# TRANSMIT AND RECEIVE TECHNIQUES FOR MIMO OFDM SYSTEMS

SUMEI SUN

NATIONAL UNIVERSITY OF SINGAPORE

2006

# TRANSMIT AND RECEIVE TECHNIQUES FOR MIMO OFDM SYSTEMS

**SUMEI SUN**

(*B. Sc.(Hons.), Peking University, M.Eng, Nanyang Technological University*)

A THESIS SUBMITTED

FOR THE DEGREE OF PH.D

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

March 2006

# Acknowledgement

I would sincerely like to thank my thesis supervisor, Professor Tjeng Thiang Tjhung, for his constant guidance, encouragement, patience, and support, without which this thesis would not have been possible. His enthusiasm and serious attitude in research has set a great example for me and I believe I will benefit from it beyond this work.

I would like to thank my colleagues, Yan Wu, Chin-Keong, Ying-Chang, Yongmei, Yuan Li, Zhongding, Woon Hau, Patrick, and Hongyi, for the interesting technical discussions and sharing, and the enjoyable environment we have created together, in which research has been full of fun.

My special thanks also go to Professor Pooi Yuen Kam, Professor Chun Sum Ng, and Dr. A. Nallanatham for sitting in my thesis committee and for their advices.

Last but not least, I would like to thank my family for their understanding, tolerance, encouragement and unconditional support, especially my two lovely children Xinyi and Jiarui who have made my life so meaningful and joyful.

# Table of Contents

# List of Figures

# List of Tables

# List of Abbreviations

| | |
|---|---|
| 2DLPT | two-dimensional linear pre-transform |
| 3G | third generation |
| ARQ | automatic repeat request |
| AWGN | additive white Gaussian noise |
| BICM | bit-interleaved coded modulation |
| BLAST | Bell Lab LAyered Space-Time |
| BMMSE | Bayesian minimum mean squared error |
| bps | bits per second |
| CC | convolutional code |
| CDD | Cyclic Delay (transmit) Diversity |
| CDDSS | Cyclic Delay Diversity with Spatial Spreading |
| CDMA | code division multiple access |
| CP | Cyclic Prefix |
| CSI | channel state information |
| CSCG | circularly symmetric complex Gaussian |
| DBLAST | Diagonal BLAST |
| DFT | Discrete Fourier Transform |
| DM | direct mapping |
| ECC | error correction code |
| EGC | equal gain combining |
| EM | expectation maximization |
| EPDF | exponential power delay profile |

| | |
|---|---|
| ETSI | European Telecommunications Standards Institute |
| EXIT | EXtrinsic Information Transfer |
| EXT | extrinsic information |
| FEC | Forward error correction |
| FFT | Fast Fourier Transform |
| GSM | Global System for Mobile Communications |
| GSTBC | Groupwise Space Time Block Code(d) |
| GSTTC | Groupwise Space Time Trellis Code(d) |
| IDFT | Inverse Discrete Fourier Transform |
| IEEE | Institute of Electrical & Electronic Engineers |
| IFFT | Inverse Fast Fourier Transform |
| ISI | intersymbol interference |
| ITU | International Telecommunication Union |
| LAN | Local Area Network |
| LDC | Linear Dispersion Code(d) |
| LLR | log-likelihood ratio |
| LMMSE | linear minimum mean squared error |
| LPT | linear pre-transform |
| LS | least squares |
| MAP | maximum *a posteriori* |
| MIMO | Multiple-Input Multiple-Output |
| MISO | Multiple-Input Single-Output |
| ML | maximum likelihood |
| MMSE | minimum mean squared error |
| MRC | maximal ratio combining |
| OFDM | Orthogonal Frequency Division Multiplexing |

| | |
|---|---|
| pdf | probability density function |
| PDF | power delay profile |
| PEP | pair-wise error probability |
| pmf | probability mass function |
| PT | pre-transform |
| RAS | receive antenna selection |
| RF | radio frequency |
| SD | sphere decoding |
| SDF | soft decision function |
| SIMO | Single-Input Multiple-Output |
| SISO | soft-input soft-output |
| SNR | Signal to Noise Ratio |
| SS | spatial spreading |
| ST | Space-Time |
| STBC | Space Time Block Code(d) |
| STC | Space Time Code(d) |
| STFP | Space-Time-Freq(uency)-Precoding |
| STSR | single-transmit single-receive |
| STTC | Space Time Trellis Code(d) |
| SWF | Statistical Water Filling |
| SVD | singular value decomposition |
| TAS | transmit antenna selection |
| UPDF | uniform power delay profile |
| VBLAST | Vertical BLAST |
| WLAN | Wireless Local Area Network |

# List of Symbols

| | |
|---|---|
| $n_T$ | number of transmit antennas |
| $n_R$ | number of receive antennas |
| $n_S$ | number of spatial streams |
| $\mathbf{R}$ | frequency domain received signal vector at each subcarrier |
| $\mathbf{X}$ | frequency domain transmitted signal vector at each subcarrier |
| $\mathbf{H}$ | frequency domain (precoded) channel matrix at each subcarrier |
| $\mathbf{V}$ | frequency domain AWGN noise vector at each subcarrier |
| $L$ | number of multipath components (sample-spaced) |
| $L_{\mathrm{CP}}$ | cyclic prefix length |
| $N$ | FFT size of an OFDM system |
| $P$ | number of subcarriers used to transmit data and pilots |
| $\mathcal{E}$ | statistical expectation |
| $CN(\mathbf{m}, \mathbf{Q})$ | complex Gaussian distribution with mean $\mathbf{m}$ and covariance matrix $\mathbf{Q}$ |
| $C$ | channel capacity |
| $C_{\mathrm{E}}$ | ergodic capacity |
| $\tilde{X}_{k,i}$ | decision statistic of signal $X_k$ at iteration $i$ |
| $\hat{X}_{k,i}$ | statistical mean estimation of signal $X_k$ at iteration $i$ |
| $\breve{X}_{k,i}$ | Bayesian MMSE estimation signal of $X_k$ at iteration $i$ |

# Summary

This thesis is concerned in general with the transmit and receive techniques for multiple-input multiple-output (MIMO) orthogonal frequency-division multiplexing (OFDM) systems in wideband frequency selective fading channels. In particular we address issues such as the space-time-frequency precoding schemes to achieve optimal or near-optimal capacity and diversity performance in MIMO-OFDM channels, optimal and efficient detection and decoding of transmitted sequence at the receiver, and optimal training signal design and low-complexity channel estimation to support coherent detection and optimal decoding.

In rich-scattering environments, a MIMO channel created by deploying multiple antenna arrays at both the transmitter and the receiver of a wireless link can provide both multiplexing gain and diversity gain. For a MIMO channel with fixed dimensions, i.e., fixed number of transmit and receive antennas, there is a tradeoff between the multiplexing gain and the diversity gain. A high diversity gain can only be achieved at the cost of reduced multiplexing gain. When deployed in wideband frequency selective channels, MIMO can be combined with OFDM to efficiently mitigate the intersymbol interference. To further exploit the frequency diversity inherent in frequency selective channels, error control coding or pre-transform can be used with OFDM. Therefore, how to achieve the required multiplexing gain and diversity gain from the spatial and frequency domains is an important design issue for MIMO-OFDM systems.

For wireless communication systems, an asymmetric MIMO channel with more transmit than receive antennas is typically created for downlink transmission, due to the size and power limitation of the mobile terminal. We address the multiplexing and diversity gains of asymmetric MIMO-OFDM channels through space-time-frequency precoding, which can map fewer spatial data streams to more transmit antennas. Both linear and nonlinear precoding schemes are considered. A unified linear system model for the precoding schemes considered is established, with which we obtain the capacity and diversity performance of the precoded MIMO-OFDM channels in a unified approach. A two-dimensional linear pre-transformed MIMO-OFDM system is proposed in this thesis which achieves full capacity and full diversity simultaneously when the number of spatial data streams is equal to the number of transmit antennas, and full diversity and maximum capacity of a symmetric MIMO channel when the number of spatial streams is less than the number of transmit antennas.

Exploitation of the diversity and multiplexing gains in the MIMO-OFDM channel relies on not only the precoding scheme at the transmitter, but also optimal and efficient receiver algorithms. For receiver design, we dedicate our effort in this thesis to the iterative algorithms. In particular, a Bayesian minimum mean squared error turbo receiver is proposed. Compared with the conventional turbo receivers in the literature which make use of only the extrinsic information from the decoder for interference estimation and cancelation, the proposed Bayesian turbo receiver uses both the decoder extrinsic information *and* the detector decision statistic for interference estimation. As a result, the estimation accuracy is greatly improved, especially in low to medium SNR regions. This also contributes to the 1.5 dB improvement at BER performance of $10^{-5}$, and the better convergence behavior of the turbo process.

To further analyze the performance of the proposed Bayesian turbo receivers, the extrinsic information transfer chart is derived and compared with that of the conventional turbo receivers, in both fixed and random MIMO channels. A much higher output mutual information is demonstrated from the Bayesian turbo detector, proving its superior performance. When plotted with the extrinsic information transfer chart of the decoder, the trajectories of the Bayesian receivers also exhibit much faster convergence than the conventional receivers.

Effective realization of the capacity and diversity potential in the MIMO-OFDM channels requires efficient space-time-frequency precoding and optimal receiver design. For the turbo receivers discussed in the thesis, accurate channel state information is needed at the receiver. Four training signal schemes are proposed, two of which to support frequency-domain channel estimation, and the other two to support time-domain channel estimation. All the training signal design schemes are optimized to achieve the minimum mean squared error performance.

# List of Publications

- **Journal Paper**

  1. **Sumei Sun**, Yan Wu, Yuan Li, and Tjeng Thiang Tjhung, *"A Bayesian MMSE Turbo Receiver for Coded MIMO OFDM systems"*, submitted to IEEE Trans. Vehicular Technology, January 2005, and first revision March 2006

- **Conference Papers**

  1. **Sumei Sun**, Yan Wu, and Tjeng Thiang Tjhung, *"A Two-Dimensional Linear Pre-Transformed (2DLPT) MIMO-OFDM System"*, ICC 2007
  2. **Sumei Sun**, Ying-Chang Liang, Yan Wu, and Tjeng Thiang Tjhung, *"Precoding for asymmetric MIMO-OFDM channels"*, ICC 2006, Turkey, June 2006
  3. **Sumei Sun**, Ying-Chang Liang, and Tjeng Thiang Tjhung, *"Space-time precoding for asymmetric MIMO channels"*, WCNC 2006, Las Vegas, April 2006
  4. **Sumei Sun**, Yan Wu, Yuan Li, and Tjeng Thiang Tjhung, *"Exit chart analysis of Bayesian MMSE turbo receiver for coded MIMO systems"*, IEEE VTC 2005 Fall, Dallas, Texas, USA., September 2005
  5. **Sumei Sun**, T. T. Tjhung, and Y. Li, *"An Iterative Receiver for Groupwise Bit-Interleaved Coded QAM STBC OFDM"*, VTC 2004 Spring, Milan, Italy, May 2004
  6. **Sumei Sun**, Y. Wu, Y. Li, and T. T. Tjhung, *"A Novel Iterative Receiver for Coded MIMO OFDM Systems"*, ICC 2004, Paris, France, June 2004

7. **Sumei Sun**, I. Wiemer, Chin Keong Ho, and T. T. Tjhung, *"Training-Sequence Assisted Channel Estimation for MIMO OFDM"*, Proceedings of WCNC 2003, pp. 38-43, Vol. 1, New Orleans, LA, USA

8. **Sumei Sun**, and T. T. Tjhung, *"Soft-decision Based Iterative Interference Cancellation for Group-Wise STBC MIMO Systems"*, Proceedings of VTC 2003 Spring, pp. 984 - 988, Vo. 2, Jeju, Korea, April 2003

# Chapter 1

# Introduction

## 1.1 Background

The last decade has seen tremendous growth in wireless communications. The data rate of mobile communication networks has evolved from 9.6 kilobits per second (kbps) of the early second generation GSM (Global System for Mobile Communications) network, to 21.4 kbps of GPRS (General Packet Radio Service), and 69.2 kbps of Extended GPRS (EGPRS) using EDGE (Enhanced Data Rate for GSM Evolution) technology. GPRS and EDGE are also classified as the "*2.5 Generation*" mobile networks in contrast to the third generation (3G) code division multiple access (CDMA) networks which can offer 384 kbps for high mobility users and 2 megabits per second (mbps) for pedestrians. The 3GPP (Third Generation Partnership Project) is working on the standard specification for delivering data services up to 10 mbps for data users, and it is predicted that for fourth generation mobile networks, the data rate has to reach 100 mbps for high mobility users and one gigabits per second (gbps) for users in hot spots. The technology and bandwidth advancement has also attracted significant increase in the number of subscribers. According to the International Telecommunication Union (ITU) record, the worldwide mobile phone subscribers by the middle of 2004 have reached 1.5 billion, which is about 25% of the world's population.

Similar to the cellular mobile communications, the data rate offered by wireless local area network (WLAN) has also grown by about 50 times over the last decade, from 1mbps of the early IEEE (Institute of Electrical & Electronic Engineers) 802.11 [1], to 11 mbps of IEEE 802.11b [1], and to 54 mbps of today's IEEE 802.11a [2] and 11g [3] systems. Currently, the IEEE 802.11 task group n (TGn) is working toward a standard to offer as high as 600 mbps WLAN system [4].

Wireless communication has become a seamless (or inseparable) part of people's life style. Getting connected anywhere and anytime is no longer just a dream.

Wireless communication system design, however, remains challenging. As predicted by the Edholm's law of data rates [5], the bandwidth of a communication system, wireless or wireline, is to increase exponentially with time until some fundamental human limit, for example, number of pixels per second the human eyeball can process, is reached at some point of time. The radio frequency (RF) bandwidth allocated by regulatory agencies, on the other hand, is limited and can not increase at a matched pace with the data rate requirement. Increasing the working signal to noise ratio (SNR) is another way of increasing data rate, as suggested by the Shannon channel capacity formula [6]. Wireless communication systems, however, are transmission power limited. Hence SNR can not be increased unlimitedly. Furthermore, data rate is a logarithm function of SNR. In the high SNR region, every 3dB SNR increase, or two times' transmission power, leads to an additional capacity of only 1 bps/Hz. Therefore, other means have to be found to fulfill the data rate demand.

In the mid 1990's, independent work from Foschini [7] and Telatar [8] showed that in a rich scattering environment, deploying multiple antenna arrays at both the transmitter and the receiver can create a multiple-input multiple-output (MIMO) channel. The MIMO channel capacity is linearly increased with the minimum number of the transmit and receive antennas. Foschini also recommended the Diagonal Bell LAboratories Space-Time (DBLAST) [9] and Vertical Bell LAboratories Space-Time (VBLAST) [10] systems to realize the capacity potential in the MIMO channel.

In addition to the continuously growing demand for higher data rate, another big challenge for wireless communications is the hostile channel the information is transmitted through. With reflections, diffractions, scattering in the radio propagation channel, constructive and destructive superposition of the reflected, diffracted or scattered paths results in received signal strength experiencing the phenomenon called "fading" [11]. Fading can be frequency selective, time selective, or doubly selective in both time and frequency. For wideband channels[1], the transmitted signals are further distorted by "multipath". Multiple replicas of the transmitted signals arrive at the receiver with different time delays and experience different attenuation and phase distortion. The detrimental intersymbol interference (ISI) caused by multipath is traditionally mitigated by equalization techniques [12]. Due to its effective ISI mitigation capability and its simple implementation, orthogonal frequency division multiplexing (OFDM) [13] [14] [15] has been widely adopted in wideband and broadband wireless communications. The wireless LAN IEEE 802.11a [2] and 802.11g [3], ETSI (European Telecommunications Standards Institute) HiperLAN/2 [16] all specify to use OFDM as the physical layer (PHY) solution.

