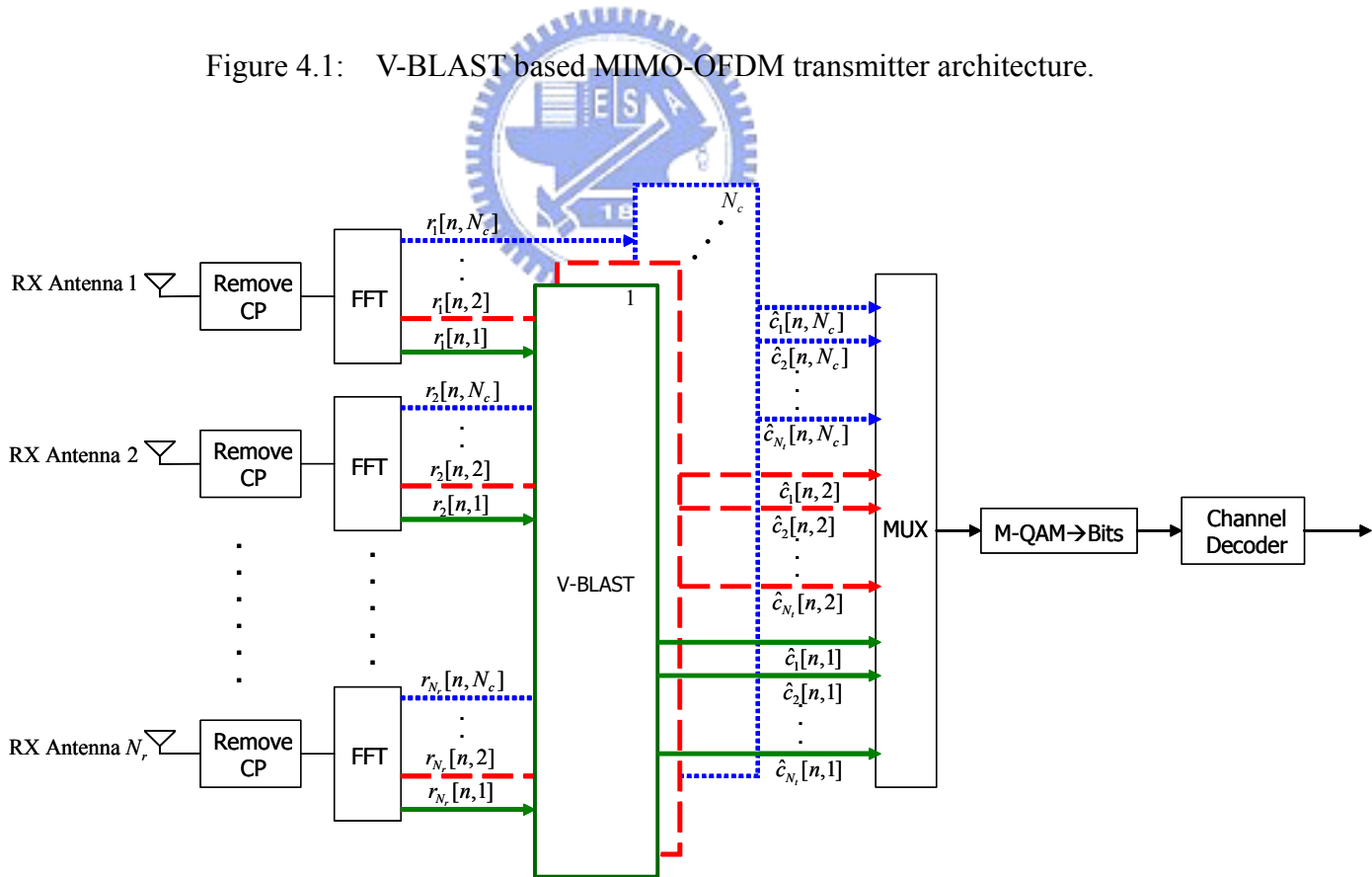

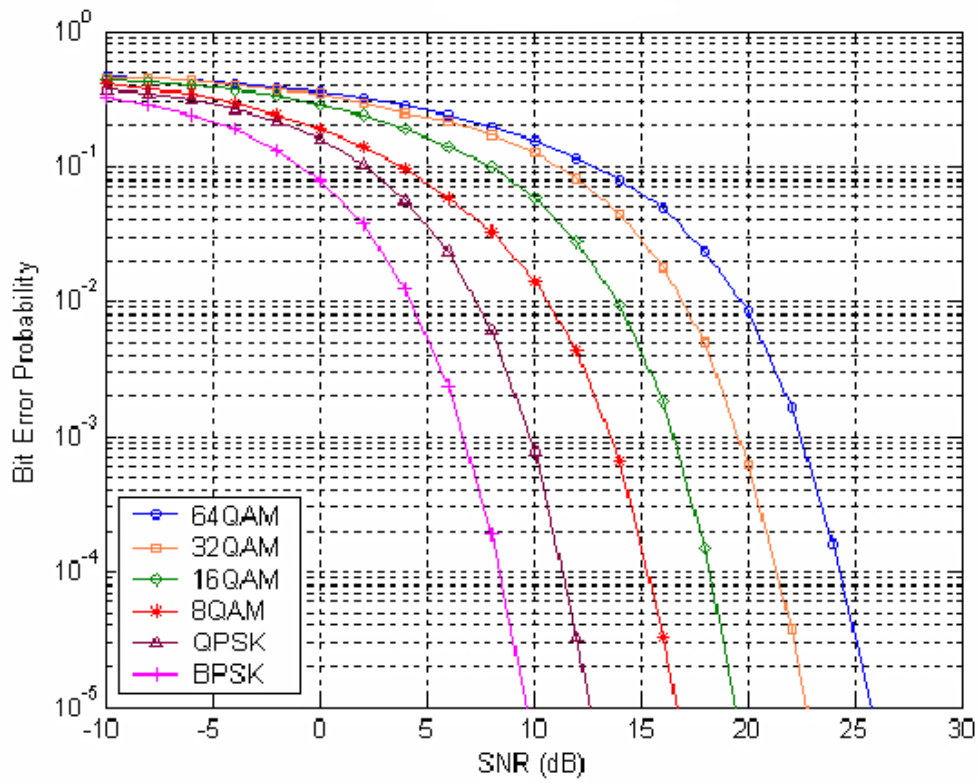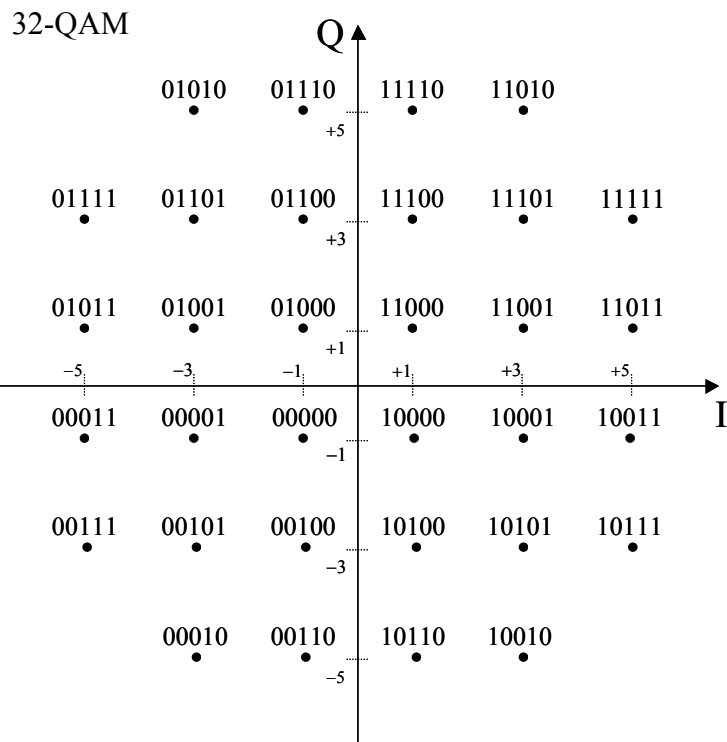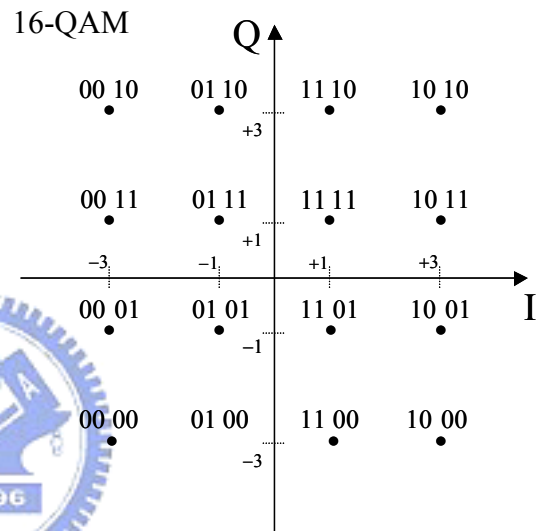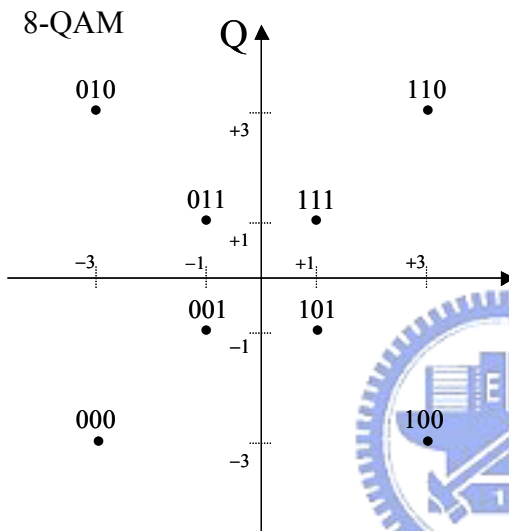Figure 4.2: V-BLAST based MIMO-OFDM receiver architecture.