To combat fading and provide reliable and robust performance, a wireless communication system has to rely on various "diversity" techniques. Traditional diversity techniques include:

**Time Diversity** Time diversity can be exploited from a time selective fading channel. Forward error correction (FEC) coding with interleaving is one popular time diversity scheme in which additional information (redundancy) is transmitted at different time instances that the channel is experiencing independent (or close to independent) fading. Diversity gains are achieved through de-interleaving and decoding [12]. Another time diversity technique which is less referred to is the automatic repeat request (ARQ) scheme [17] in which re-transmission is requested by the receiver to the transmitter through a feedback channel when it detects incorrect decoding of information. Depending on the ARQ schemes adopted by the network, either the same set of information or the re-encoded and re-packetized information is re-transmitted. The receiver will then perform either code combining or diversity combining [18] to recover the information. The incremental redundancy (IR) ARQ scheme [19] is also a time diversity scheme which transmits additional redundant information of an error correction code word to help correctly decode the original information sequence.

**Frequency Diversity** Frequency diversity is available for exploitation when the channel is experiencing frequency selective fading. Spread spectrum modulation exploits the frequency diversity through transmitting the raw information over a wide frequency in which each subbands experience independent fading. The receiver can achieve the diversity gain through maximal ratio combining

[1]Channels with bandwidth $BW$ wider than the coherence bandwidth is considered as "Wideband channels" [11].

(MRC) the independently faded signals over each subbands [20]. For OFDM modulated signals, the frequency diversity is exploited by using FEC coding and interleaving [15].

**(Receive) Space Diversity** Traditionally, space diversity is exploited at the receiver by using multiple receive antenna elements and combining algorithms such as MRC, equal gain combining (EGC), receive antenna selection (RAS), or receive antenna switching. Macro-cell diversity or soft handoff used in CDMA systems [21] is also a space diversity technique. Different combining algorithms have different level of complexity and lead to different level of diversity gains. As these techniques are realized solely at the receiver, we call them *receive* space diversity.

A system can exploit more than one type of diversity gains. For example, an OFDM system can use FEC coding and interleaving to exploit frequency diversity, ARQ scheme to exploit time diversity, and multiple receive antenna to exploit space diversity.

Time and frequency diversity techniques are realized at the cost of additional redundancy, be it the additional redundancy introduced by FEC coding in single carrier and OFDM systems, or the additional redundancy by transmitting a narrowband signal over a channel with much wider bandwidth in spread spectrum systems. Receiving space diversity does not cost any additional redundancy. Its realization, however, will depend on the availability of multiple antenna elements at the receiver, which may sometimes not be possible due to the size limitation of the wireless terminal. The base station, on the other hand, is not so size-constrained and hence can accommodate more antenna elements. Therefore, space diversity exploitation at the transmitter have to be explored.

In 1991, Wittneben proposed a base station modulation diversity approach in [22] to achieve diversity gains through transmitting the same information from different base stations. He further extended this work to transmit antenna diversity gain in [23]. J. Winters studied the transmit diversity gains in Rayleigh fading channels in [24] and showed that transmit diversity can achieve the same gain as the receive diversity. Publication of Tarokh *et. al.* on space-time code design in [25] started the years of active research in space-time code design and realization of *transmit* space diversities.

## 1.2 Focus of This Thesis

This thesis is concerned with in general the design of transmit and receive techniques for a MIMO-OFDM system in wideband frequency selective MIMO channels, and more specifically the appropriate space-time precoding schemes and transmitter and receiver designs for a MIMO-OFDM system in block fading multipath frequency selective channels. Several space-time pre-coding schemes are studied. Their ergodic and outage capacity performances are analyzed and their tradeoff between capacity and diversity gains is investigated. A two-dimensional linearly transformed MIMO-OFDM system is proposed to maximize the frequency and space diversity gains.

For the receiver, we focus on the iterative turbo receiver algorithms. Simplification of the soft decision functions have been proposed which introduce only marginal performance degradation. More importantly, a family of Bayesian minimum mean squared error (MMSE) turbo receivers are proposed. The proposed Bayesian turbo receivers can significantly improve the BER and FER performance over conventional turbo receivers, especially when punctured high rate error correction code (ECC) is used in the system. The proposed Bayesian turbo receivers can also improve the convergence speed, hence effectively reducing the processing delay. The extrinsic information transfer (EXIT) chart of the proposed Bayesian turbo receiver is derived and compared with the conventional turbo receivers. The EXIT chart analysis results verify the superior performance of the proposed Bayesian turbo receiver over the conventional receivers.

For coherent detection, channel state information is essential at the receiver. To accurately acquire the channel estimates, efficient training signal is required. The preamble design for training sequence assisted channel estimation is studied. Both the time domain and frequency domain channel estimation algorithms are looked into, and the corresponding preamble design is proposed which can optimize the mean squared error (MSE) of the channel estimates.

## 1.3 Thesis Organization

The rest of the thesis is organized as follows. In Chapter 2, the ergodic and outage capacity of the MIMO channel is reviewed, under the condition of perfect channel state information (CSI) available at either only the receiver but not at the transmitter, or both the transmitter and the receiver. Then an overview is given on the various space-time coding schemes, with the emphasis on the orthogonal space-time block codes (STBC), space-time trellis codes (STTC), and quasi-orthogonal space-time block codes (QSTBC). We also show analytically that when FEC code is serially concatenated with orthogonal STBC, additional diversity gain can be exploited if the channel is fast fading, or alternatively when the channel is slow fading, additional coding gain can be exploited. A brief discussion of the capacity and diversity tradeoff is also given in Chapter 2.

In Chapter 3, we formulate the linear signal model for MIMO OFDM systems. Various space-time-frequency precoding (STFP) techniques are considered. By combining precoding with the MIMO propagation channel, all the precoding schemes considered can be expressed by the common linear signal model. This unifies the capacity and diversity analysis in Chapter 4. It has also made the derivation of the turbo receiver algorithms in Chapter 5 applicable to all these precoded MIMO-OFDM systems.

Chapter 4 is dedicated to the capacity and diversity analysis of the various space-time precoded MIMO-OFDM channels. Both the ergodic capacity and the outage capacity with unconstrained complex Gaussian input signals are studied. The mutual information of the precoded channels for fixed-order modulation signals is also investigated. The mutual information knowledge will provide more realistic guidance for precoding scheme selection in practical systems. A two-dimensional linear pre-transformed (2DLPT) MIMO-OFDM system is proposed which can achieve full capacity and full diversity.

Chapter 5 is focussed on the study of iterative turbo receivers for coded MIMO-OFDM systems. It is further divided into two parts. The first part is dedicated to simplification of soft decision functions (SDF's) in conventional turbo receivers. In order to effectively realize the huge capacity of the MIMO-OFDM channels, higher order modulation, e.g., 8PSK, 16QAM, or 64QAM, signals need to be transmitted. The estimation of these high-order modulation signals with the soft output extrinsic information from the decoder, however, requires calculation of several exponential terms, hence complex in practical implementation. In view of this, simplified linear SDF's are derived which introduce negligible BER performance degradation, as demonstrated from simulations.

In the second part of Chapter 5, we propose a family of Bayesian turbo receivers. Different from the conventional turbo receivers, Bayesian signal estimation theory is used to estimate the interference signals. Hence both the *a priori* information, i.e., the extrinsic information from the decoder, and the observation, i.e., the received signal or filter output of the interference canceller, is used. As a result, the estimation accuracy of the interference signals is greatly improved. The improved estimation accuracy can lead to significant performance improvement, as shown through our simulated BER and FER results. Two types of filtering schemes have been considered in the interference cancellation (IC) process of the Bayesian turbo receiver, namely, the matched filtering (MF), i.e., the maximal ratio combining (MRC) filtering, and the linear MMSE (LMMSE) filtering. These two types of turbo receivers are referred to as the *IC-MRC* turbo receiver and the *LMMSE-IC* turbo receiver, respectively.

In Chapter 6, we derive the extrinsic information transfer (EXIT) chart of the proposed Bayesian turbo receivers and compare with that of the conventional turbo receivers. Our EXIT chart analysis shows that the Bayesian IC-MRC turbo receiver has superior performance to not only the conventional IC-MRC turbo receiver, but also the conventional LMMSE-IC turbo receiver. The performance improvement lies in two ways - the much higher output mutual information of the Bayesian detector, and the reduced number of iterations to achieve convergence in the turbo receiver. This result makes the Bayesian IC-MRC turbo receiver practically appealing. This is because MRC filtering performs only multiplication and summation, whereas the LMMSE filtering, on the other hand, required the much more complex operations of complex-valued matrix inversion for each signal stream at each iteration.

The capacity and diversity analysis of precoded MIMO-OFDM channels in Chapter 4, the Bayesian turbo receiver studies in Chapter 5 and Chapter 6 are all based on the assumption of perfect CSI available at the receiver. In Chapter 7, we study training signal-based CSI estimation. Both frequency domain and time domain channel estimation schemes are considered when designing the preamble sequence. Their corresponding mean squared error (MSE) is derived and used as the objective function for optimal training signal design. Two optimal training signal schemes are proposed for both the frequency and time domain channel estimation, supporting very simple channel estimation computation and lead to minimum MSE.

Chapter 8 concludes the work reported in this dissertation. Recommendation for further continuation of the research work in this dissertation is also given in this Chapter.

## 1.4 Contributions of This Thesis

The major original contributions of this thesis are summarized below.

- Studied systematically the capacity and diversity performance of the various open-loop space-time-frequency precoded MIMO-OFDM systems. In particular, we derived the ergodic capacity of spatial spreading MIMO systems by making use of the random matrix theory.

- Proved that cyclic delay transmission in MIMO-OFDM systems transfers the spatial diversity to frequency diversity by making use of the linear algebraic model of OFDM systems.

- Proposed a two-dimensional linear pre-transformed MIMO-OFDM system structure which can achieve full capacity and full diversity;

- Proposed the linear soft decision functions for high-order modulation signals in turbo receivers which can significantly reduce the computational complexity in signal estimation but at the same time maintain the BER performance;

- Proposed the Bayesian turbo receivers which makes use of both the extrinsic information from the soft output decoder and the soft output from the detector to obtain the Bayesian estimate of the interference signals. The Bayesian signal estimation is further extended to the LMMSE-IC turbo receivers. Significant performance improvement is obtained from the Bayesian turbo receivers;

- Developed the EXIT chart analytical model of the Bayesian turbo receivers. With this model, the EXIT chart is derived and compared with the conventional turbo receivers. From the EXIT chart analysis, the superior performance in terms of both higher output mutual information and the reduced number of iterations for convergence is proved;

- Systematically studied the training signal design for both frequency-domain and time-domain channel estimation in MIMO-OFDM systems. With the objective of minimum mean squared error, two preambles schemes, i.e., the orthogonal training signal and the switched-subcarrier training signal, are proposed for frequency domain channel estimation. Similarly, two preambles schemes are also derived for minimum mean squared error time-domain channel channel estimation, i.e., the switched-subcarrier training signal and cyclic delayed training signal. All the four training signal schemes involve very simple filtering calculation to obtain the channel estimates.

## 1.5 Notations

Throughout the rest of the thesis, unless otherwise mentioned, the time domain data are represented with lower-case, frequency-domain data with upper-case, vectors and matrices with bold face letters. The symbols $(\cdot)^T$, $(\cdot)^H$, and $(\cdot)^{-1}$ represent matrix transposition, Hermitian, and inversion, respectively, and the delimiter $(\cdot)^y$ defines a space of dimension $y$. All vectors are defined as column vectors with row vectors represented by transposition.

# Chapter 2

# Introduction to MIMO

Space-time (ST) processing is one of the most active research areas in wireless communications during the last decade, covering the theoretical aspects of capacity limit of a MIMO channel, performance limit of a space-time system, ST coding/decoding and modulation/demodulation techniques, and solutions to integrate the technology into practical systems. In this chapter, we give a general overview of the space-time transmission techniques. We start from the definition of a MIMO channel model. We then derive its capacity limit with CSI perfectly known at only the receiver but not known at the transmitter. We also briefly discuss the capacity limit when the CSI is perfectly known at both the transmitter and the receiver. Following the capacity discussions, we give an introduction to the MIMO diversity techniques. The various space-time codes are reviewed, and their decoding schemes are compared. Finally we conclude the chapter by discussing the capacity and diversity gain trade-off in the MIMO channels.

## 2.1 The MIMO Channel Model

A narrowband flat fading MIMO channel with $n_T$ transmit and $n_R$ receive antennas is defined as

$$\mathbf{y} = \mathbf{h}\mathbf{x} + \mathbf{n}, \tag{2.1}$$

where $\mathbf{y} \in \mathcal{C}^{n_R}$ and $\mathbf{x} \in \mathcal{C}^{n_T}$ are the complex-valued channel output and input signals, $\mathbf{n} \in \mathcal{C}^{n_R}$ denotes the zero mean complex additive white Gaussian noise (AWGN) with variance $\sigma^2$ per real dimension, i.e., $\mathbf{n} \sim CN(0, 2\sigma^2\mathbf{I}_{n_R})$, and $\mathbf{h} \in \mathcal{C}^{n_R \times n_T}$ with its entries $\{h_{ij}\}$ denoting the complex-valued fading coefficients corresponding to transmit antenna $j$ and receive antenna $i$. Fig. 2.1 depicts a simple illustration of such a $n_T \times n_R$ MIMO channel.

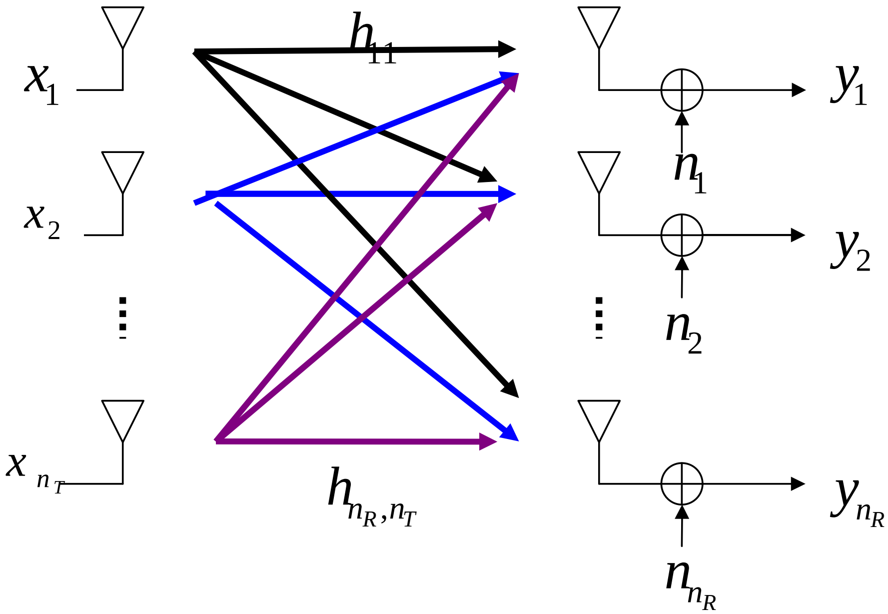


**Figure 2.1: Illustration of a narrowband $n_T \times n_R$ MIMO channel model.**

The MIMO channels can be divided into three categories:

**Deterministic Channel.** $h_{ij}$'s are deterministic values.

**Ergodic Channel.** $h_{ij}$'s are random variables, and each channel use corresponds to an independent realization of $h_{ij}$'s.

**Non-Ergodic Channel.** $h_{ij}$'s are random variables, but remain fixed once they are chosen.

Among the three channels, the last two are of more interest for MIMO communication systems design. Their corresponding suitable capacity measures are respectively the *ergodic capacity*, and the *outage capacity*. The reason to use *ergodic capacity* to measure an *ergodic channel* is due to the fact that a long enough code word transmitted in an *ergodic channel* will experience all states of the channel and hence it averages out the channel randomness. As for *non-ergodic channel*, a code word can only experience one channel realization no matter how long it is. The *outage capacity* is therefore defined as the rate such that there exists a code which can achieve with a pre-defined error probability for a set of channels. In Chapter

4, both ergodic capacity and outage capacity will be studied for precoded MIMO-OFDM channels.

## 2.2 Channel Capacity with CSI Perfectly Known Only at Receiver

When the CSI is perfectly known at the receiver but not known at the transmitter, we first look at the capacity for each channel realization by performing singular value decomposition (SVD) on the channel matrix $\mathbf{h}$ as

$$\mathbf{h} = \mathbf{\Omega}\mathbf{\Sigma}\mathbf{\Gamma}^H, \tag{2.2}$$

where $\mathbf{\Omega} \in \mathcal{C}^{n_R \times n_R}$ and $\mathbf{\Gamma} \in \mathcal{C}^{n_T \times n_T}$ are unitary matrices, and

$$\mathbf{\Sigma} = \mathrm{diag}\{\sigma_1,\ \cdots,\ \sigma_r,\ 0,\ \cdots,\ 0\} \in \Re^{n_R \times n_T}$$

is the singular value matrix of $\mathbf{h}$ whose rank is assumed to be $r = \min\{n_R,\ n_T\}$.

(2.1) can therefore be re-written as

$$\mathbf{y} = \mathbf{\Omega}\mathbf{\Sigma}\mathbf{\Gamma}^H\mathbf{x} + \mathbf{n}. \tag{2.3}$$

Pre-multiplying (2.3) with $\mathbf{\Omega}^H$, we have

$$\tilde{\mathbf{y}} = \mathbf{\Sigma}\mathbf{\Gamma}^H\mathbf{x} + \tilde{\mathbf{n}}, \tag{2.4}$$

where $\tilde{\mathbf{y}} = \mathbf{\Omega}^H\mathbf{y}$, and $\tilde{\mathbf{n}} = \mathbf{\Omega}^H\mathbf{n}$, and $\tilde{\mathbf{n}} \sim CN\left(\mathbf{0}, 2\sigma^2\mathbf{I}\right)$. If we further define $\tilde{\mathbf{x}} = \mathbf{\Gamma}^H\mathbf{x}$, (2.1) is turned into

$$\tilde{\mathbf{y}} = \mathbf{\Sigma}\tilde{\mathbf{x}} + \tilde{\mathbf{n}}, \tag{2.5}$$

which is effectively $n_R$ parallel single-input single-output channels

$$\begin{cases} \tilde{y}_i = \sigma_i\tilde{x}_i + \tilde{n}_i, & i = 1, 2, \cdots, r, \\ \tilde{y}_i = \tilde{n}_i, & i = r+1, \cdots, n_R. \end{cases} \tag{2.6}$$

When the transmitter has no knowledge on $\mathbf{h}$, allocating the transmission power equally to the $n_T$ transmit antennas will lead to maximum capacity [26] [8]. Supposing we normalize the total transmit power to unity, we have

$$\mathcal{E}\left\{\mathbf{x}\mathbf{x}^H\right\} = \frac{1}{n_T}\mathbf{I}_{n_T}. \tag{2.7}$$

Each parallel single-input single-output channel will then achieve capacity when the channel input $x_i$ is Gaussian [6]

$$\begin{aligned} C_i &= \log_2(1+\rho_i), \quad i = 1, 2, \cdots, r, \\ C_i &= 0, \qquad\qquad\qquad i = r+1, \cdots, n_R, \end{aligned} \tag{2.8}$$

where $\rho_i = \frac{1}{n_T}\frac{\sigma_i^2}{N_o} = \frac{1}{n_T}\frac{\lambda_i}{N_o}$ is the SNR at the channel output, and $\lambda_i = \sigma_i^2$ is the $i$th eigenvalue of matrix $\mathbf{h}\mathbf{h}^H$ and $\mathbf{h}^H\mathbf{h}$.

The MIMO channel capacity for each channel realization is thus

$$C = \sum_{i=1}^{r} C_i = \sum_{i=1}^{r} \log_2(1+\rho_i) = \sum_{i=1}^{r} \log_2\left(1 + \frac{\lambda_i}{n_T N_o}\right) \text{ (bits/channel use)}, \tag{2.9}$$

As matrices $\mathbf{h}\mathbf{h}^H$ and $\mathbf{h}^H\mathbf{h}$ have the same eigenvalues, the $n_R \times n_T$ MIMO channel $\mathbf{h}$ and the $n_T \times n_R$ MIMO channel $\mathbf{h}^H$ have the same capacity if the *receive SNR* is set to the same. This property is called "reciprocity" by Telatar [8].

As

$$\sum_{i=1}^{r} \log_2\left(1 + \frac{\lambda_i}{n_T N_o}\right) = \log_2\left[\prod_{i=1}^{r}\left(1 + \frac{\lambda_i}{n_T N_o}\right)\right] = \log_2 \det\left(\mathbf{I} + \frac{1}{n_T N_o}\mathbf{h}\mathbf{h}^H\right),$$

the MIMO capacity per realization is also written as [7][8]

$$C(\mathbf{h}) = \log_2 \det\left(\mathbf{I} + \frac{1}{n_T N_o}\mathbf{h}\mathbf{h}^H\right) \text{ (bits/channel use)}. \tag{2.10}$$

### 2.2.1 Ergodic Capacity

The ergodic capacity is defined as

$$C_{\mathrm{E}} = \mathcal{E}\left\{C(\mathbf{h})\right\}, \tag{2.11}$$

where the expectation is taken over all the realizations of $\mathbf{h}$. Analytical evaluation of $C_{\mathrm{E}}$ requires the statistics of $\mathbf{h}$, or eigenvalues $\{\lambda_i\}$ of $\mathbf{h}\mathbf{h}^H$. If the joint probability density function (pdf) of $\{h_{ij}\}$, $p\left(h_{11},\ h_{12},\ \cdots,\ h_{n_R,n_T}\right)$, is known, the ergodic capacity is obtained as

$$C_{\mathrm{E}} = \underbrace{\int_{h_{11}} \cdots \int_{h_{n_R,n_T}}}_{n_R \times n_T} \log_2 \det\left(\mathbf{I} + \frac{1}{n_T N_o}\mathbf{h}\mathbf{h}^H\right) p\left(h_{11},\ \cdots,\ h_{n_R,n_T}\right) dh_{11} \cdots dh_{n_R,n_T}. \tag{2.12}$$

Alternatively, if the joint pdf of $\{\lambda_i\}$, $p(\lambda_1, \cdots, \lambda_r)$, is known, the ergodic capacity can also be obtained as

$$C_{\mathrm{E}} = \int_{\lambda_1} \cdots \int_{\lambda_r} \sum_{i=1}^{r} \log_2\left(1 + \frac{\lambda_i}{n_T N_o}\right) p(\lambda_1, \cdots, \lambda_r)\, d\lambda_1 \cdots d\lambda_r. \tag{2.13}$$

For circularly symmetric complex Gaussian (CSCG) channels with $h_{ij} \sim \mathcal{CN}(0, 1)$ [8], the joint pdf of unordered eigenvalues $\{\lambda_i\}$ is given in [27] as

$$p(\lambda_1, \cdots, \lambda_r) = \frac{2^{-rs}\pi^{r(r-1)}}{r!\tilde{\Gamma}_r(s)\tilde{\Gamma}_r(r)} \exp\left(-\frac{1}{2}\sum_{i=1}^{r}\lambda_i\right) \prod_{i=1}^{r} \lambda_i^{s-r} \prod_{i<j} (\lambda_i - \lambda_j)^2 \tag{2.14}$$

where $s = \max\{n_R, n_T\}$, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r$, and $\tilde{\Gamma}_m(a)$ is the complex multivariate gamma function defined by

$$\tilde{\Gamma}_m(a) = \pi^{m(m-1)/2} \prod_{i=1}^{m} \Gamma(a - i + 1),$$

with $\Gamma(a)$ being the gamma function.

Based on (2.14), Telatar worked out the ergodic capacity of $n_T \times n_R$ CSCG channel as

$$\int_0^{\infty} \log_2\left(1 + \frac{\lambda}{n_T N_o}\right) \sum_{k=0}^{r-1} \frac{k!}{(k+s-r)!} \left[L_k^{s-r}(\lambda)\right]^2 \lambda^{s-r} \exp(-\lambda) d\lambda, \tag{2.15}$$

where $r = \min\{n_R, n_T\}$, $s = \max\{n_R, n_T\}$, and

$$L_k^{n-m}(x) = \frac{1}{k!} \exp(-x) x^{m-n} \frac{d^k}{dx^k}(\exp(x) x^{k+n-m})$$

is the Laguerre polynomial of order $k$ [28].

In Appendix 2A of this Chapter, we give the Laguerre polynomials and the ergodic capacity formulas for the CSCG MIMO channels that are going to studied in Chapter 4.

**Linear Increase of MIMO Capacity with $r$** From the strong law of large numbers, we have for fixed $n_R$ and as $n_T \rightarrow \infty$

$$\frac{1}{n_T}\mathbf{h}\mathbf{h}^H \rightarrow \mathbf{I}_{n_R}$$

hence from (2.10), we have

$$C(\mathbf{h}) = \log_2 \det\left(\mathbf{I} + \frac{1}{N_o}\mathbf{I}\right) = n_R \log_2\left(1 + \frac{1}{N_o}\right) \text{ (bits/channel use)}, \tag{2.16}$$

i.e., the capacity increases linearly with $n_R$.

When we fix $n_T$ and make $n_R \rightarrow \infty$, in order to prove the linear relation between capacity $C$ and the number of antennas $n_T$, we need to scale the channel matrix as $\frac{1}{\sqrt{n_R}}\mathbf{h}$. Without this scaling, the receive SNR will grow to infinity. The channel capacity is then

$$\begin{aligned} C(\mathbf{h}) &= \log_2 \det\left(\mathbf{I} + \frac{1}{n_T n_R N_o}\mathbf{h}\mathbf{h}^H\right) = \log_2 \det\left(\mathbf{I} + \frac{1}{n_T n_R N_o}\mathbf{h}^H\mathbf{h}\right) = \log_2 \det\left(\mathbf{I} + \frac{1}{n_T N_o}\mathbf{I}\right) \\ &= n_T \log_2\left(1 + \frac{1}{n_T N_o}\right) \text{ (bits/channel use)}, \end{aligned} \tag{2.17}$$

by making use of the fact that $\frac{1}{n_R}\mathbf{h}^H\mathbf{h} \rightarrow \mathbf{I}_{n_T}$ when $n_R \rightarrow \infty$, from the strong law of large numbers.

### 2.2.2 Outage Capacity

For non-ergodic channels, the Shannon capacity is zero. This is because no matter how long a code word we can take, there is a non-zero probability that the realized $\mathbf{h}$ is incapable of supporting however a small rate. Therefore, the *outage capacity* is a more appropriate measure which is defined as the transmission rate $R$ that exceeds the instantaneous channel capacity

$$C(\mathbf{h}) = \log_2 \det\left(\mathbf{I} + \frac{1}{n_T N_o}\mathbf{h}\mathbf{h}^H\right)$$

with probability $P$. $P$ is defined as the *outage probability* [8], i.e.,

$$P_{\text{out}}(R) = p\left(R > C(\mathbf{h})\right). \tag{2.18}$$

Equation (2.18) can be evaluated by Monte Carlo simulations. For approximation, the asymptotic result of $C(\mathbf{h})$ tending to a Gaussian random variable when $n_R$ and $n_T$ grow to infinity, can be used. The details can be referred to [29][30].

## 2.3 Channel Capacity with CSI Perfectly Known at Both Transmitter and Receiver

When both the transmitter and the receiver have perfect CSI, by using the result of Information Theory concerning parallel Gaussian channels [26][6], from (2.6), we need to allocate the transmission power to the $r$ parallel channels via "water-filling". Supposing the power allocated to the $i$th parallel channel is

$$P_i = 2\mathcal{E}\{\text{Re}(\tilde{x}_i)\}^2 = 2\mathcal{E}\{\text{Im}(\tilde{x}_i)\}^2 \tag{2.19}$$

subject to the power constraint

$$\sum_{i=1}^{r} P_i = 1,$$

then the mutual information between the input and the output of the channel is

$$I(\tilde{\mathbf{x}};\tilde{\mathbf{y}}) \leq \sum_{i=1}^{r} \log_2\left(1+\frac{P_i\lambda_i}{n_T N_o}\right).$$

Maximization of the mutual information subject to the power constraint leads to the channel capacity. It can be solved by using the Lagrange multipliers, through defining the cost function

$$J(P_1,\ P_2,\ \cdots\ ,\ P_r) = \sum_{i=1}^{r} \log\left(1+\frac{P_i\lambda_i}{n_T N_o}\right) + \nu(1-\sum_{i=1}^{r} P_i)$$

and differentiating with respect to $P_i$, and we have

$$\frac{\lambda_i}{n_T N_o + P_i\lambda_i} - \nu = 0$$

leading to

$$P_i = \frac{1}{\nu} - \frac{n_T N_o}{\lambda_i}.$$

As $P_i$ needs to be non-negative, we therefore have the following “water-filling” solution

$$P_i = \left(\mu - \frac{n_T N_o}{\lambda_i}\right)^+ \tag{2.20}$$

satisfying the power constraint $\sum_{i=1}^{r} P_i = 1$. $\mu$ is called the “water level” and $(x)^+$ is defined as

$$(x)^+ = \begin{cases} x & \text{if } x \geq 0, \\ 0 & \text{if } x < 0. \end{cases}$$

The principle of water-filling for an $r = 4$ MIMO channel is illustrated in Fig. 2.2.

## 2.4 MIMO Diversity and Space-Time Codes

Besides capacity gain, the MIMO channels can also be used to exploit diversity gains and improve the robustness of wireless communication systems against fading. This is achieved by transmitting space-time coded signals through the $n_T$ antennas, and processing the received signals at the $n_R$ antennas by maximal ratio combining (MRC) and maximum likelihood (ML) decoding.