Figure 4.3:   The average BER of various M-QAM modulation schemes over Rayleigh
                fading channel.

BPSK

Q

−1   +1
0   1   I

QPSK

Q

01   11
+1
−1   +1
00   10   I
−1

8-QAM

Q

010   110
+3
011   111
+1
−3   −1   +1   +3
001   101   I
−1
000
−3

16-QAM

Q

00 10   01 10   11 10   10 10
+3
00 11   01 11   11 11   10 11
+1
−3   −1   +1   +3
00 01   01 01   11 01   10 01   I
−1
00 00   01 00   11 00   10 00
−3

32-QAM

Q

01010   01110   11110   11010
+5
01111   01101   01100   11100   11101   11111
+3
01011   01001   01000   11000   11001   11011
+1
−5   −3   −1   +1   +3   +5
00011   00001   00000   10000   10001   10011   I
−1
00111   00101   00100   10100   10101   10111
−3
00010   00110   10110   10010
−5

64-QAM

Q

000 100   001 100   011 100   010 100 | 110 100   111 100   101 100   100 100
$+7$

000 101   001 101   011 101   010 101 | 010 101   111 101   101 101   100 101
$+5$

000 111   001 111   011 111   010 111 | 110 111   111 111   101 111   100 111
$+3$

000 110   001 110   011 110   010 110 | 110 110   111 110   101 110   100 110
$+1$

-7 -5 -3 -1 $+1$ $+3$ $+5$ $+7$ I

000 010   001 010   011 010   010 010 | 100 010   111 010   101 010   100 010
-1

000 011   001 011   011 011   010 011 | 100 011   111 011   101 011   100 011
-3

000 001   001 001   011 001   010 001 | 100 001   111 001   101 001   100 001
-5

000 000   001 000   011 000   010 000 | 100 000   111 000   101 000   100 000
-7

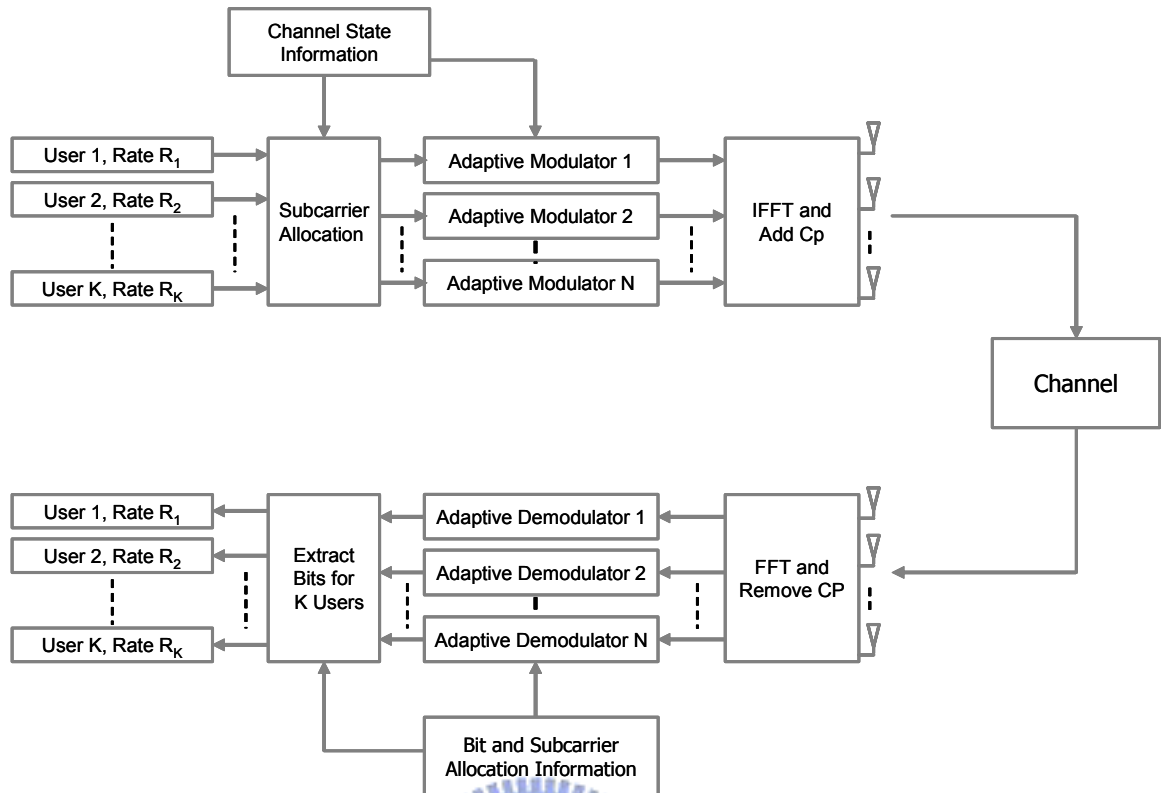Figure 4.4: BPSK, QPSK, 8-QAM, 16-QAM, 32-QAM, and 64-QAM constellation diagrams.

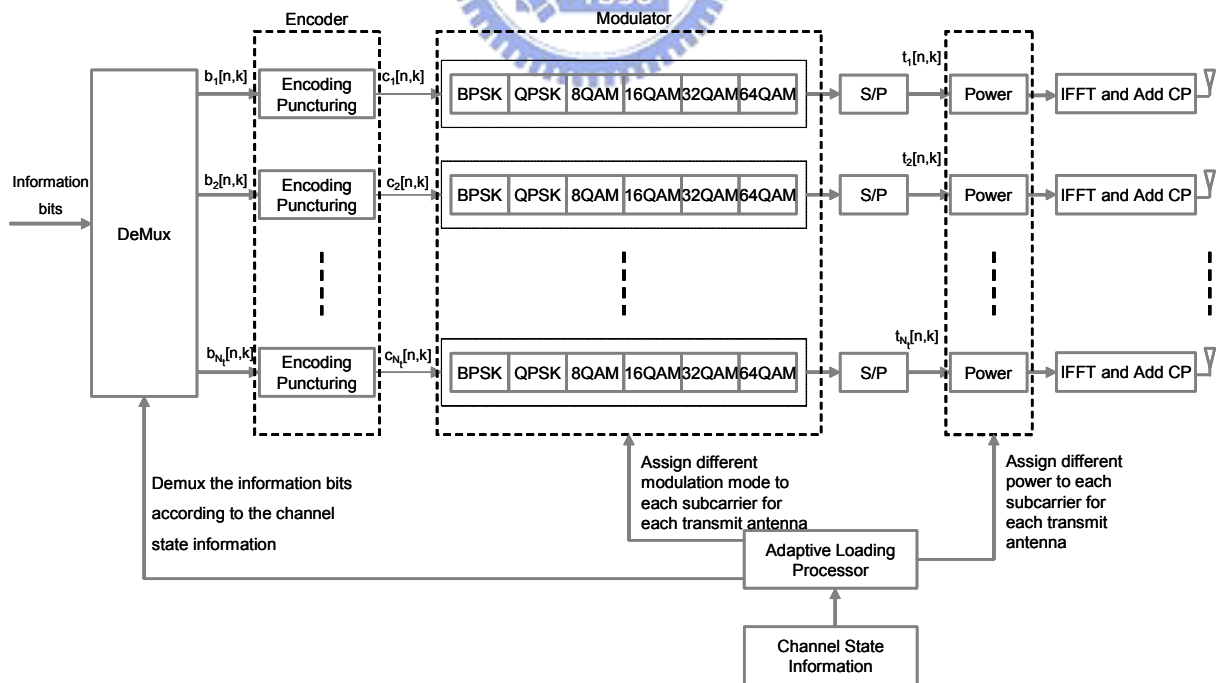Figure 4.5: Block diagrams of the multiuser adaptive MIMO-OFDM system.