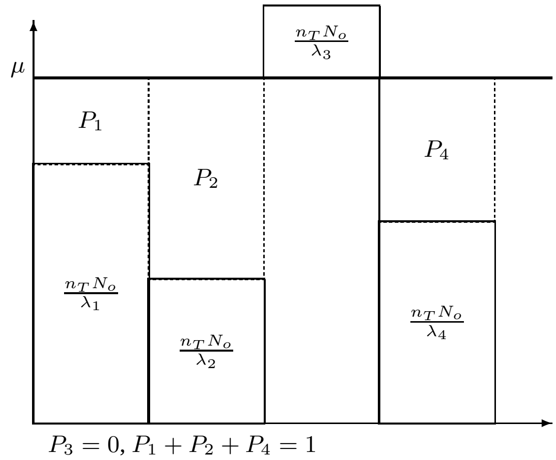


**Figure 2.2: Illustration of "water-filling" principle.**

Supposing the space-time encoder takes in $M$ bits and produce an $L$-symbol-long space-time codeword, as

$$\mathbf{x} = [\mathbf{x}_1 \ \mathbf{x}_2 \ \cdots \ \mathbf{x}_L]$$

with $\mathbf{x}_l = [x_{l,1} \ x_{l,2} \ \cdots \ x_{l,n_T}]^T$, we have the corresponding received signal as

$$\mathbf{y}_l = \mathbf{h}_l \mathbf{x}_l + \mathbf{n}_l \tag{2.21}$$

for a fast fading channel and

$$\mathbf{y}_l = \mathbf{h} \mathbf{x}_l + \mathbf{n}_l \tag{2.22}$$

for a quasi-static fading channel. By *"fast fading"*, we mean the channel coefficients remain constant during one symbol interval, but vary randomly from one symbol to another; By *"quasi-static fading"*, we mean the channel coefficients remain constant during a frame but vary randomly from one frame to another. In other words, for a fast fading channel, the coherence time is longer than the symbol interval but shorter than the space-time codeword interval, and for a quasi-static fading channel, the coherence time is longer than the space-time codeword duration.

When perfect CSI is available at the receiver, we have

$$p(\mathbf{y}|\mathbf{x}) = \prod_{l=1}^{L} \frac{1}{(2\pi\sigma^2)^{n_R}} \exp\left(-\frac{|\mathbf{y}_l - \mathbf{h}_l \mathbf{x}_l|^2}{2\sigma^2}\right) \tag{2.23}$$

for fast fading channel, and

$$p(\mathbf{y}|\mathbf{x}) = \prod_{l=1}^{L} \frac{1}{(2\pi\sigma^2)^{n_R}} \exp\left(-\frac{|\mathbf{y}_l - \mathbf{h}\mathbf{x}_l|^2}{2\sigma^2}\right) \tag{2.24}$$

for a slow quasi-static fading channel.

The ML decision of the transmitted space-time codeword for fast-fading channel is thus

$$\begin{aligned}
\hat{\mathbf{x}}_{\mathrm{ML}} &= \arg \max_{\mathbf{x}\in\Omega^{n_T L}} p(\mathbf{y}|\mathbf{x}) && (2.25)\\
&= \arg \max_{\mathbf{x}\in\Omega^{n_T L}} \prod_{l=1}^{L} \frac{1}{(2\pi\sigma^2)^{n_R}} \exp\left(-\frac{|\mathbf{y}_l - \mathbf{h}_l\mathbf{x}_l|^2}{2\sigma^2}\right) && (2.26)\\
&= \arg \max_{\mathbf{x}\in\Omega^{n_T L}} \sum_{l=1}^{L} \left(-|\mathbf{y}_l - \mathbf{h}_l\mathbf{x}_l|^2\right) && (2.27)\\
&= \arg \min_{\mathbf{x}\in\Omega^{n_T L}} \sum_{l=1}^{L} |\mathbf{y}_l - \mathbf{h}_l\mathbf{x}_l|^2 , && (2.28)
\end{aligned}$$

where $\Omega$ denotes the modulation signal set for the space-time code in use.

Similarly, the ML decision of the transmitted space-time codeword for quasi-static fading channel is

$$\hat{\mathbf{x}}_{\mathrm{ML}} = \arg \min_{\mathbf{x}\in\Omega^{n_T L}} \sum_{l=1}^{L} |\mathbf{y}_l - \mathbf{h}\mathbf{x}_l|^2 . \tag{2.29}$$

Similar to error control codes, space-time codes can be categorized into block codes, trellis codes, and turbo codes. Different codes have different diversity and coding gains, and different decoding complexity. In this section, we will give a brief overview of orthogonal space-time block codes (STBC), space-time trellis codes (STTC), and quasi-orthogonal STBC (QSTBC).

### 2.4.1 Orthogonal STBC

The *"Orthogonal STBC"*, or OSTBC, encoding is a non-linear mapping, which takes input sequence $\{s_1,\ s_2,\ \cdots,\ s_Q\}$ and maps to a row-orthogonal matrix $\mathbf{x}_{n_T\times L}$, i.e.,

$$\mathbf{x}_{n_T\times L} = \mathcal{M}_{\mathrm{OSTBC}}\left(s_1,\ s_2,\ \cdots,\ s_Q\right),$$

and

$$\mathbf{x}\mathbf{x}^H = \alpha\mathbf{I}_{n_T}, \tag{2.30}$$

where $\alpha$ is a constant that is related to the total signal power transmitted by the STBC codeword. The code rate of STBC is defined as

$$R_{\mathrm{STBC}} = \frac{Q}{L}. \tag{2.31}$$

The most popular OSTBC is the *"Alamouti Code"* (AC) for $n_T = 2$ [31], whose mapping is defined as

$$\mathcal{M}_{\mathrm{AC}}\left(s_1,\ s_2\right) = \begin{bmatrix} s_1 & -s_2^* \\ s_2 & s_1^* \end{bmatrix} \tag{2.32}$$

for both real and complex signals. The code rate of AC is $R_{\mathrm{AC}} = 1$. It has been proven in [32] that the $2 \times 1$ AC is capacity optimal.

While rate-1 OSTBC is available for $n_T = 2$ for both real and complex signals, it is not the case for $n_T > 2$. Tarokh *et. al.* have constructed rate-1 real OSTBC's for $n_T \leq 8$ with entries of the form $\pm s_1,\ \pm s_2,\ \cdots,\ \pm s_Q$ in [33]. But they showed that rate-1 complex OSTBC exists only for $n_T = 2$, i.e., the AC. For $n_T = 3$ and $n_T = 4$ cases, rate $\frac{1}{2}$ and rate $\frac{3}{4}$ OSTBC's are given by Tarokh *et. al.* in [33].

The most attractive advantage of OSTBC is its full diversity order of $n_R n_T$, and its simple ML decoding by linear processing. As given in [34], the ML decision metric for each signal $s_q$, $q = 1,\ \cdots,\ Q$, can be decoupled and optimized individually. The drawback of OSTBC is its limited "coding" gain. But this can be remedied by concatenating a FEC code before the STBC encoding. Fig. 2.3 depicts a concatenated bit-interleaved coded modulation (BICM) [35] STBC transmission system. In Appendix 2B, we prove that when a FEC code is concatenated with the Alamouti STBC, and when the channel is quasi-static over the STBC codeword interval, but changing independently from one STBC codeword to another, a diversity order of $2d_{\min}n_R$ is achieved, of which $2n_R$ is from the spatial domain and $d_{\min}$, which is also the minimum Hamming distance of the FEC code, is from the time domain (or FEC coding domain). When the channel is quasi-static over both the STBC and the FEC codeword intervals, the system achieves the diversity order of $2n_R$, and coding gain of $d_{\min}$.

### 2.4.2 STTC

OSTBC can achieve maximum diversity order of $n_R n_T$ with simple linear decoding. However, OSTBC alone does not have any or very limited coding gain. The STTC, on the other hand, is designed to achieve

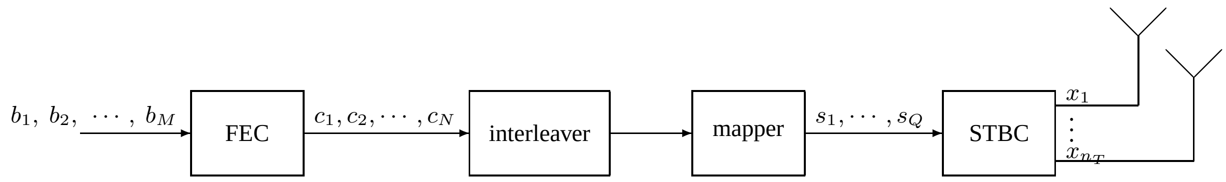


**Figure 2.3: Illustration of a concatenated BICM-STBC transmitter.**

both the maximum possible diversity gain and coding gain [25]. The name of "trellis code" comes from the fact that the encoding process can be represented by a trellis [25].

The optimal STTC design is derived based on the pair-wise error probability (PEP), which is defined as the probability that the decoder selects an erroneous codeword $\hat{\mathbf{x}}$ instead of the transmitted codeword $\mathbf{x}$. If we assume perfect CSI at the receiver, we have the PEP written as

$$\begin{aligned} P(\mathbf{x},\ \hat{\mathbf{x}}|\mathbf{h}) &= Prob\left(p(\mathbf{y}|\mathbf{x},\mathbf{h}) < p(\mathbf{y}|\hat{\mathbf{x}},\mathbf{h})\right) = Prob\left(\|\mathbf{y}-\mathbf{h}\mathbf{x}\|^2 > \|\mathbf{y}-\mathbf{h}\hat{\mathbf{x}}\|^2\right) \\ &= Prob\left(\sum_{l=1}^{L}\sum_{i=1}^{n_R}\left\|y_i - \sum_{j=1}^{n_T} h_{i,j}^l x_j^l\right\|^2 > \sum_{l=1}^{L}\sum_{i=1}^{n_R}\left\|y_i - \sum_{j=1}^{n_T} h_{i,j}^l \hat{x}_j^l\right\|^2\right) \\ &= Prob\left[\underbrace{\sum_{l=1}^{L}\sum_{i=1}^{n_R}\left(\left\|y_i - \sum_{j=1}^{n_T} h_{i,j}^l x_j^l\right\|^2 - \left\|y_i - \sum_{j=1}^{n_T} h_{i,j}^l \hat{x}_j^l\right\|^2\right)}_{d_{\mathbf{h}}^2(\mathbf{x},\hat{\mathbf{x}})} > 0\right] \\ &= Q\left(\sqrt{\frac{E_s}{2N_o} d_{\mathbf{h}}^2(\mathbf{x},\hat{\mathbf{x}})}\right) \\ &\leq \frac{1}{2}\exp\left(-d_{\mathbf{h}}^2(\mathbf{x},\hat{\mathbf{x}})\frac{1}{4N_o n_T}\right) \end{aligned}$$

where $E_s = \frac{1}{n_T}$ is the symbol energy at each transmit antenna, and $Q(\cdot)$ is the complementary error function [12].

Depending on the fading channel models, e.g., Rayleigh or Rician fading, slow or fast fading, number of receive antennas, etc., different criteria may have to be used to maximize both the coding gain and the diversity gain. In [25], Tarokh *et. al.* first developed the "*rank criterion*" to maximize the diversity gain, and the "*determinant criterion*" to optimize the coding gain for slow Rayleigh fading channels. As for slow Rician fading channels, the "*rank criterion*" and the "*coding advantage criterion*" is derived to maximize the diversity and coding gains. For fast fading Rayleigh channels, the design criteria become the "*distance*" and the "*product*" criteria for diversity and coding gains, respectively. A comprehensive summary of STTC design criteria in various channels can be found in [36].

As the STTC encoder structure can not guarantee the geometrical uniformity of the code [37], the search for the optimal encoder trellis has to be conducted over all possible pairs of paths in the code trellis. The decoding complexity of STTC is exponential with the code word length, the number of transmit antennas, the modulation order, and the number of states in the trellis. Due to all these issues, adoption of STTC in practical systems falls behind OSTBC.

### 2.4.3 Quasi-Orthogonal STBC (QSTBC)

Orthogonal STBC has the advantage of full transmit diversity order, simple and decoupled decoding for each symbol, and easy concatenation with FEC for further coding gain and time-diversity exploitation. However, rate-1 OSTBC is only available for $n_T = 2$ with complex signals. Therefore, quasi-orthogonal STBC (QSTBC) was proposed by Jafarkhani in [38] which can achieve full-rate (rate-1) but only half the maximum transmit diversity. Same as OSTBC, the QSTBC encoding is a non-linear mapping, which can be written as

$$\mathbf{x}_{n_T \times L} = \mathcal{M}_{\text{QSTBC}}\left(s_1,\ s_2,\ \cdots,\ s_Q\right),$$

where $\{s_1,\ s_2,\ \cdots,\ s_Q\}$ is the input sequence and $\mathbf{x}_{n_T \times L}$ is the corresponding QSTBC codeword. The original QSTBC proposed by Jafarkhani in [38] has the coding rate of 1, i.e., $Q = L$, and full transmit diversity.

The QSTBC decoding can be decoupled into groups of symbols instead of single symbols. Therefore the complexity is higher than OSTBC.

One example of a rate-1 Jafarkhani-QSTBC for $n_T = 4$ is given as follows:

$$\mathbf{x} = \begin{bmatrix} x_1 & -x_2^* & x_3 & -x_4^* \\ x_2 & x_1^* & x_4 & x_3^* \\ x_3 & -x_4^* & x_1 & -x_2^* \\ x_4 & x_3^* & x_2 & x_1^* \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 \\ \mathbf{A}_2 & \mathbf{A}_1 \end{bmatrix}, \tag{2.33}$$

where $\mathbf{A}_i,\ i = 1,\ 2$ denotes the $i$th Alamouti codeword. There are many variations of the Jafarkhani-QSTBC design in (2.33), as indicated in [38]. But they all have the same performance.

In order to achieve full-rate and full-diversity simultaneously, QSTBC based on constellation rotation was proposed in [39] [40]. In this scheme, the two groups of symbols are drawn from the original constellation and the rotated constellation, respectively. With the optimal rotation angle, full transmit diversity can always be made possible.

If the restriction of full transmit diversity is relaxed, high-rate QSTBC can be designed, e.g., Yuen *et. al.* proposed a rate-2 QSTBC for $n_T = 4$ in [41] as

$$\mathbf{x} = \sqrt{\frac{1}{2}} \begin{bmatrix} x_1 + x_3 & -x_2^* - x_4^* \\ x_2 + x_4 & x_1^* + x_3^* \\ x_1 - x_3 & -x_2^* + x_4^* \\ x_2 - x_4 & x_1^* - x_3^* \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 + \mathbf{A}_2 \\ \mathbf{A}_1 - \mathbf{A}_2 \end{bmatrix}, \tag{2.34}$$

and a rate-4 QSTBC for $n_T = 4$ as

$$\mathbf{x} = \begin{bmatrix} x_1 + x_3 + x_5^* + x_7^* & -x_2^* - x_4^* + x_6 + x_8 \\ x_2 + x_4 + x_6^* + x_8^* & x_1^* + x_3^* - x_5 - x_7 \\ x_1 - x_3 + x_5^* - x_7^* & -x_2^* + x_4^* + x_6 - x_8 \\ x_2 - x_4 + x_6^* - x_8^* & x_1^* - x_3^* - x_5 + x_7 \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 + \mathbf{A}_2 + \mathbf{A}_3^* + \mathbf{A}_4^* \\ \mathbf{A}_1 - \mathbf{A}_2 + \mathbf{A}_3^* - \mathbf{A}_4^* \end{bmatrix}, \tag{2.35}$$

where $\mathbf{A}_i$, $i = 1,\ 2,\ 3,\ 4$ denotes the $i$th Alamouti codeword.

These QSTBC codes are sometimes also referred as linear dispersion codes (LDC) which were first proposed by Hassibi and Hochwald [32] due to the fact that the signals $x_q$, $l = 1,\ 2,\ \cdots,\ Q$ are transmitted over all the transmit antennas within the QSTBC codeword. In Chapter 4, the capacity and diversity performance of the rate-2 QSTBC and LDC MIMO-OFDM channels will be studied in detail.

## 2.5 Diversity and Capacity Tradeoff in MIMO Channels

As reviewed in the previous sections in this chapter, a MIMO channel can provide two types of gains - the diversity gain and the capacity gain. If the MIMO channel is used to transmit independent information streams in parallel (spatial multiplexing [42]), the data rate (capacity) of the system increases. Schemes to exploit the capacity potential include the various BLAST architectures [9, 10]. If the MIMO channel is used to transmit signals that carry the same information, for example, OSTBC and STTC, the diversity gain of the system increases. A MIMO channel can also be used to exploit both the capacity and diversity gains simultaneously. Hybrid systems, e.g., groupwise STBC (GSTBC) [43] [44] and groupwise STTC systems [45][46], are the most straightforward approaches in which the transmit antennas are divided into groups. Independent information is transmitted in different antenna groups to exploit the multiplexing gain; Within each antenna group, STBC or STTC is applied to exploit the transmit diversity gain. Besides the hybrid schemes, high-rate QSTBC and LDC [32][47] are space-time block codes which were designed based on some design criteria to achieve both multiplexing and diversity gains. For example, the LDC by Hassibi and Hochwald [32] was designed to maximize a given MIMO channel capacity. After the channel capacity is maximized, the diversity order will then be optimized.

In [48], Tse and Li provided a framework to show that there is a fundamental tradeoff between the capacity and diversity gains for a given $n_T \times n_R$ MIMO channel. If $r \leq \min(n_T,\ n_R)$ antennas are used to exploit the spatial multiplexing gain, then only the rest of the antennas can be used to exploit diversity gains. If higher diversity order is desired from the given MIMO channel, the transmission rate will have to be reduced correspondingly.

In Chapter 4, the capacity and diversity performance of precoded MIMO-OFDM channels will be studied. A 2DLPT precoding scheme for MIMO-OFDM systems will be proposed. We will show that when the 2DLPT transform in unitary, this precoding scheme can achieve simultaneously full capacity and full diversity.

## Appendix 2A - Ergodic Capacity for i.i.d. CSCG MIMO Channels

### 2A.1 $4 \times 2$

For a $4 \times 2$ i.i.d. CSCG MIMO channel, we have

$$n_R = 2, \quad n_T = 4, \quad m = 2, \quad n = 4.$$

We then have

$$\begin{cases} k=0, \quad L_0^2(x)=1, & \frac{k!}{(k+n-m)!}=\frac{1}{2}, \\ k=1, \quad L_1^2(x)=-x+3, & \frac{k!}{(k+n-m)!}=\frac{1}{6}, \end{cases}$$

hence the ergodic capacity for the $4 \times 2$ CSCG MIMO channel is

$$C_{4\times 2}=\int_0^{\infty}\log_2\left(1+\frac{\rho}{4}x\right)\left[\frac{1}{2}+\frac{1}{6}\,(x-3)^2\right]x^2\exp(-x)dx. \tag{2.36}$$

### 2A.2 $8 \times 4$

For a $8 \times 4$ i.i.d. CSCG MIMO channel, we have

$$n_R=4, \quad n_T=8, \quad m=4, \quad n=8.$$

We then have

$$\begin{cases} k=0, \quad L_0^4(x)=1, & \frac{k!}{(k+n-m)!}=\frac{1}{24}, \\ k=1, \quad L_1^4(x)=-x+5, & \frac{k!}{(k+n-m)!}=\frac{1}{120}, \\ k=2, \quad L_2^4(x)=\frac{1}{2}(30-12x+x^2), & \frac{k!}{(k+n-m)!}=\frac{1}{360}, \\ k=3, \quad L_3^4(x)=\frac{1}{6}(210-126x+21x^2-x^3), & \frac{k!}{(k+n-m)!}=1, \end{cases}$$

hence the ergodic capacity for the $8 \times 4$ CSCG MIMO channel is

$$\begin{aligned} C_{8\times 4} &= \int_0^{\infty}\log_2\left(1+\frac{\rho}{8}x\right)\left[\frac{1}{24}+\frac{(x-5)^2}{120}+\frac{\left(30-12x+x^2\right)^2}{1440}\right. \\ &\qquad \left.+\frac{\left(210-126x+21x^2-x^3\right)^2}{840\times 36}\right]x^4\exp(-x)dx. \end{aligned} \tag{2.37}$$

### 2A.3 $4 \times 4$

For a $4 \times 4$ i.i.d. CSCG MIMO channel, we have

$$n_R=4, \quad n_T=4, \quad m=4, \quad n=4.$$

We then have

$$\begin{cases} k=0, \quad L_0(x)=1, & \frac{k!}{(k+n-m)!}=1 \\ k=1, \quad L_1(x)=-x+1, & \frac{k!}{(k+n-m)!}=1, \\ k=2, \quad L_2(x)=\frac{1}{2}(2-4x+x^2), & \frac{k!}{(k+n-m)!}=1, \\ k=3, \quad L_3(x)=\frac{1}{6}(6-18x+9x^2-x^3), & \frac{k!}{(k+n-m)!}=1, \end{cases}$$

hence the ergodic capacity for the $4 \times 4$ CSCG MIMO channel is

$$\begin{aligned} C_{4\times 4} &= \int_0^\infty \log_2\left(1+\frac{\rho}{4}x\right)\left[1+(x-1)^2+\frac{\left(2-4x+x^2\right)^2}{4}\right. \\ &\left.+\frac{\left(6-18x+9x^2-x^3\right)^2}{36}\right]\exp(-x)dx. \end{aligned} \tag{2.38}$$

## 2A.4 $2 \times 4$

For a $2 \times 4$ i.i.d. CSCG MIMO channel, we have

$$n_R = 4, \quad n_T = 2, \quad m = 2, \quad n = 4.$$

We then have

$$\begin{cases} k=0, \quad L_0^2(x) = 1, & \frac{k!}{(k+n-m)!} = \frac{1}{2}, \\ k=1, \quad L_1^2(x) = -x+3, & \frac{k!}{(k+n-m)!} = \frac{1}{6}, \end{cases}$$

hence the ergodic capacity for the $2 \times 4$ CSCG MIMO channel is

$$C_{2\times 4} = \int_0^\infty \log_2\left(1+\frac{\rho}{2}x\right)\left[\frac{1}{2}+\frac{1}{6}(x-3)^2\right]x^2\exp(-x)dx. \tag{2.39}$$

## 2A.5 $4 \times 8$

For a $4 \times 8$ i.i.d. CSCG MIMO channel, we have

$$n_R = 8, \quad n_T = 4, \quad m = 4, \quad n = 8.$$

We then have

$$\begin{cases} k=0, \quad L_0^4(x) = 1, & \frac{k!}{(k+n-m)!} = \frac{1}{24}, \\ k=1, \quad L_1^4(x) = -x+5, & \frac{k!}{(k+n-m)!} = \frac{1}{120}, \\ k=2, \quad L_2^4(x) = \frac{1}{2}(30-12x+x^2), & \frac{k!}{(k+n-m)!} = \frac{1}{360}, \\ k=3, \quad L_3^4(x) = \frac{1}{6}(210-126x+21x^2-x^3), & \frac{k!}{(k+n-m)!} = 1, \end{cases}$$

and the ergodic capacity for the $8 \times 4$ CSCG MIMO channel is

$$\begin{aligned} C_{4\times 8} &= \int_0^\infty \log_2\left(1+\frac{\rho}{8}x\right)\left[\frac{1}{24}+\frac{(x-5)^2}{120}+\frac{\left(30-12x+x^2\right)^2}{1440}\right. \\ &\left.+\frac{\left(210-126x+21x^2-x^3\right)^2}{840\times 36}\right]x^4\exp(-x)dx. \end{aligned} \tag{2.40}$$

# Appendix 2B - Performance Bound of BICM-STBC System

In this section, we derive the union bound of BICM-STBC system with ML decision and ML decoding. We first consider BPSK modulation and i.i.d. CSCG MIMO channels which remain constant within each STBC code word but change from one codeword to another, i.e., the MIMO channels are quasi-static with respect to the STBC codeword interval, but fast fading with respect to the FEC codeword interval. We then analyze the system performance in i.i.d. CSCG channels which remain constant throughput the entire FEC code word. These two channels are also termed as *fast* and *slow quasi-static* fading channels.

Making use of the fact that the OSTBC designed for $n_T$ transmit antennas provides exactly the same performance as the $n_T n_R$ order receive MRC if the SNR normalization is taken care of, as proven in [33] and [34], we start the union bound derivation from the PEP of a coded system with $L$-branch MRC.

## 2B.1 Fast Fading Channel

Expressing the received signal at symbol interval $i$ and MRC branch $l$ as

$$r_{i,l} = h_{i,l} c_i \sqrt{E_c} + n_{i,l},$$

where $c_i$ is the BPSK modulated coded bits, $\{h_{i,l}\}$ are i.i.d. complex Gaussian, and $n_{i,l}$ is the AWGN noise with variance $\sigma^2$, we have the MRC decision statistic as

$$y_i = \sum_{l=1}^{L} h_{i,l} r_{i,l} = \left( \sum_{l=1}^{L} h_{i,l}^2 \right) c_i \sqrt{E_c} + \sum_{l=1}^{L} h_{i,l} n_{i,l}.$$

Defining

$$\alpha_i = \sqrt{\sum_{l=1}^{L} h_{i,l}^2},$$

we have that $\alpha_i^2$ is chi-square distributed with $2L$ degrees of freedom, with the pdf of

$$p(h) = \frac{1}{(2\gamma)^L (L-1)!} h^{L-1} \exp(\frac{-h}{2\gamma^2}),$$

where $2\gamma^2$ is the variance of $h_{i,l}$, $l = 1,\ 2, \cdots, L$.

Scaling the decision statistic $y_i$ with $\alpha_i$, we have

$$z_i = \frac{y_i}{\alpha_i} = \alpha_i c_i \sqrt{E_c} + v_i,$$

where

$$\begin{aligned} v_i &= \frac{\sum_{l=1}^{L} h_{i,l} n_{i,l}}{\alpha_i}, \\ \mathcal{E}\{v_i\} &= 0, \\ \mathcal{E}\{|v_i|^2\} &= \mathcal{E}\left\{ \frac{\sum_{l=1}^{L} h_{i,l} n_{i,l} \sum_{m=1}^{L} h_{i,m} n_{i,m}}{\alpha_i^2} \right\} = \sigma^2. \end{aligned}$$

We therefore have

$$p\left\{z_i | c_i, \alpha_i\right\} = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left( -\frac{|z_i - \alpha_i c_i \sqrt{E_c}|^2}{2\sigma^2} \right).$$

If we denote the transmitted code sequence of length $n$ (for block code, $n$ is the code word length) as $\mathbf{C}$, and the corresponding scaled MRC output sequence as $\mathbf{Z}$, we have

$$\begin{aligned} p\{\mathbf{Z}|\mathbf{C}\} &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left( -\frac{\left|z_i - \alpha_i c_i \sqrt{E_c}\right|^2}{2\sigma^2} \right) \\ &= \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \exp\left[ -\frac{\sum_{i=1}^{n} \left( |z_i|^2 - 2 z_i \alpha_i c_i \sqrt{E_c} + \alpha_i^2 E_c \right)}{2\sigma^2} \right] \\ &= \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \exp\left[ -\frac{\sum_{i=1}^{n} \left( |z_i|^2 + \alpha_i^2 E_c \right)}{2\sigma^2} \right] \exp\left( \frac{\sum_{i=1}^{n} z_i \alpha_i c_i \sqrt{E_c}}{\sigma^2} \right). \end{aligned}$$

We then have the PEP as

$$\begin{aligned} P_{2,e}\left(\tilde{\mathbf{C}}|\mathbf{C}\right) &= Prob\left\{ p\{\mathbf{Z}|\tilde{\mathbf{C}}\} > p\{\mathbf{Z}|\mathbf{C}\} \right\} \\ &= Prob\left\{ \exp\left( \frac{\sum_{i=1}^{n} z_i \alpha_i \tilde{c}_i \sqrt{E_c}}{\sigma^2} \right) > \exp\left( \frac{\sum_{i=1}^{n} z_i \alpha_i c_i \sqrt{E_c}}{\sigma^2} \right) \right\} \\ &= Prob\left\{ \sum_{i=1}^{n} z_i \alpha_i \sqrt{E_c} \tilde{c}_i > \sum_{i=1}^{n} z_i \alpha_i \sqrt{E_c} c_i \right\} \\ &= Prob\left\{ \sum_{i=1}^{n} \alpha_i z_i \left( \tilde{c}_i - c_i \right) > 0 \right\}. \end{aligned}$$

If the Hamming distance between $\tilde{\mathbf{C}}$ and $\mathbf{C}$ is $d_m$, i.e., $(\tilde{c}_i - c_i) = -2c_i$ for $d_m$ symbols, and

$(\tilde{c}_i - c_i) = 0$ for $n - d_m$ symbols, we have the PEP as

$$\begin{aligned} P_{2,e}\left(\tilde{\mathbf{C}}|\mathbf{C}\right) = Prob\left\{\sum_{j=1}^{d_m} \alpha_j z_j(-2c_j) > 0\right\} \\ &= Prob\left\{\sum_{j=1}^{d_m} \alpha_j c_j \left(\alpha_j c_j \sqrt{E_c} + v_j\right) < 0\right\} \\ &= Prob\left\{\sum_{j=1}^{d_m} \alpha_j^2 \sqrt{E_c} < -\sum_{j=1}^{d_m} \alpha_j c_j v_j\right\} \\ &= Prob\left\{\sqrt{E_c}\sum_{j=1}^{d_m}\sum_{l=1}^{L}|h_{j,l}|^2 < \underbrace{-\sum_{j=1}^{d_m} c_j v_j \sqrt{\sum_{l=1}^{L}|h_{j,l}|^2}}_{\eta}\right\}, \end{aligned}$$

where

$$\begin{aligned} \mathcal{E}\{\eta\} &= 0, \\ \mathcal{E}\{|\eta|^2\} &= \sigma^2 \sum_{j=1}^{d_m}\sum_{l=1}^{L}|h_{i,l}|^2, \end{aligned}$$

i.e., $\eta$ is zero mean Gaussian.

We therefore have the PEP as

$$\begin{aligned} P_{2,e}(d_m|\mathbf{h}) &= \mathrm{Q}\left(\frac{\sqrt{E_c \sum_{j=1}^{d_m}\sum_{l=1}^{L}|h_{j,l}|^2}}{\sigma}\right) \\ &= \mathrm{Q}\left(\sqrt{\frac{2E_c}{N_o}\sum_{j=1}^{d_m}\sum_{l=1}^{L}|h_{j,l}|^2}\right), \end{aligned} \tag{2.41}$$

which has a diversity order of $d_m L$.

The average PEP is therefore

$$\overline{P}_{2,e}(d_m) = \left[\frac{1}{2}(1-\mu)\right]^{d_m L} \sum_{k=0}^{d_m L - 1} \left(\begin{array}{c} d_m L - 1 + k \\ k \end{array}\right) \left[\frac{1}{2}(1+\mu)\right]^k,$$

where $\mu = \sqrt{\frac{\frac{E_c}{N_o}}{1+\frac{E_c}{N_o}}} = \sqrt{\frac{2\gamma^2 \frac{E_c}{N_o}}{1+2\gamma^2\frac{E_c}{N_o}}}$.