Figure 4.6: V-BLAST based multiuser adaptive MIMO-OFDM transmitter acchitecture.
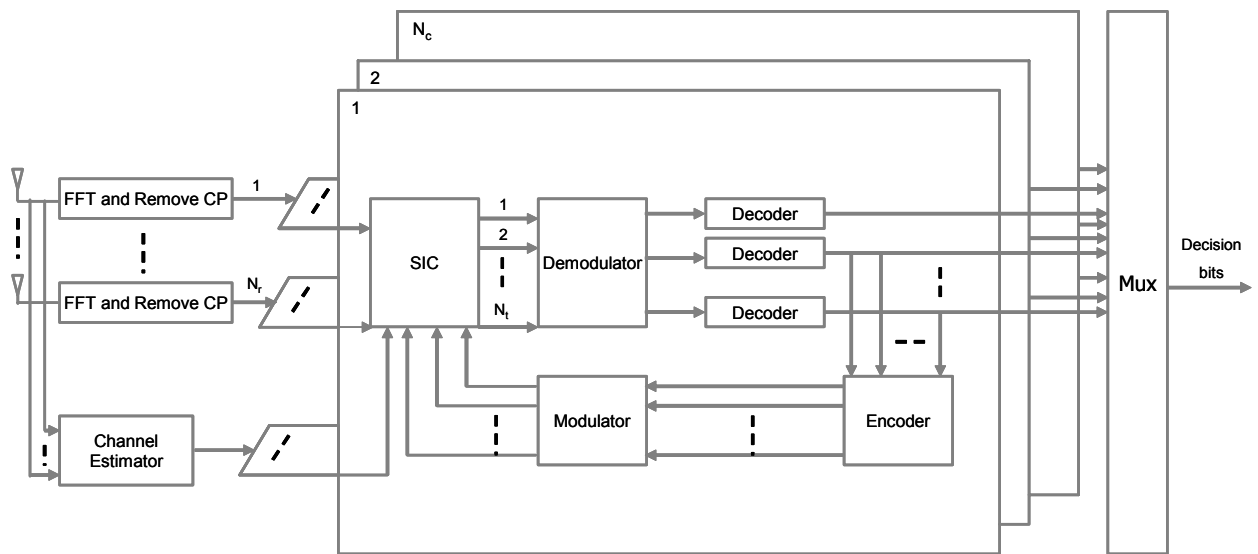
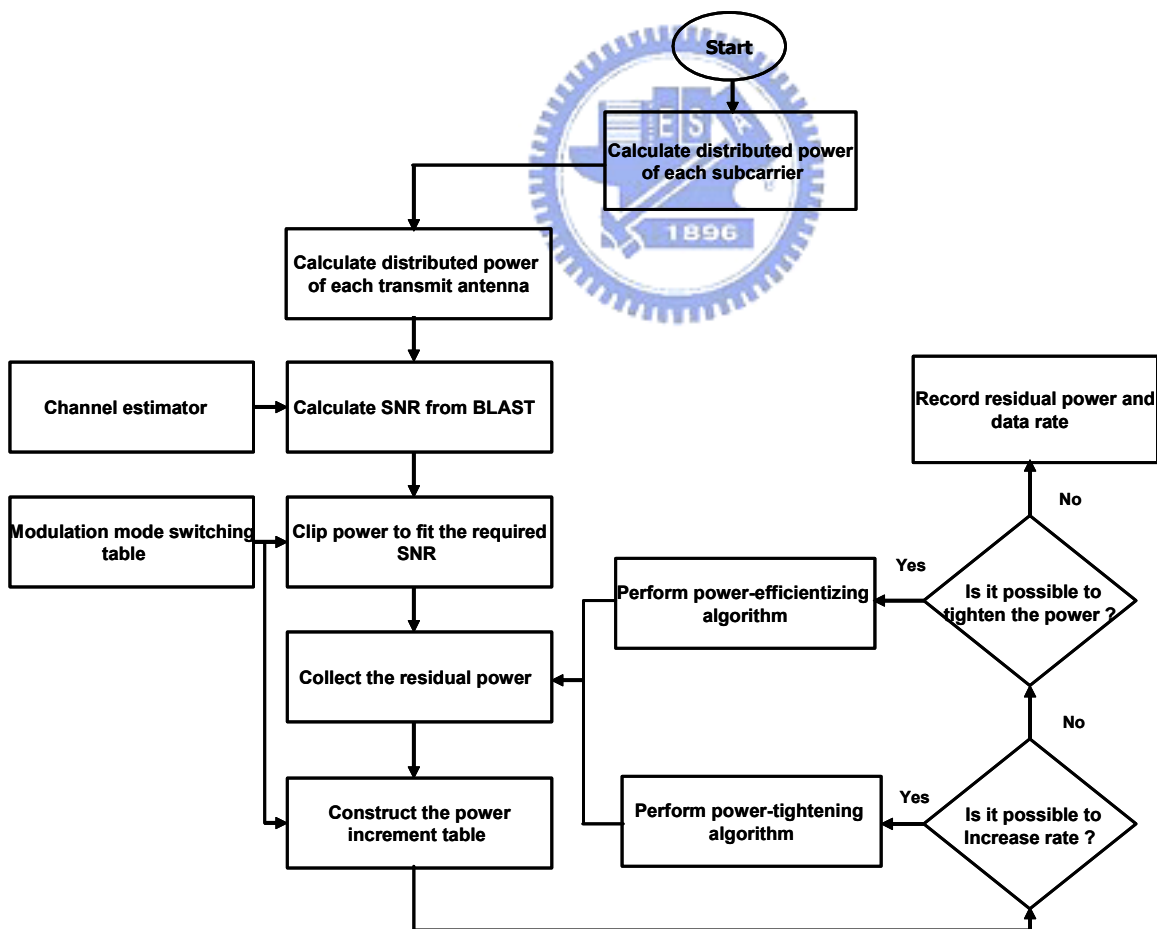Figure 4.7:   V-BLAST based multiuser adaptive MIMO-OFDM receiver architecture.

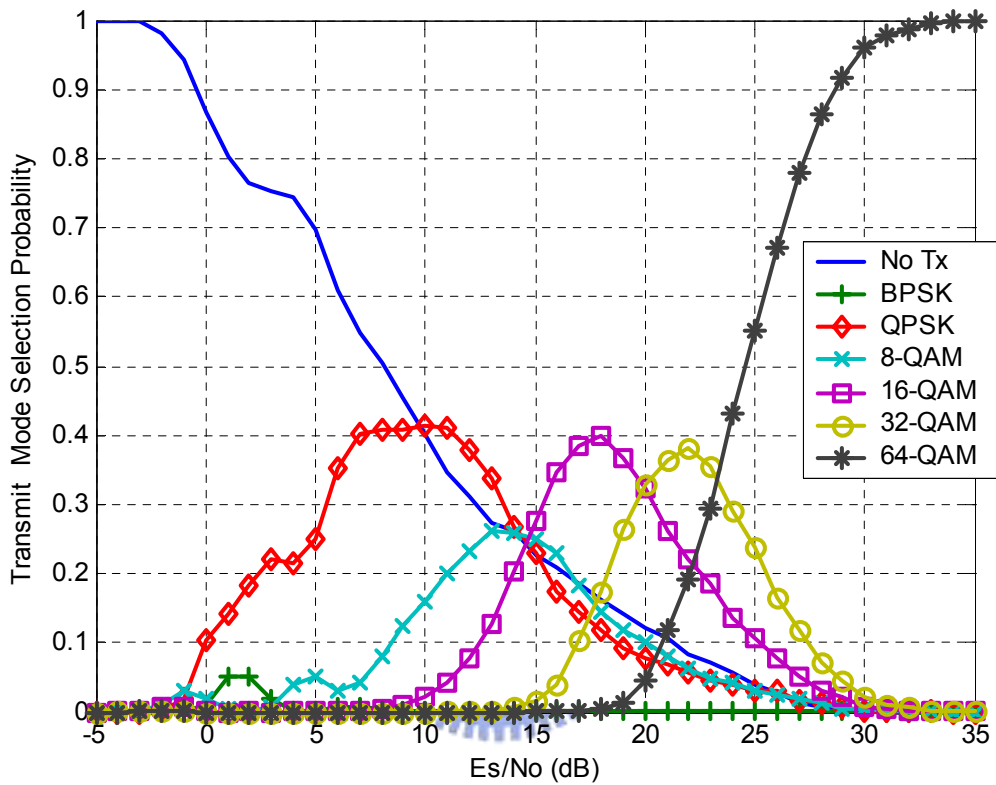Figure 4.8:   Flow chart of the first stage adaptive bit loading algorithm.