For block code, we can therefore obtain the union bound of code word error rate as

$$P(\text{code word error}) \le \sum_{d=d_{min}}^{n} N_d P_{2,e}(d),$$

where $N_d$ is the number of code word with Hamming weight $d$.

For convolutional code (CC) $(n, k)$ where $k$ is the number of input bits and $n$ is the number of output bits per time interval, the union bound for soft decision decoding is

$$P(\text{code word error}) < \sum_{d=d_{free}}^{\infty} \beta_d P_{2,e}(d),$$

where $d_{free}$ is the minimum free distance of the code, $\{\beta_d\}$ are obtained from expanding the first derivative of the transfer function $T(D, N)$ [17] as follows

$$\left.\frac{dT(D,N)}{dN}\right|_{N=1} = \sum_{d=d_{free}}^{\infty} \beta_d D^d.$$

The bit error probability is obtained as

$$P_b < \frac{1}{k} \sum_{d=d_{free}}^{\infty} \beta_d P_{2,e}(d). \tag{2.42}$$

Figure 2.4 depicts the simulated uncoded and coded performance as well as the derived bound for fast-fading $2 \times 1$ and $2 \times 2$ Alamouti STBC system. The constraint length $K = 3$ $R_c = \frac{1}{2}$ convolution code is used. From the figure we can see clearly that FEC introduces diversity gain in fast fading channels, exhibited by the slope change of the BER versus SNR curves. We also see a very good match between the simulation and the bound. We have therefore confirmed that concatenating FEC with STBC can exploit the diversity gain in a fast fading channel.

## 2B.2 Slow Fading Channel

For slow fading channel whose coefficients remain unchanged during the code word interval, the MRC gain coefficient $\alpha_i$ remains constant for all $i = 1, \cdots, n$, i.e.

$$\alpha = \sqrt{\sum_{l=1}^{L} h_l^2}.$$

The scaled MRC output $z_i$ is thus

$$z_i = \sqrt{E_c}\alpha c_i + v_i$$

with

$$\begin{aligned}
v_i &= \frac{\sum_{l=1}^{L} h_l n_{i,l}}{\alpha}, \\
\mathcal{E}\{v_i\} &= 0, \\
\mathcal{E}\{|v_i|^2\} &= \mathcal{E}\left\{\frac{\sum_{l=1}^{L} h_l n_{i,l} \sum_{m=1}^{L} h_m n_{i,m}}{\alpha^2}\right\} = \sigma^2.
\end{aligned}$$

Making use of the results from the previous section, we can quickly write the PEP for slow fading channel as

$$\begin{aligned}
P_{2,e}\left(\tilde{\mathbf{C}}|\mathbf{C}\right) &= Prob\left\{p\{\mathbf{Z}|\tilde{\mathbf{C}}\} > p\{\mathbf{Z}|\mathbf{C}\}\right\} \\
&= Prob\left\{\exp\left(\frac{\sum_{i=1}^{n} z_i \alpha \tilde{c}_i \sqrt{E_c}}{\sigma^2}\right) > \exp\left(\frac{\sum_{i=1}^{n} z_i \alpha c_i \sqrt{E_c}}{\sigma^2}\right)\right\} \\
&= Prob\left\{\sum_{i=1}^{n} z_i \alpha \sqrt{E_c}\tilde{c}_i > \sum_{i=1}^{n} z_i \alpha \sqrt{E_c} c_i\right\} \\
&= Prob\left\{\sum_{i=1}^{n} z_i \left(\tilde{c}_i - c_i\right) > 0\right\} \\
&= Prob\left\{\sum_{j=1}^{d_m} z_j(-2c_j) > 0\right\} \\
&= Prob\left\{\sum_{j=1}^{d_m} c_j \left(\alpha c_j \sqrt{E_c} + v_j\right) < 0\right\} \\
&= Prob\left\{\sum_{j=1}^{d_m} \alpha\sqrt{E_c} < \underbrace{-\sum_{j=1}^{d_m} c_j v_j}_{\eta}\right\} \\
&= Prob\left\{\eta > d_m \alpha \sqrt{E_c}\right\}.
\end{aligned}$$

As $\eta$ is Gaussian, and

$$\begin{aligned}
\mathcal{E}\{\eta\} &= 0, \\
\mathcal{E}\{|\eta|^2\} &= d_m \sigma^2,
\end{aligned}$$

we have the PEP as

$$P_{2,e}(d_m|\mathbf{h}) = \mathrm{Q}\left(\sqrt{\frac{d_m^2 \alpha^2 E_c}{d_m \sigma^2}}\right) = \mathrm{Q}\left(\sqrt{\frac{2E_c d_m}{N_o}\sum_{l=1}^{L}|h_l|^2}\right), \tag{2.43}$$

which shows that the diversity gain is of order $L$ from the orthogonal STBC, and coding gain of $d_m$ from the FEC, in contrast to the full diversity gain of order $Ld_m$ in fast fading channel case.

The average PEP is therefore

$$\overline{P}_{2,e}(d_m) = \left[\frac{1}{2}(1-\mu)\right]^L \sum_{k=0}^{L-1} \begin{pmatrix} L-1+k \\ k \end{pmatrix} \left[\frac{1}{2}(1+\mu)\right]^k,$$

where $\mu = \sqrt{\frac{\frac{\overline{E_c}}{N_o}}{1+\frac{\overline{E_c}}{N_o}}} = \sqrt{\frac{2\gamma^2 d_m \frac{E_c}{N_o}}{1+2d_m\gamma^2\frac{E_c}{N_o}}}$.

For block code, we can therefore obtain the union bound of code word error rate as

$$P(\text{code word error}) \leq \sum_{d=d_{min}}^{n} N_d P_{2,e}(d),$$

where $N_d$ is the number of code word with Hamming weight $d$.

For CC $(n,k)$ where $k$ is number of input bits and $n$ is the number of output bits per time interval, the union bound for soft decision decoding is

$$P(\text{code word error}) < \sum_{d=d_{free}}^{\infty} \beta_d P_{2,e}(d),$$

where $d_{free}$ is the minimum free distance of the code, $\{\beta_d\}$ are obtained from expanding the first derivative of the transfer function $T(D,N)$ [17] and the bit error probability is obtained as

$$P_b < \frac{1}{k} \sum_{d=d_{free}}^{\infty} \beta_d P_{2,e}(d). \tag{2.44}$$

Figure 2.5 depicts the uncoded and coded performance for slow-fading $2\times1$ and $2\times2$ Alamouti STBC system. Same as in the study on fast fading channels, we use the constraint length $K=3$ $R_c=\frac{1}{2}$ convolution code and QPSK modulation. From the figure we can see clearly that FEC introduces only coding gain in slow fading channels, exhibited by the parallel shift of the curves.

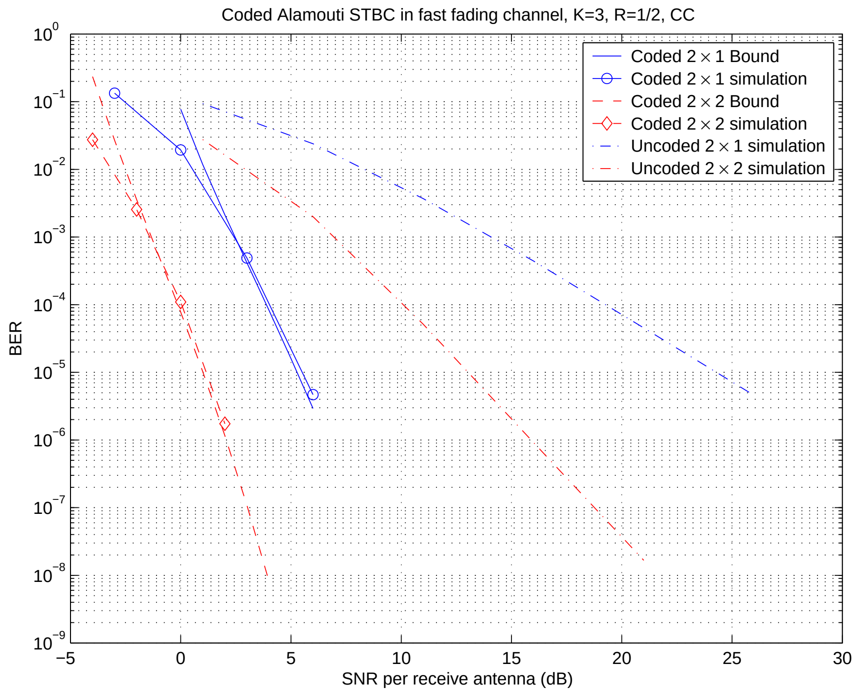


**Figure 2.4: Convolutional coded STBC system performance. Bound analysis and simulation result. K=3, $R_c = \frac{1}{2}$, BPSK.**

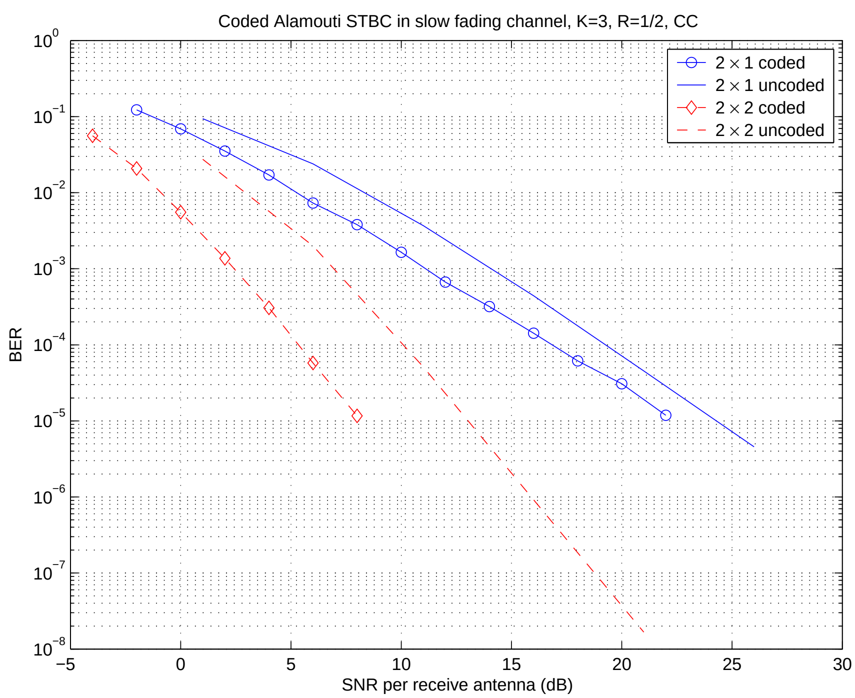


**Figure 2.5: Convolutional coded STBC system performance. Bound analysis and simulation result. K=3, $R_c = \frac{1}{2}$, BPSK.**

# Chapter 3

# An Overview of MIMO-OFDM

In this chapter, we will develop the MIMO-OFDM system model and give an overview of MIMO-OFDM. We start with the signal model formulation for conventional single-antenna cyclic prefix (CP)-based OFDM systems, and extend it to the MIMO systems. We will then generalize the linear signal model and incorporate the various STFP schemes. The generalized linear signal model will facilitate the capacity and diversity analysis of the various STFP schemes. The receiver algorithms derived based on one particular STFP scheme can be also easily extended to other systems.

## 3.1 A General MIMO-OFDM System Model

In this section, we develop a mathematical model of MIMO-OFDM systems with $n_T$ transmit and $n_R$ receive antennas. The OFDM modulation is composed of two steps of operation - inverse fast Fourier transform (IFFT) and CP insertion. We denote $N$ as the total number of subcarriers, or the FFT size, $P$ as the number of subcarriers used to transmit data (and pilot signals) where $P \leq N$, $L$ the number of sample-spaced multipaths in each of the MIMO channels defined by the transmit-receive antenna pairs, and $L_{\mathrm{CP}}$ the CP length in samples. Without loss of generality, we assume that $L \leq L_{\mathrm{CP}}$, the inter-symbol interference (ISI) from the multipath channel can therefore be completely mitigated, as shown in the following derivation.

### 3.1.1 Signal Model for Single-Input Single-Output OFDM

The OFDM modulated signal $\mathbf{x}$ for a single-antenna system can be written as

$$\mathbf{x} = \mathbf{T}_{\mathrm{CP}}\mathbf{F}^H\mathbf{X}, \tag{3.1}$$

where $\mathbf{x}$ is a column vector of dimension $(N + L_{\mathrm{CP}}) \times 1$, $\mathbf{F}$ is the $N \times N$ Fourier transform matrix with its elements defined as $f_{mn} = \frac{1}{\sqrt{N}} \exp\left[-j\frac{2\pi}{N}(m-1)(n-1)\right]$, $m,\ n = 1, 2, \cdots, N$, and $\mathbf{F}^H$ represents the IFFT operation on the frequency domain signal vector $\mathbf{X}$. $\mathbf{T}_{\mathrm{CP}}$ is a circulant matrix [49] of size $(N + L_{\mathrm{CP}}) \times N$ with its first row written as:

$$\left( \underbrace{0 \quad \cdots \quad 0}_{N-L_{\mathrm{CP}}} \; 1 \; \underbrace{0 \quad \cdots \quad 0}_{L_{\mathrm{CP}}-1} \right).$$

$\mathbf{T}_{\mathrm{CP}}$ adds CP to the IFFT output. $\mathbf{X}$ is of size $N$ When $P = N$, $\mathbf{X}$ is the signal to be transmitted; when $P < N$, some subcarriers at the edges of the allocated bandwidth are used as the guard band. The subcarrier allocation scheme defined in IEEE 802.11a WLAN [2] is illustrated in Fig. 3.1. Out of the sixty-four subcarriers, i.e., N=64, eleven subcarriers are used as guard subcarriers, five of which at the higher frequency band (compared to direct current, or $0$th subcarrier), i.e., subcarriers $27 \sim 31$, and six at the lower frequency band, i.e., subcarriers $-32 \sim -27$. No data or pilot is transmitted at the direct current (dc) subcarrier, either. In this case $\mathbf{X}$ is formed as follows:

$$\mathbf{X}^T = \left( 0 \quad \mathbf{X}_h^T \quad \mathbf{0}_{N-P-1}^T \quad \mathbf{X}_l^T \right),$$

where $\mathbf{X}_h$ denotes the frequency domain signal at the higher frequency subcarriers (in the IEEE 802.11a case, subcarriers $1 \sim 26$), $\mathbf{X}_l$ denotes the signal at the lower frequency subcarriers (subcarriers $-26 \sim -1$ in the IEEE 802.11a case), and $\mathbf{0}_{N-P-1}$ denotes the all zero vector with length (N-P-1).