Figure 4.9: Simulated probabilities of each modulation mode utilized by the ZF V-BLAST based multiuser adaptive MIMO-OFDM system (with space-frequency loading) under the exponential decay Rayleigh fading channel with $\tau_{rms} = 50$ ns, and $f_d = 0$ Hz . $(N_t, N_r) = (4, 4)$. The number of users is 10. Other parameters are listed in Table 4.2.
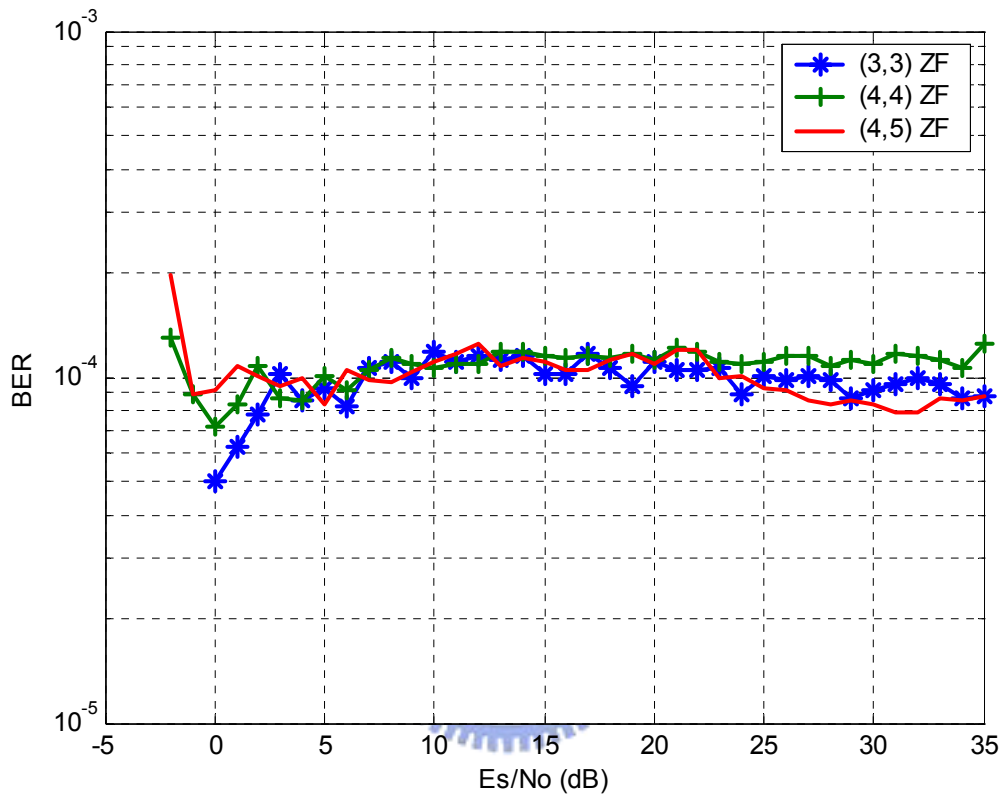
Figure 4.10:   BER versus E_s/N_0 for the ZF V-BLAST based multiuser adaptive MIMO-OFDM system without using residual power. The exponential decay Rayleigh fading channel is employed with τ_{rms} =50 ns, and $f_d = 0$ Hz. (N_t, N_r) = (3, 3), (4, 4), and (4, 5). The number of users is 10. Other parameters are listed in Table 4.2.
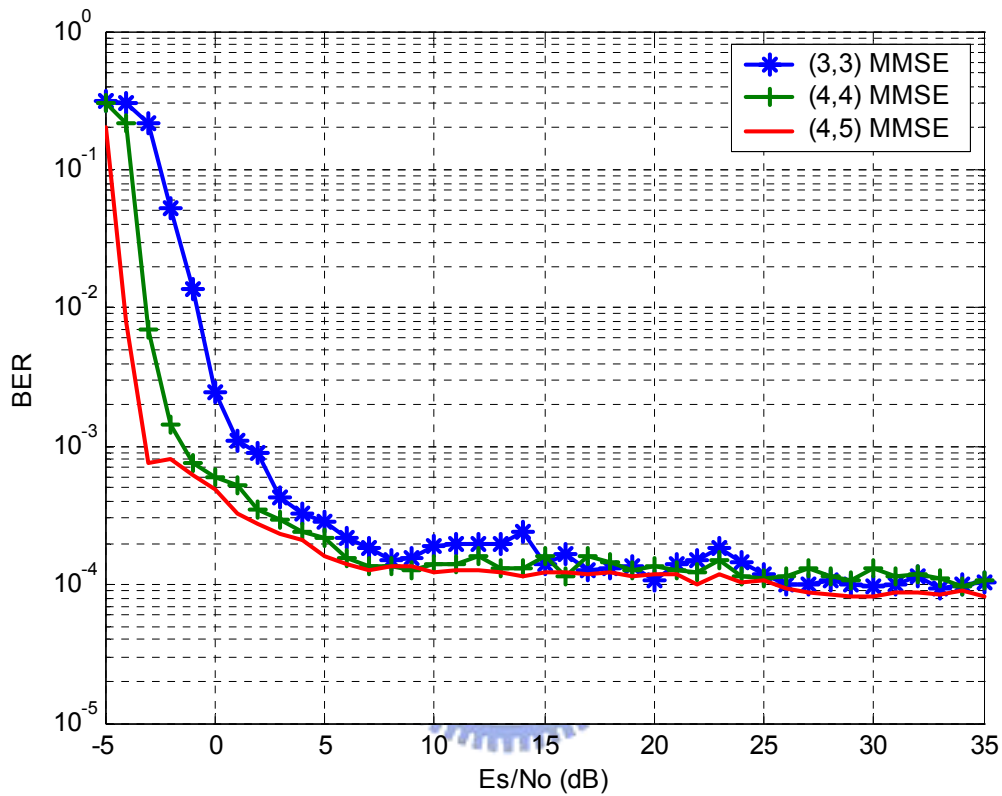
Figure 4.11:   BER versus E_s/N_0 for the MMSE V-BLAST based multiuser adaptive MIMO-OFDM system without using residual power. The exponential decay Rayleigh fading channel is employed with τ_{rms} =50 ns, and $f_d = 0$ Hz . $(N_t, N_r) = (3, 3), (4, 4)$, and $(4, 5)$. The number of users is 10. Other parameters are listed in Table 4.2.
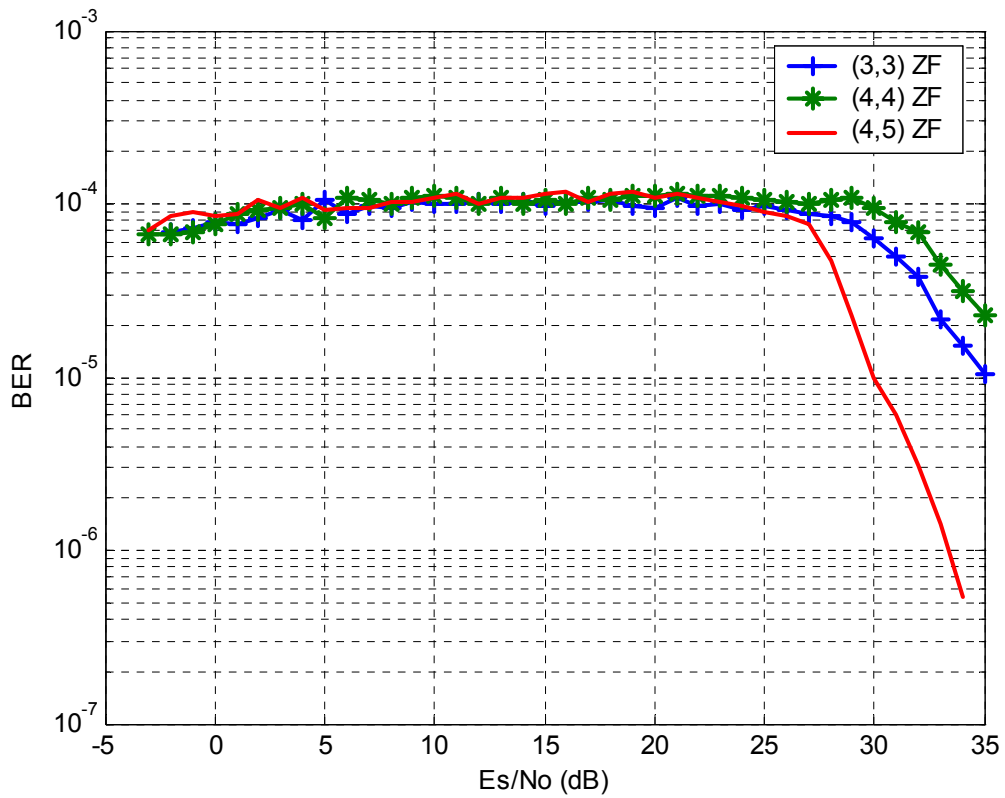
Figure 4.12: BER versus E_s/N_0 for the ZF V-BLAST based multiuser adaptive MIMO-OFDM system using residual power. The exponential decay Rayleigh fading channel is employed with τ_{rms} =50 ns, and $f_d = 0$ Hz. $(N_t, N_r) = (3, 3), (4, 4)$, and $(4, 5)$. The number of users is 10. Other parameters are listed in Table 4.2.
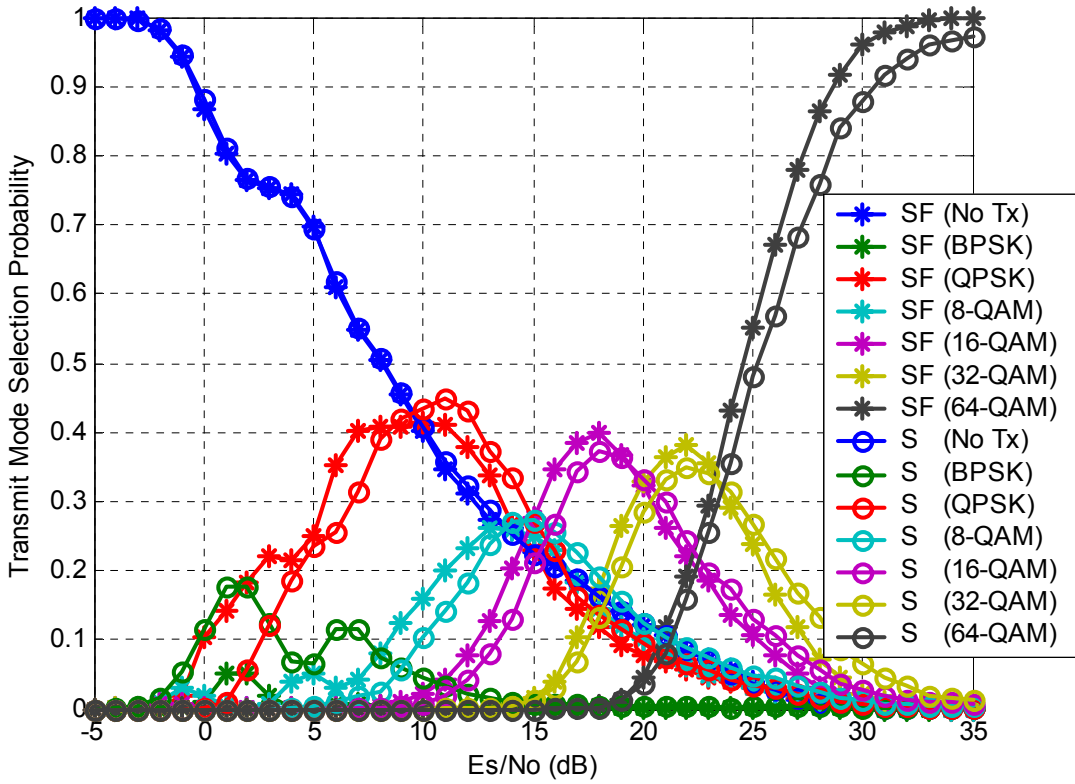
Figure 4.13:   Simulated probabilities of each modulation mode utilized by the ZF
V-BLAST based multiuser adaptive MIMO-OFDM system (with
space-frequency and space loading, respectively) under the exponential
decay Rayleigh fading channel with $\tau_{rms} = 50$ ns, and $f_d = 0$ Hz.
$(N_t, N_r) = (4, 4)$. The number of users is 10. Other parameters are listed
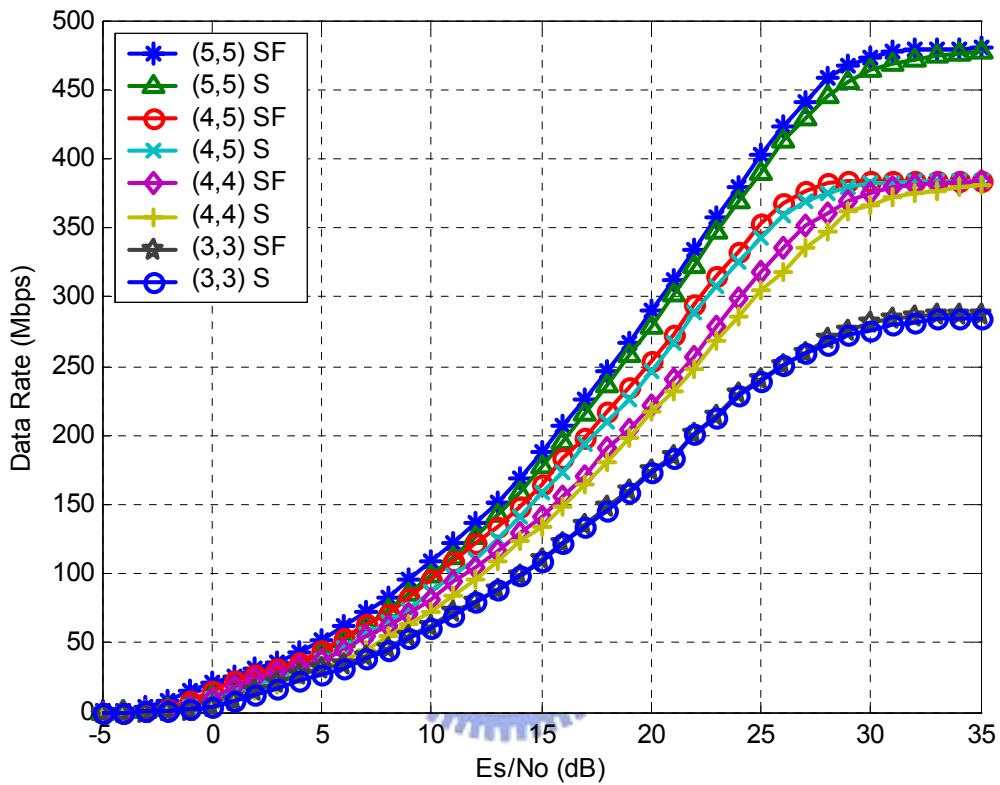in Table 4.2.

Figure 4.14: Data rate versus E_s/N_0 for the ZF V-BLAST based multiuser adaptive MIMO-OFDM system (with space-frequency and space loading, respectively) under the exponential decay Rayleigh fading channel with $\tau_{rms} = 50$ ns, and $f_d = 0$ Hz. $(N_t, N_r) = (3, 3), (4, 4), (4, 5)$, and $(5, 5)$. The number of users is 10. Other parameters are listed in Table 4.2.
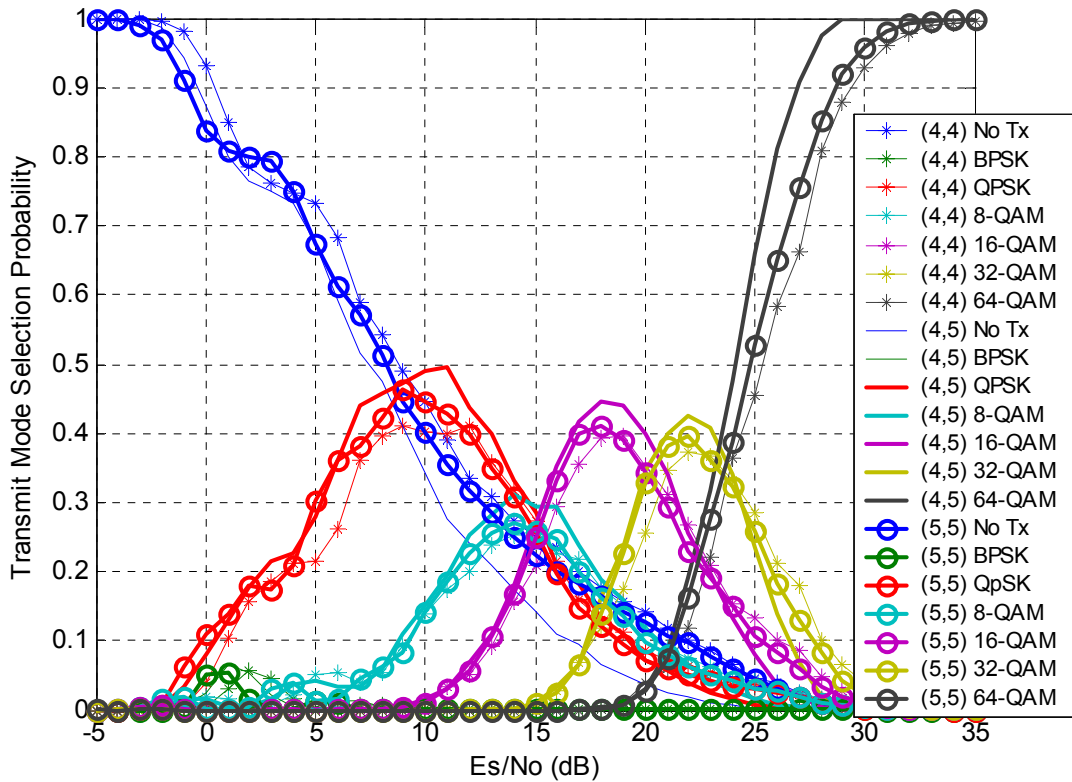
Figure 4.15: Simulated probabilities of each modulation mode utilized by the ZF V-BLAST based multiuser adaptive MIMO-OFDM system (with space-frequency loading) under the exponential decay Rayleigh fading channel with $\tau_{rms} = 50$ ns, and $f_d = 0$ Hz . $(N_t, N_r) = (4, 4), (4, 5)$, and $(5, 5)$. The number of users is 10. Other parameters are listed in Table 4.2.
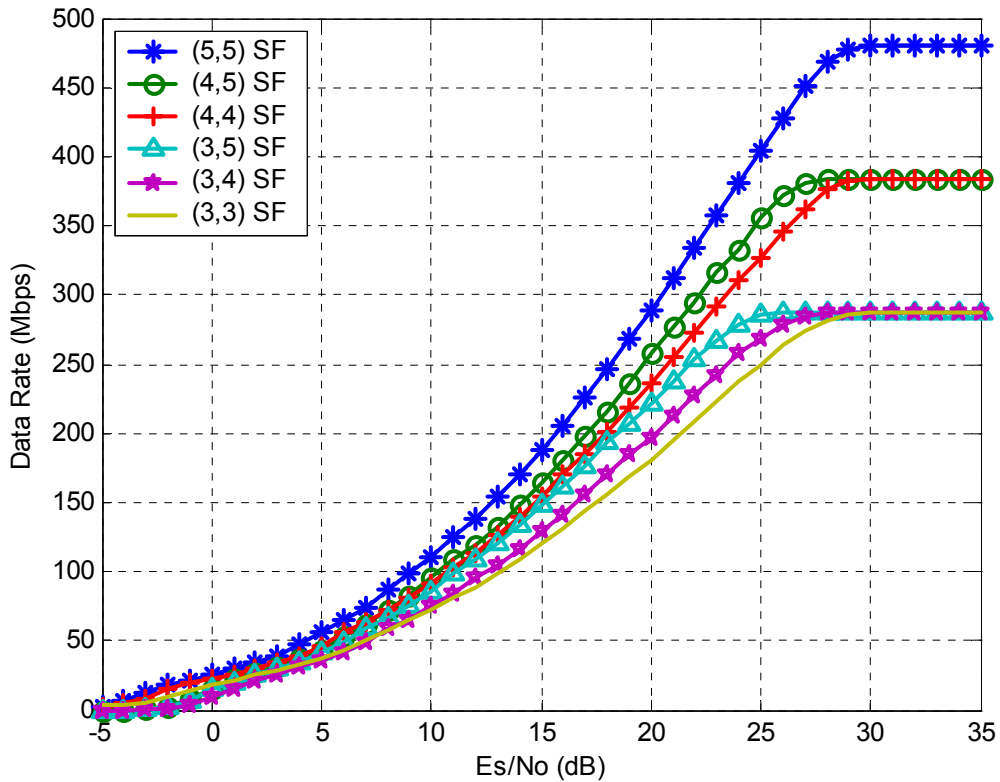
Figure 4.16: Data rate versus E_s/N_0 for the ZF V-BLAST based multiuser adaptive MIMO-OFDM system (with space-frequency loading) under the exponential decay Rayleigh fading channel with $\tau_{rms} = 50$ *n*s, and $f_d = 0$ Hz. $(N_t, N_r) = (3, 3), (3, 4), (3, 5), (4, 4), (4, 5),$ and $(5, 5)$. The number of users is 10. Other parameters are listed in Table 4.2.
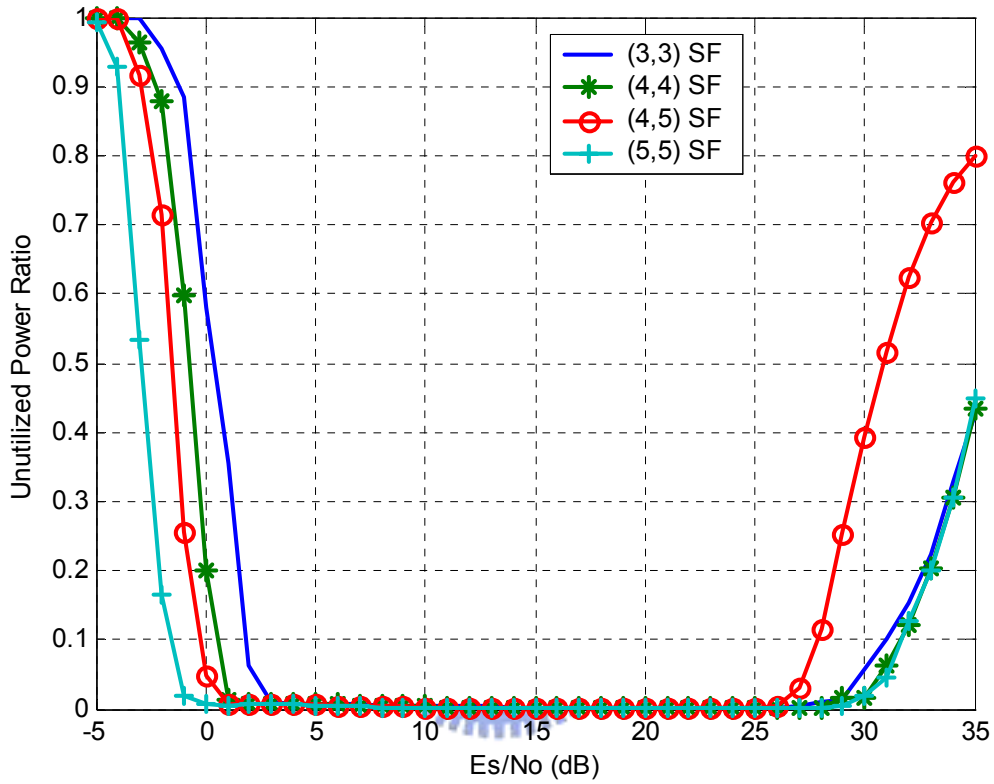