Assuming a sample-spaced multipath channel with $L$ equally-spaced multipaths and the $l$th element having complex gain of $h_l$, we can write the received signal as

$$\mathbf{r}_i = \mathbf{h}_0\mathbf{x}_i + \mathbf{h}_1\mathbf{x}_{i-1} + \mathbf{v}_i, \tag{3.2}$$

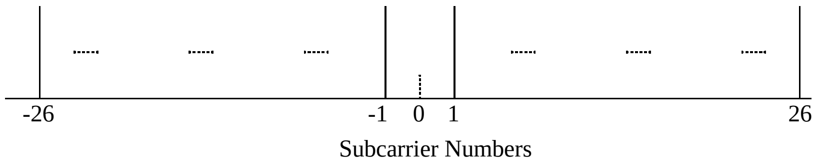


**Figure 3.1: Illustration of Subcarrier Allocation with Guard Bands**

where $i$ represents the $i$th received block of data, $\mathbf{h}_0$ and $\mathbf{h}_1$ are both size $(N+L_{\mathrm{CP}})\times(N+L_{\mathrm{CP}})$ Toeplitz matrices [49] with $(h_0, h_1, \cdots, h_{L-1}, 0, \cdots, 0)^T$ as the first column and $(h_0, 0, \cdots, 0)$ as the first row of $\mathbf{h}_0$, and $(0, \cdots, 0)^T$ as the first column and $(0, \cdots, 0, h_{L-1}, h_{L-2}, \cdots, h_1)$ as the first row of $\mathbf{h}_1$, i.e.,

$$\mathbf{h}_0 = \begin{bmatrix} h_0 & 0 & 0 & \cdots & 0 & 0 \\ h_1 & h_0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ h_{L-1} & h_{L-2} & h_{L-3} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & h_1 & h_0 \end{bmatrix}_{(N+L_{\mathrm{CP}})\times(N+L_{\mathrm{CP}})},$$

and

$$\mathbf{h}_1 = \begin{bmatrix} 0 & \cdots & h_{L-1} & h_{L-2} & \cdots & h_1 \\ 0 & \cdots & 0 & h_{L-1} & \cdots & h_2 \\ \vdots & \ddots & 0 & 0 & \ddots & 0 \end{bmatrix}_{(N+L_{\mathrm{CP}})\times(N+L_{\mathrm{CP}})}.$$

$\mathbf{v}_i$ denotes the complex AWGN with zero mean and variance $\sigma^2$. Therefore $\mathbf{r}_i$ consists of three parts: the desired signal $\mathbf{x}_i$, the inter-OFDM symbol interference from the previous OFDM symbol $\mathbf{x}_{i-1}$, and the AWGN.

Frame synchronization process identifies the starting point of the OFDM block, following which the CP portion in the received signal is removed. This is written as:

$$\mathbf{y}_i = \mathbf{R}_{\mathrm{CP}}\mathbf{r}_i, \tag{3.3}$$

where $\mathbf{R}_{\mathrm{CP}}$ is a circulant matrix of size $N \times (N + L_{\mathrm{CP}})$ whose first row is written as

$$(\underbrace{0 \;\cdots\; 0}_{L_{\mathrm{CP}}}\; 1\; \underbrace{0 \;\cdots\; 0}_{N-1}).$$

Therefore, $\mathbf{R}_{\text{CP}}\mathbf{h}_1 = \mathbf{0}$, the inter-OFDM symbol interference is removed from the received signal. Hence we drop the index $i$ and rewrite (3.3) as:

$$\mathbf{y} = \mathbf{R}_{\text{CP}}\mathbf{h}_0\mathbf{x} + \mathbf{v}, \tag{3.4}$$

where $\mathbf{v}$ is the AWGN vector with length $N$. Performing FFT on $\mathbf{y}$, we have

$$\begin{aligned}\mathbf{Y} &= \mathbf{F}\mathbf{R}_{\text{CP}}\mathbf{h}_0\mathbf{x} + \mathbf{F}\mathbf{v} \\ &= \mathbf{F}\mathbf{R}_{\text{CP}}\mathbf{h}_0\mathbf{T}_{\text{CP}}\mathbf{F}^H\mathbf{X} + \mathbf{V},\end{aligned} \tag{3.5}$$

where $\mathbf{V}$ represents the frequency domain noise which is still white Gaussian, and $\mathbf{R}_{\text{cp}}\mathbf{h}_0\mathbf{T}_{\text{cp}}$ is a $N \times N$ circulant matrix with the first column written as

$$\mathbf{h}_t = [h_0, h_1, \cdots, h_{L-1}, 0, \cdots, 0]^T. \tag{3.6}$$

Therefore, $\mathbf{R}_{\text{cp}}\mathbf{h}_0\mathbf{T}_{\text{cp}}$ can be diagonalized and we have

$$\mathbf{F}\mathbf{R}_{\text{cp}}\mathbf{h}_0\mathbf{T}_{\text{CP}}\mathbf{F}^H = \mathbf{H} = \text{diag}\left(H_0, H_1, \cdots, H_{N-1}\right),$$

where $\mathbf{H}$ is the frequency response of the channel with the $n$th diagonal element expressed as [50]

$$H_n = \sum_{l=0}^{L-1} h_l \exp(-j\frac{2\pi}{N}nl) = \sqrt{N}\mathbf{F}^n\mathbf{h}_t, \tag{3.7}$$

where $\mathbf{F}^n$ is the $n$th row of the FFT matrix $\mathbf{F}$.

Hence, we can rewrite (3.5) as

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{V}. \tag{3.8}$$

**Theorem 1.** *The frequency domain channel response $H_n$'s, $n = 0,\ 1,\ \cdots,\ N-1$ have the same statistical property for wide-sense stationary uncorrelated scattering (WSSUS) complex Gaussian multipath channels in which only the first component may have non-zero mean.*

*Proof:* For WSSUS complex Gaussian multipath channels, we have for $h_l = h_{l,Re} + jh_{l,Im}$,

$$\begin{aligned}\mathcal{E}\left\{h_{l,Re}^2\right\} &= \mathcal{E}\left\{h_{l,Im}^2\right\}, && l \neq 0, \\ \mathcal{E}\left\{h_{l,Re}^2\right\} &= a_0^2 + \mathcal{E}\left\{h_{l,Im}^2\right\}, && l = 0, \\ \mathcal{E}\left\{h_{l,Im}\right\} &= 0, && l = 0,\ 1, \cdots, L-1, \\ \mathcal{E}\left\{h_{l,Re}\right\} &= a_0\delta(l), && l = 0,\ 1, \cdots, L-1\end{aligned}$$

where $a_0$ is the Rician component which is assumed to appear in the real part of $h_0$.

We also have

$$
\begin{aligned}
\mathcal{E}\left\{h_{l,Re}h_{l,Im}\right\} &= 0, \\
\mathcal{E}\left\{h_l^2\right\} &= \mathcal{E}\left\{h_{l,Re}^2 - h_{l,Im}^2 + 2jh_{l,Re}h_{l,Im}\right\} = a_0^2\delta(l), \\
\mathcal{E}\left\{h_l h_{m\neq l}\right\} &= 0 \\
\mathcal{E}\left\{h_l h_{m\neq l}^*\right\} &= 0.
\end{aligned}
$$

From (3.7), $H_n$ is complex Gaussian random variable (r.v.) when the time domain coefficients $h_l$, $l = 0,\ 1,\ \cdots,\ L-1$, are complex Gaussian. Therefore, in order to prove Theorem 1, we only need to show that all the $H_n$'s have the same first order and second order statistics.

The first-order mean value of $H_n$ is

$$
\mathcal{E}\left\{H_n\right\} = \mathcal{E}\left\{\sqrt{N}\mathbf{F}^n\mathbf{h}_t\right\} = \sqrt{N}\mathbf{F}^n\mathcal{E}\left\{\mathbf{h}_t\right\} = \sqrt{N}F(n,0)\mathcal{E}\left\{h_0\right\} = \mathcal{E}\left\{h_0\right\},
$$

i.e., all $H_n$'s have the same mean.

Now let's look at the second order statistics.

$$
\begin{aligned}
\mathcal{E}\left\{|H_n|^2\right\} &= \mathcal{E}\left\{H_n H_n^*\right\} = \mathcal{E}\left\{N\mathbf{F}^n\mathbf{h}_t\mathbf{h}_t^H\left(\mathbf{F}^n\right)^H\right\} = N\mathbf{F}^n\mathcal{E}\left\{\mathbf{h}_t\mathbf{h}_t^H\right\}\left(\mathbf{F}^n\right)^H \\
&= N\sum_{l=0}^{L-1}F(n,l)\mathcal{E}\left\{|h_l|^2\right\}F(n,l)^* = L\sum_{l=0}^{L-1}\mathcal{E}\left\{|h_l|^2\right\}, \\
\mathcal{E}\left\{H_n^2\right\} &= N\mathcal{E}\left\{\left(\sum_{l=0}^{L-1}F(n,l)h_l\right)^2\right\} \\
&= N\sum_{l=0}^{L-1}F(n,l)^2\mathcal{E}\left\{h_l^2\right\} \\
&= \mathcal{E}\left\{h_0^2\right\} \\
Var(H_n) &= \mathcal{E}\left\{\left(H_n - \mathcal{E}(H_n)\right)\left(H_n - \mathcal{E}(H_n)\right)^*\right\} \\
&= \mathcal{E}\left\{|H_n|^2 - \mathcal{E}(H_n)H_n^* - H_n\mathcal{E}(H_n^*) + |\mathcal{E}(H_n)|^2\right\} \\
&= \mathcal{E}\left\{|H_n|^2\right\} - |\mathcal{E}(H_n)|^2 \\
&= L\sum_{l=0}^{L-1}\mathcal{E}\left\{|h_l|^2\right\} - |\mathcal{E}\left\{h_0\right\}|^2,
\end{aligned}
$$

i.e., all $H_n$'s have the same second-order statistics.

We hence prove that $H_n$, $n = 0,\ 1,\ \cdots,\ N-1$ are statistically the same for WSSUS complex Gaussian multipath channels. ■

**Theorem 2.** *The maximum frequency diversity order of an OFDM system is $L$, where $L$ is the number of multipaths in the frequency selective channel.*

*Proof:* From (3.8), assuming perfect CSI at the receiver, we have the PEP of maximum likelihood detection (MLD) as

$$\begin{aligned} P(\mathbf{X} \to \mathbf{X}_e) &= P\left(p(\mathbf{Y}|\mathbf{X},\mathbf{H}) < p(\mathbf{Y}|\mathbf{X}_e,\mathbf{H})\right) \\ &= P\left(|\mathbf{Y} - \mathbf{H}\mathbf{X}|^2 > |\mathbf{Y} - \mathbf{H}\mathbf{X}_e|^2\right) \\ &= Q\left(\sqrt{\frac{d^2(\mathbf{H}\mathbf{X}, \mathbf{H}\mathbf{X}_e)}{2N_0}}\right) \\ &\leq \exp\left(-\frac{|\mathbf{H}\mathbf{X} - \mathbf{H}\mathbf{X}_e|^2}{4N_0}\right), \end{aligned}$$

where $\mathbf{X}$ is the transmitted sequence, and $\mathbf{X}_e$ is the erroneously detected sequence. $N_0 = 2\sigma^2$.

Defining $\mathbf{e} = \mathrm{diag}\,(\mathbf{X} - \mathbf{X}_e)$, and $\mathbf{H} = \mathrm{diag}\,(\mathbf{F}\mathbf{h}_t)$, we have

$$\begin{aligned} \|\mathbf{H}\mathbf{X} - \mathbf{H}\mathbf{X}_e\|^2 &= \|\mathbf{e}\mathbf{F}\mathbf{h}_t\|^2 \\ &= \mathbf{h}_t^H \mathbf{F}^H \mathbf{e}^H \mathbf{e}\mathbf{F}\mathbf{h}_t \\ &= \mathbf{h}_t^H \mathbf{F}^H \mathbf{Q}\mathbf{\Lambda}_e \mathbf{Q}^H \mathbf{F}\mathbf{h}_t, \end{aligned}$$

where $\mathbf{e}\mathbf{e}^H = \mathbf{Q}\mathbf{\Lambda}_e\mathbf{Q}^H$ is the eigen-decomposition of the error vector covariance matrix $\mathbf{e}\mathbf{e}^H$. The number of non-zero eigenvalues $d_c$ is determined by the free distance of the FEC.

Defining $\bar{\mathbf{h}}_t = \mathbf{Q}^H\mathbf{F}\mathbf{h}_t$, there are $L$ non-zero independent complex Gaussian elements in $\bar{\mathbf{h}}_t$, and we have

$$\|\mathbf{H}\mathbf{X} - \mathbf{H}\mathbf{X}_e\|^2 = \bar{\mathbf{h}}_t^H \mathbf{\Lambda}_e \bar{\mathbf{h}}_t = \sum_{l=0}^{L-1} |\bar{h}_l|^2 \lambda_{e,l}$$

where we assume that $d_c \geq L$.

We therefore can write the PEP as

$$P(\mathbf{X} \to \mathbf{X}_e) = \prod_{l=0}^{L-1} \exp\left(-\frac{|\bar{h}_l|^2 \lambda_{e,l}}{4N_0}\right),$$

and when the multipath components $h_l$'s are i.i.d. complex Gaussian with zero mean and variance $\frac{1}{L}$, we have the average PEP written as

$$P_e(\mathbf{X} \to \mathbf{X}_e) = \left(\frac{1}{N_0}\right)^{-d_l} \left(\prod_{l=0}^{d_l-1} \lambda_{e,l}\right)^{-1}, \tag{3.9}$$

where $d_l = \min(L, d_c)$ is the *diversity order of the OFDM system*, and we have $\max(d_l) = L$, achieved when the free distance of the FEC $d_c$ is larger than the multipath order.

■

### 3.1.2 Signal Model for MIMO-OFDM

With the SISO-OFDM model defined in (3.8), we can now work out the MIMO OFDM system model in a very straight-forward manner, as:

$$\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{X}} + \boldsymbol{\mathcal{V}}, \tag{3.10}$$

where $\boldsymbol{\mathcal{X}}$ is a dimension-$n_T N$ column vector obtained by stacking the transmitted signal vectors $\mathbf{X}_m$, $m = 1,\ 2,\ \cdots,\ n_T$ from the $n_T$ transmit antennas, $\boldsymbol{\mathcal{Y}}$ is a dimension-$n_R N$ column vector obtained by stacking the received signal vector $\mathbf{Y}_n$, $n = 1,\ 2,\ \cdots,\ n_R$, from the $n_R$ receive antennas, and $\boldsymbol{\mathcal{H}}$ is the frequency domain MIMO-OFDM channel of size $(n_R N) \times (n_T N)$ which is written as

$$\boldsymbol{\mathcal{H}} = \begin{bmatrix} \mathbf{H}_{1,1} & \mathbf{H}_{1,2} & \cdots & \mathbf{H}_{1,n_T} \\ \mathbf{H}_{2,1} & \mathbf{H}_{2,2} & \cdots & \mathbf{H}_{2,n_T} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{H}_{n_R,1} & \mathbf{H}_{n_R,2} & \cdots & \mathbf{H}_{n_R,n_T} \end{bmatrix},$$

where $\mathbf{H}_{n,m}$ is a $N \times N$ diagonal matrix corresponding to the single-antenna frequency domain channel defined by the $m$th-transmit $n$th-receive antenna pair, $\boldsymbol{\mathcal{V}}$ is the AWGN noise vector of dimension $n_R N$ obtained by stacking the AWGN noise vector at the $n_R$ receive antennas.

Therefore, for each receive antenna $n$, the received signal at subcarrier $k$ can be expressed as:

$$\begin{aligned} \mathcal{Y}_{n,k} &= \sum_{m=1}^{n_T} \mathcal{H}_{(n-1)\times N+k,(m-1)\times N+k)} \mathcal{X}_{(m-1)\times N+k} \\ &= \sum_{m=1}^{n_T} H_{n,m,k} X_{m,k}, \end{aligned} \tag{3.11}$$

where $n = 1,\ 2,\ \cdots,\ n_R$, $m = 1,\ 2,\ \cdots,\ n_T$, and $k = 0,\ ,1,\ \cdots,\ N-1$.