Figure 4.17: Unutilized power ratio versus E_s/N_0 for the ZF V-BLAST based multiuser adaptive MIMO-OFDM system (with space-time loading) under the exponential decay Rayleigh fading channel with $\tau_{rms} = 50$ ns, and $f_d = 0$ Hz. $(N_t, N_r) = (3, 3), (4, 4), (4, 5),$ and $(5, 5)$. The number of users is 10. Other parameters are listed in Table 4.2.
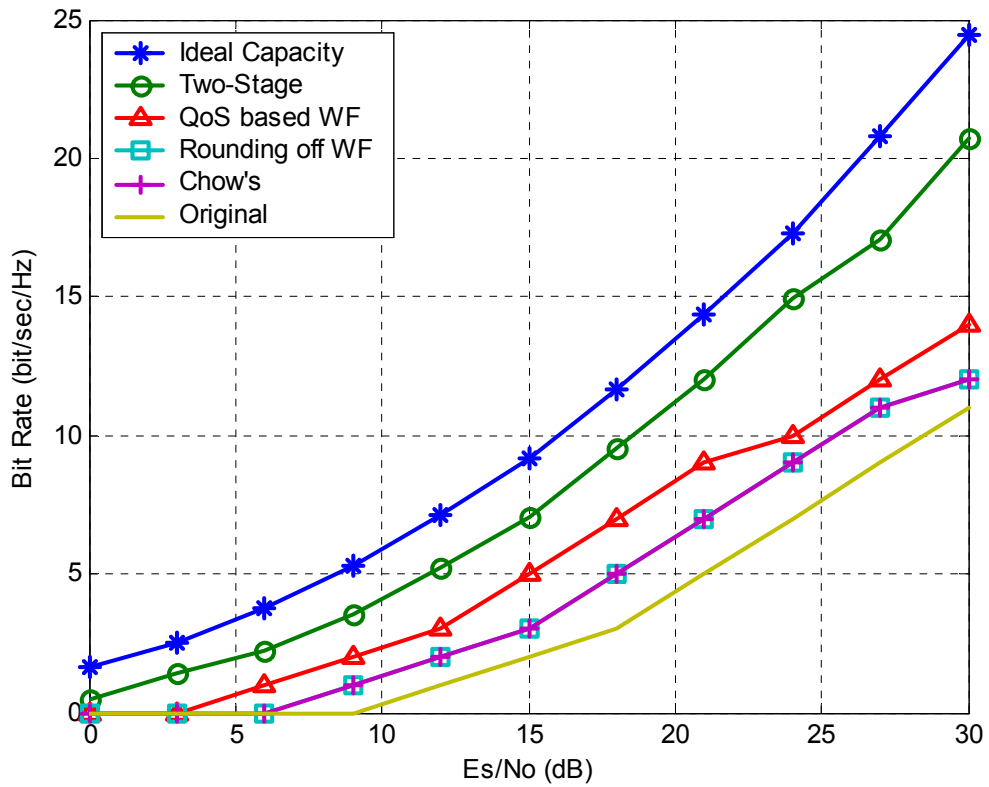
Figure 4.18: Bit rate versus E_s/N_0 for the ZF V-BLAST based multiuser adaptive MIMO-OFDM system with different bit loading algorithms under the exponential decay Rayleigh fading channel with $\tau_{rms} = 50$ ns, and $f_d = 0$ Hz . (N_t, N_r) = (4, 4). The number of users is 10. Other parameters are listed in Table 4.2.
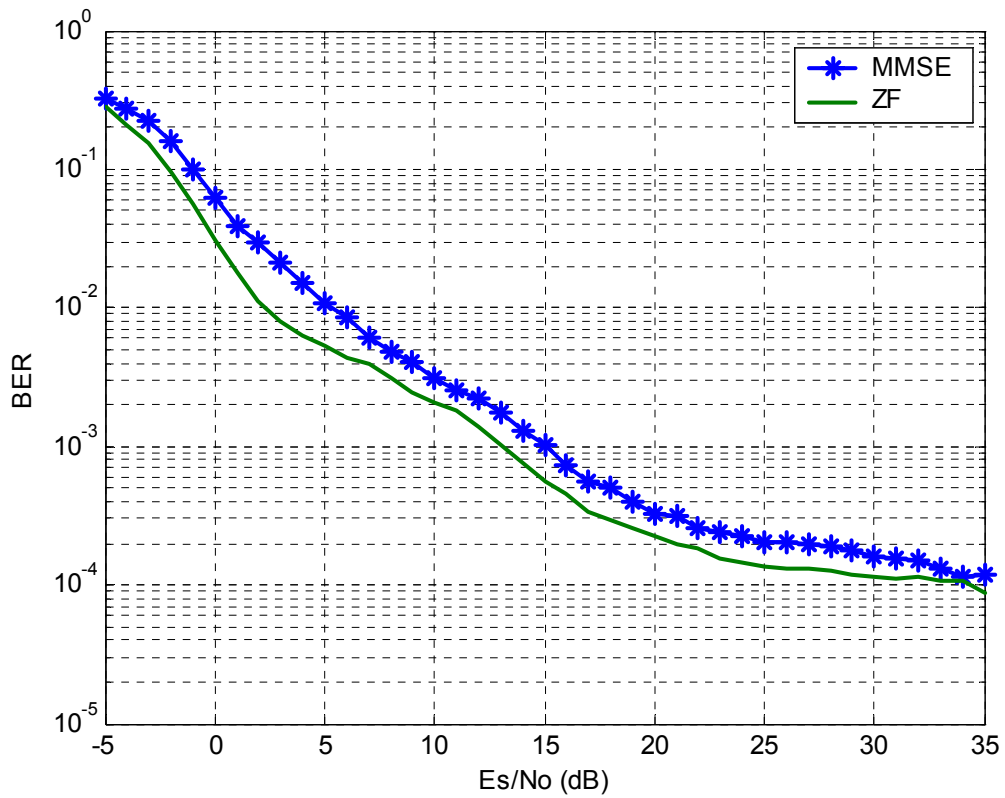
Figure 4.19:   BER versus E_s/N_0 for the V-BLAST based multiuser adaptive MIMO-OFDM system using residual power with different detection criteria. The exponential decay Rayleigh fading channel is employed with $\tau_{rms} = 50$ ns, and $f_d = 0$ Hz. $(N_t, N_r) = (4, 4)$. ΔH is equal to the noise power. The number of users is 10. Other parameters are listed in Table 4.2.
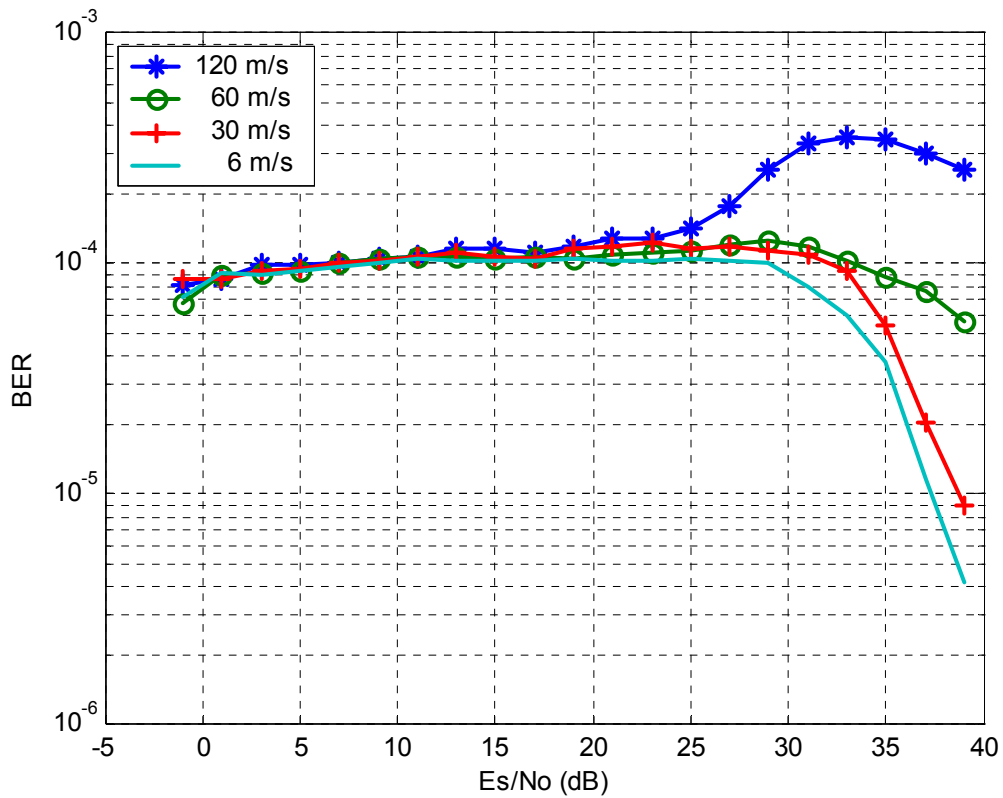
Figure 4.20:   BER versus E_s/N_0 for the ZF V-BLAST based multiuser adaptive MIMO-OFDM system under the exponential decay Rayleigh fading channel with $\tau_{rms} = 50$ ns. $(N_t, N_r) = (4,4)$. v = 120, 60, 30, and 6 m/s. The number of users is 10. Other parameters are listed in Table 4.2.

Table 4.1: SNR threshold table for various M-QAM at the target BER = 10^{-4}.

	No TX	BPSK	QPSK	8-QAM	16-QAM	32-QAM	64-QAM
S	$-\infty$	8.4	11.4	15.5	18.5	21.6	24.8

Table 4.2: Simulation parameters for the proposed V-BLAST based multiuser adaptive MIMO-OFDM system.
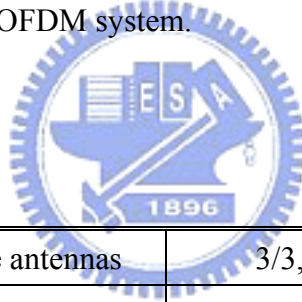
Number of transmit/receive antennas	3/3, 3/4, 3/5, 4/4, 4/5, and 5/5
Carrier frequency	5 GHz
Channel bandwidth	20 MHz
Number of carriers, FFT size	64
Number of users	10
OFDM symbol duration	$3.2\,\mu$s
Guard interval	$0.8\,\mu$s
M-QAM available	0, 1, 2, 3, 4, 5, and 6
Number of OFDM symbols in a packet	100
Channel model	Exponential delay profile, τ_{rms} = 50 ns

Chapter 5

Conclusion

In this thesis, the multiuser adaptive MIMO-OFDM system is proposed incorporating dynamic subcarrier allocation and adaptive bit loading algorithms. The dynamic subcarrier allocation algorithm solves the subcarrier assignment problem and the adaptive bit loading algorithm decides the transmission parameters on each subcarrier. In Chapter 3, the two-stage subcarrier allocation algorithm is presented. This algorithm considers users' different quality of service (QoS) requirements and allocates to each user a group of suitable subcarriers according to their channel qualities. The core idea of this algorithm is to divide the subcarrier assignment problem into two stages. In the first stage, it determines the number of subcarriers assigned to each user. In the second stage, it decides which subcarriers are given to which user. Through judiciously assigning subcarriers to users, the overall transmission rate of the multiuser MIMO-OFDM system can be increased. This algorithm is also compared with other two dynamic subcarrier allocation algorithms, and it is shown that the two-stage subcarrier allocation algorithm outperforms other two algorithms in terms of the total transmission rate and execution time under an exponential decay Rayleigh fading channel.

After subcarrier assignment, the two-stage adaptive bit loading algorithm is presented in Chapter 4. Under the total transmit power constraint, this algorithm aims to meet the target BER and further enhances the total transmission rate of the system. For instance, the subcarriers with good channel qualities are more likely to employ higher modulation order to increase the overall transmission rate of the system, while

the subcarriers with poor channel qualities are more likely to employ lower modulation order to maintain the target BER. The core idea of this algorithm is to separate the joint space-frequency optimization problem into two stages. In the first stage, the adaptive bit loading algorithm is applied to each subcarrier to obtain an optimal bit and power allocation. In the second stage, the same loading algorithm is processed over those active subchannels surviving from the first stage to further exhaust the residual power and achieve a rate enhancement. It can be demonstrated that this algorithm works well in meeting the target BER in the low SNR scenarios and further increases the total transmission rate of the system in the high SNR scenarios.

Besides, the main feature of the proposed system is that the adaptive bit loading algorithm is transparent to the number of transmit and receive antennas. That is, the same version of link adaptation software can be employed in the modem regardless of the number of antennas being used or their radio frequency (RF) characteristics. Such an idea is widely used in the concept of software-defined radio (SDR).

In the multiuser adaptive MIMO-OFDM system, the perfect channel state information (CSI) is assumed known to the transmitter. In practice, it is impossible to obtain perfect CSI due to noisy channel estimation and the unavoidable delay between performing channel estimation and using estimation result for actual transmission. Therefore, a major issue of the proposed system is what type of CSI can be made practically available to the transmitter in wireless communication environments, where fading channels are randomly varying. In recent research, partial CSI proves to have great practical value. As with perfect CSI, partial CSI is made available either through a feedback channel from the receiver to the transmitter, or when the transmitter acts as the receiver in a time or frequency division duplex operation. Based on partial CSI, the potential further work can be directed to the re-design of dynamic subcarrier allocation and adaptive bit loading algorithms suited to multiuser MIMO-OFDM systems.

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 1rchitecture 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, 2000.

[5] A. F. Naguib and R. Calderband, "Space-time coding and signal processing for high data rate wireless communications," *Wirel. Commun. Mob. Comput.,* vol. 1, no. 1, pp. 13-34, Jan. 2001.

[6] Z. Liu, G. B. Giannakis, S. Zhou, and B. Muquet, "Space-time coding for broadband wireless communications," *Wirel. Commun. Mob. Comput.*, vol. 1, no. 1, pp. 35-53, Jan. 2001.

[7] B. Holter, "On the capacity of the MIMO channel –a tutorial introduction-," *NORSIG'01*, pp. 129-143, 2001.

[8] 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.

[9] 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.

[10] I. BarhumiLi, G. Leus, and M. Moonen, "Optimal training sequence for channel estimation in MIMO OFDM systems in mobile wireless channels," *International Zurich Seminar on Broasband Communications*, pp. 44-1 – 44-6, Feb. 2002.

[11] G. J. Foschini, G. D. Golden, R. A. Valenzuela, and P. W. Wolniansky, "Simplified processing for high spectral efficiency wireless communication employing multi-element arrays," *IEEE J. Select. Areas Commu.*, vol. 17, no. 11, pp. 1841-1852, Nov. 1999.

[12] P. J. Smith and M. Shafi, "On a Gaussian approximation to the capacity of wireless MIMO systems," *Proc. IEEE ICC'02,* vol. 1, no. 28, pp. 406-410, May. 2002.

[13] M. Ergen, S. Coleri, and P. Varaiya, "QoS aware adaptive resource allocation techniques for fair scheduling in OFDMA based broadband wireless access systems," *IEEE Trans. Broadcasting*, vol. 49, pp. 362-370, Dec. 2003.

[14] I. Koutsopoulos and L. Tassiulas, "Adaptive resource allocation in SDMA-based wireless broadband networks with OFDM signaling," *IEEE INFOCOM*, vol. 3, pp. 1376-1385, Jun. 2002.

[15] A. Goldsmith and S. Chua, "Variable-rate variable-power MQAM for fading channels," *IEEE Trans. Commu.*, vol. 45, pp. 1218-1230, Oct. 1997.

[16] S. Catreux, D. Gesbert, V. Ercge, and R. W. Heath JR, "Adaptive modulation and MIMO coding for broadband wireless data networks," *IEEE Communications Magazine*, Jun. 2002.

[17] 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.

[18] H. Sampath, S. Talwar, J. Tellado, V. Erceg, and A. Paulraj, "A fourth-generation MIMO-OFDM broadband wireless system: design, performance, and field trial results," *IEEE Communications Magazine*, vol. 40, no. 9, pp. 143-149, Sep. 2002.

[19] T. S. Rappaport, A. Annamalai, R. M. Buehrer, and W. H. Tranter, "Wireless communications: past events and a future perspective," *IEEE Communications Magazine*, vol. 40, no. 5, pp. 5-14, May. 2002.

[20] G. G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Trans. Commun.,* vol. 46, pp. 357-366, Mar. 1998.

[21] R. Knopp and P. A. Humblet, "Information capacity and power control in single-cell multiuser communications," *Proc. IEEE ICC'95,* pp. 331-335, Jun. 1995.

[22] C. Wong, R. Cheng, K. Letaief, and R. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE J. Select. Areas Commun.*, vol. 17, no. 10, pp. 1747-1758, Oct. 1999.

[23] D. Kivanc and H. Lui, "Subcarrier allocation and power control for OFDMA," *Conference on Signals, Systems, and Computers,* vol. 1, pp. 147–151, 2000.

[24] H. Yin and H. Liu, "An efficient multiuser loading algorithm for OFDM-based broadband wireless systems," *IEEE GLOBECOM,* vol. 1, pp. 103-107, Dec. 2000.

[25] K.W. Ng, R. S. Cheng, and R. D. Murch, "Iterative bit & power allocation for V-BLAST based MIMO OFDM system in frequency selective fading channel," *IEEE WCNC2002,* vol. 1, pp. 271-275, Mar. 2002.

[26] E. Lawrey, "Muitiuser OFDM," *Proc. IEEE ICCPA'99,* vol. 2, pp. 761-764, Aug. 1999.

[27] W. Rhee and J. M. Cioffi, "Increasing in capacity of multiuser OFDM system using dynamic subchannel allocation," *IEEE VTC 2000-Springl,* vol. 2, pp. 1085-1089, May. 2000.

[28] I. Kim, H. L. Lee, B. Kim, and Y. H. Lee, "On the use of linear programming for dynamic subchannel and bit allocation in muitiuser OFDM," *IEEE GLOBECOM,* vol. 6, pp. 3648-3652, Nov. 2001.

[29] A. Seyedi and G. J. Saulnier, "A distributed algorithm for dynamic sub-channel assignment in a multiuser OFDM communication system," *Proc. IEEE Signal Processing Workshop,* vol. 6, pp. 397-400, Aug. 2001.

[30] A. Pandharipande and S. Dasgupta, "Optimum multiuser OFDM systems with unequal subchannel assignment," *IEEE ICC'03,* vol. 5, pp. 3423-3427, May. 2003.

[31] L. Zhen, Z. Geqing, W. Weihua, and S. Junde, "Improved algorithm of multiuser dynamic subcarrier allocation in OFDM system," *IEEE ICCT'03,* vol. 2, pp. 1144-1147, Apr. 2003.

[32] S. Pfleschinger, G. Munz, and J. Speidel, "Efficient subcarrier allocation for multiple access in OFDM systems," *InOWo'02*, pp. 21-25, Sep. 2002.

[33] J. Gross, H. Karl, F. Fitzek, and A. Wolisz, "Comparison of heuristic and optimal subcarrier assignment algorithms," *Proc. IEEE ICWN'03*, pp. 249-255, Jun. 2003.

[34] P. S. Chow, J. M. Cioffi, and J. A. C. Bingham, "A practical discrete multitone transceiver loading algorithm for data transmission over spectrally shaped channels," *IEEE Trans. Commu.*, vol. 43, no. 3, pp. 773-775, Apr. 1995.

[35] J. Campello, "Practical bit loading for DMT," *IEEE ICC'99*, vol. 2, pp. 801-805, Jun. 1999.

[36] J. Campello De Souza, "Discrete bit loading for multicarrier modulation systems," PhD. Dissertation, Stanford University, 1999.

[37] R. F. H. Fischer and J. B. Huber, "A new loading algorithm for discrete multitone transmission," *IEEE GLOBECOM*, vol. l, pp. 724-728, Nov. 1996.

[38] X. Zhang and B. Ottersten, "Power allocation and bit loading for spatial multiplexing in MIMO systems," *IEEE ICASSP'03*, vol. 5, pp. 53-56, Apr. 2003.

[39] J. Terry and J. Heiskala, "*OFDM Wireless LANs: A Theoretical and Practical Guide*," Sams, 2001.

[40] B. O'Hara and A. Petrick, "*The IEEE 802.11 Handbook: A Designer's Companion*," New York: IEEE Press, 1999.