Defining

$$
\begin{aligned}
\boldsymbol{\mathcal{R}}_k &= \begin{bmatrix} \mathcal{Y}_{1,k} & \mathcal{Y}_{2,k} & \cdots & \mathcal{Y}_{n_R,k} \end{bmatrix}^T, \\
\boldsymbol{\mathcal{H}}_k &= \begin{bmatrix} \mathcal{H}_{k,k} & \mathcal{H}_{k,2k} & \cdots & \mathcal{H}_{k,n_Tk} \\ \mathcal{H}_{2k,k} & \mathcal{H}_{2k,2k} & \cdots & \mathcal{H}_{2k,n_Tk} \\ & \cdots & & \\ \mathcal{H}_{n_Rk,k} & \mathcal{H}_{n_Rk,2k} & \cdots & \mathcal{H}_{n_Rk,n_Tk} \end{bmatrix} = \begin{bmatrix} H_{1,1,k} & H_{1,2,k} & \cdots & H_{1,n_T,k} \\ H_{2,1,k} & H_{2,2,k} & \cdots & H_{2,n_T,k} \\ \vdots & \vdots & \ddots & \vdots \\ H_{n_R,1,k} & H_{n_R,2,k} & \cdots & H_{n_R,n_T,k} \end{bmatrix}, \\
\boldsymbol{\mathcal{S}}_k &= \begin{bmatrix} \mathcal{X}_k & \mathcal{X}_{2k} & \cdots & \mathcal{X}n_Tk \end{bmatrix}^T, \\
\boldsymbol{\mathcal{N}}_k &= \begin{bmatrix} \mathcal{V}_k & \mathcal{V}_{2k} & \cdots & \mathcal{V}_{n_Tk} \end{bmatrix}^T,
\end{aligned}
$$

we can write (3.11) as

$$\boldsymbol{\mathcal{R}}_k = \boldsymbol{\mathcal{H}}_k \boldsymbol{\mathcal{S}}_k + \boldsymbol{\mathcal{N}}_k. \tag{3.12}$$

**Theorem 3.** *The frequency domain MIMO-OFDM channel response $\boldsymbol{\mathcal{H}}_k$'s, $k = 0,\ 1,\ \cdots,\ N-1$ have the same statistical property for spatially uncorrelated WSSUS complex Gaussian multipath channels in which only the first component may have non-zero mean.*

*Proof:* From Theorem 1, the elements $H_{n_r,n_t,k}$ are complex Gaussian and they are statistically the same for different $k$'s. Therefore, we only need to prove that the correlation coefficients between elements in $\boldsymbol{\mathcal{H}}_k$ is independent of $k$, as follows.

$$
\begin{aligned}
\mathcal{E}\left\{H_{m,n,k}H^*_{p,q,k}\right\} &= \mathcal{E}\left\{\left(\sum_{l=0}^{L-1} h_{m,n,l}e^{-j\frac{2\pi}{N}kl}\right)\left(\sum_{d=0}^{L-1} h_{p,q,d}e^{-j\frac{2\pi}{N}kd}\right)^*\right\} \\
&= \mathcal{E}\left\{\sum_{l=0}^{L-1}\sum_{d=0}^{L-1} h_{m,n,l}h^*_{p,q,d}e^{-j\frac{2\pi}{N}k(l-d)}\right\} \\
&= \sum_{l=0}^{L-1}\sum_{d=0}^{L-1} \mathcal{E}\left\{h_{m,n,l}h^*_{p,q,d}\right\} e^{-j\frac{2\pi}{N}k(l-d)} \\
&= \sum_{l=0}^{L-1}\sum_{d=0}^{L-1} \sigma_l^2 \rho_r^{\frac{1}{2}}(m,p,l)\rho_t^{\frac{1}{2}}(n,q,l)\delta(l-d)e^{-j\frac{2\pi}{N}k(l-d)} \\
&= \sum_{l=0}^{L-1} \sigma_l^2 \rho_r^{\frac{1}{2}}(m,p,l)\rho_t^{\frac{1}{2}}(n,q,l)
\end{aligned} \tag{3.13}
$$

with the assumption of WSSUS multipath models. Therefore, the MIMO channel $\boldsymbol{\mathcal{H}}_k$ have the same statistical properties for all the subcarriers. Here $\rho_t(n, q, l)$ and $\rho_r(m, p, l)$ denote the correlation coefficients of the $l$th path at the transmitter and the receiver, which are determined by the angle of departure and angle of arrival respectively [51][52].

■

**Corollary 1** The frequency domain channel is spatially uncorrelated if the time-domain multipath components are spatially uncorrelated.

*Proof:* This is straightforward from (3.13). ■

From now onward, unless otherwise stated, we will assume that there is no spatial correlation in the channel. For signal notation, we will no longer differentiate the MIMO signals from the single-antenna signals by using calligraphic letters unless otherwise stated.

## 3.2 STFP and FEC Encoding in MIMO-OFDM Systems

In this section, we will generalize the signal model developed in Section 3.1 by incorporating the various STFP schemes and discuss application of FEC coding in such systems.

A bit-interleaved coded and modulated STFP MIMO system is illustrated in Fig. 3.2. In this figure, the random information bits are first demultiplexed into parallel streams (layers) by the "Spatial DeMux" unit. In each parallel stream, information bits are encoded, bit interleaved, and mapped to constellation points of the adopted modulation scheme in the "BICM" unit, based on the bit-interleaved coded modulation (BICM) principle [35]. BICM provides both a large Hamming distance and a large Euclidean distance, hence is a robust coded modulation scheme for wireless channels. It also splits the coded modulation design into two parts - selection of the encoder, and design of the modulation scheme. In this thesis, we consider Gray mapping rules to map the coded and interleaved bits to symbols. In order to achieve "turbo" gain in iterative decoding of BICM (BICM-ID), other mapping rules have been proposed. The details can be referred to works by Li and Ritcey [53] [54] [55] [56], Schreckenbach *et. al.* [57] [58].

The BICM outputs from the various parallel streams are further processed by the "Space-Time-

Freq(uency)-Precoding", i.e., the STFP unit. Depending on the system requirement and the transmit and receive antennas available, different STFP schemes can be adopted. For example, if maximum capacity is targeted from a system with no less receive than transmit antennas, and if no CSI is available at the transmitter, simple spatial multiplexing, i.e., Vertical BLAST (VBLAST) structure can be used. The corresponding STFP processing is represented by a linear transform with identity matrix $\mathbf{I}_{n_T}$. If maximum transmit and receive diversity is desired from the MIMO channel, the full-diversity space-time coding schemes, e.g., STBC [31] [33], STTC [25] [59], etc., can be used. A linear frequency-domain transform can be applied to the BICM output before the STBC encoding to exploit the frequency domain diversity and improve the system performance, as suggested in [60]. The linear pre-transform can be applied to a VBLAST-OFDM system following a similar approach in [60]. To achieve compliance with legacy standard with multiple transmit antennas, cyclic delay diversity (CDD) can be applied [61, 62, 63]. CDD can also be combined with other multiple antenna processing schemes, e.g., transmit beamforming, as shown in [64], or spatial spreading (SS), as will be discussed in this thesis. If both transmit diversity and capacity gains are desired, groupwise STBC (GSTBC) [65], LDC [66], QSTBC [38, 39], etc., can be used. When the transmitter has no knowledge about the channel, equal rate is assigned to the different spatial streams, hence same BICM scheme should be used. Each spatial stream should have the same power allocation as well.

When perfect CSI is available at the transmitter, SVD-based beamforming in conjunction with water-filling performed at each subcarrier is optimum in achieving the channel capacity. In this case, the STFP is a linear transform represented by matrix $\mathbf{\Omega}$ as given in (2.2). To reduce the transmitter complexity, subchannel grouping (SCG) and statistical water-filling (SWF) was proposed in [67], and it was proved that SCG and SWF can achieve ergodic capacity of MIMO-OFDM channels. With SCG, the MIMO-OFDM channels are partitioned into several parallel Gaussian channels with different SNR and different diversity order. To realize the channel capacity, a multiple-codebook variable rate (MCVR) coded modulation was proposed in [68] in which different coded modulation scheme was used for different parallel channel, and the power ratio among the parallel channels was also adjusted according to the SWF principle in order to optimize the power utilization.

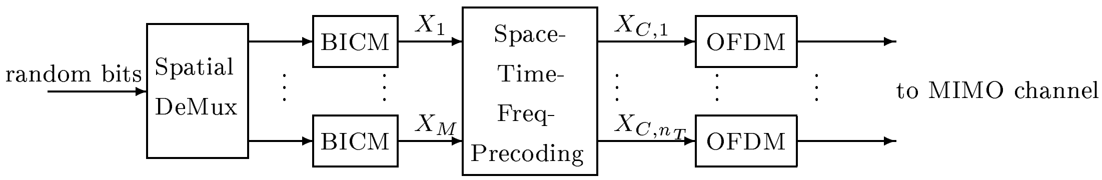


**Figure 3.2: A coded MIMO-OFDM transmitter.**

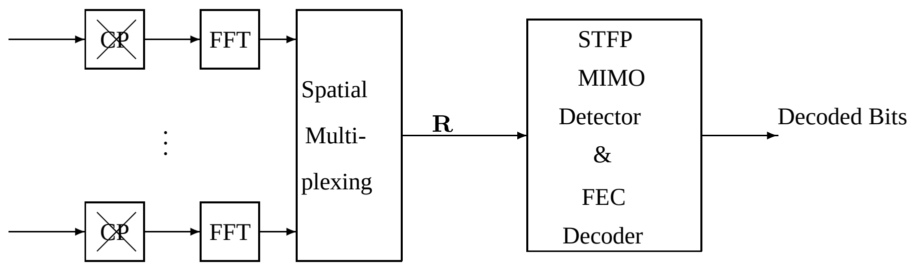


**Figure 3.3: Block Diagram of A Generalized MIMO OFDM Receiver.**

The output of STFP is then OFDM modulated and transmitted by different antennas.

Assuming perfect timing and frequency synchronization, we remove the CP part from the received data and then convert the signals to frequency domain by FFT, as shown in Fig. 3.3. Our aim here is the develop a unified linear signal model for each subcarrier through the "Spatial Multiplexing" unit, written as

$$\mathbf{R} = \mathbf{H}\mathbf{X} + \mathbf{V}, \tag{3.14}$$

where $\mathbf{R}$, $\mathbf{X}$, $\mathbf{H}$ and $\mathbf{V}$ denote the received and transmitted signals, the space-time-frequency precoded MIMO channel, and the complex AWGN noise, respectively. Depending on the STFP scheme adopted, the dimensions of $\mathbf{R}$, $\mathbf{X}$, $\mathbf{H}$ and $\mathbf{V}$ may vary for the same $n_T \times n_R$ MIMO channel. With such a general signal model, we can analyze the capacity and diversity performance within the same framework and identify the best possible precoding scheme. The receiver algorithms we develop in Chapter 5 can also be applied to the various STFP MIMO-OFDM systems in a straightforward manner.

### 3.2.1 VBLAST-OFDM

For a $n_T \times n_R$ VBLAST-OFDM system [65], the signals $\mathbf{R}$, $\mathbf{H}$, $\mathbf{X}$ and $\mathbf{V}$ are defined as

$$\begin{cases} \mathbf{R} & \stackrel{\text{def}}{=} & \begin{bmatrix} R_1 & R_2 & \cdots & R_{n_R} \end{bmatrix}^T, \\ \mathbf{H} & \stackrel{\text{def}}{=} & \begin{bmatrix} H_{1,1} & H_{1,2} & \cdots & H_{1,n_T} \\ \vdots & \vdots & \vdots & \vdots \\ H_{n_R,1} & H_{n_R,2} & \cdots & H_{n_R,n_T} \end{bmatrix}, \\ \mathbf{X} & \stackrel{\text{def}}{=} & \begin{bmatrix} X_1 & X_2 & \cdots & X_{n_T} \end{bmatrix}^T, \\ \mathbf{V} & \stackrel{\text{def}}{=} & \begin{bmatrix} V_1 & V_2 & \cdots & V_{n_R} \end{bmatrix}^T, \end{cases} \tag{3.15}$$

where $R_i$ denotes the received signal at antenna $i$, $H_{i,j}$ denotes the channel response between transmit antenna $j$ and receive antenna $i$, $X_j$ denotes the transmitted symbol at antenna $j$, and $V_i$ denotes the independent and identically distributed (i.i.d.) complex AWGN at receive antenna $i$. $V_i$ has zero mean and variance $\sigma^2$ per real dimension.

For VBLAST systems, generalized decision feedback equalizer (GDFE) with no error propagation is capacity lossless when the transmitter knows *a priori* the rate information for each streams [69]. For linear receivers, VBLAST will have poor performance when $n_R < n_T$ due to lack of degree of freedom [10][70], hence it is in general required that the receive antenna number is no less than the transmit antenna number. When this is the case, an iterative turbo receiver with capacity approaching FEC codes, e.g., turbo codes, low-density parity check (LDPC) codes, can have capacity approaching performance without requirement for any CSI-related information or rate information [71] [72].

### 3.2.2 GSTBC-OFDM

In a GSTBC-OFDM system [44][73][65], the transmitted signals are divided into groups. Spatial multiplexing is applied to signals among the groups; Within the group, Alamouti STBC is performed before the OFDM modulation. Therefore, an even number of transmit antennas is required to support GSTBC. However, even when the transmitter has an odd number of antennas, the concept of GSTBC is still applicable. In this case, the first $(n_T - 1)$ antennas can be used to transmit GSTBC-signals, and the $n_T$th antenna transmit is used to transmit the non-STBC coded signal, as in some variable rate STBC's proposed in [43]. However, in such a system, either rate feedback or power adjustment is needed in order to have optimal performance.

Another way of applying GSTBC in an odd-number transmit antenna system is to combine the GSTBC encoding with spatial spreading (SS) - encoding $\frac{n_T-1}{2}$ streams of data by Alamouti STBC, and then spreading the $(n_T-1)$ STBC-coded streams by SS to $n_T$ antennas. The SS is modeled in 3.2.5.

Here we first assume an even-numbered transmit antennas, and the signals $\mathbf{R}$, $\mathbf{H}$, $\mathbf{X}$ and $\mathbf{V}$ are defined as

$$\begin{cases} \mathbf{R} & \stackrel{\text{def}}{=} & \begin{bmatrix} R_{1,1} & -R_{1,2}^* & \cdots & R_{n_R,1} & -R_{n_R,2}^* \end{bmatrix}^T, \\ \mathbf{H} & \stackrel{\text{def}}{=} & \begin{bmatrix} H_{1,1} & H_{1,2} & \cdots & H_{1,n_T-1} & H_{1,n_T} \\ -H_{1,2}^* & H_{1,1}^* & \cdots & -H_{1,n_T}^* & H_{1,n_T-1}^* \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ H_{n_R,1} & H_{n_R,2} & \cdots & H_{n_R,n_T-1} & H_{n_R,n_T} \\ -H_{n_R,2}^* & H_{n_R,1}^* & \cdots & -H_{n_R,n_T}^* & H_{n_R,n_T-1}^* \end{bmatrix}, \\ \mathbf{X} & \stackrel{\text{def}}{=} & \begin{bmatrix} X_1 & X_2 & \cdots & X_{n_T} \end{bmatrix}^T, \\ \mathbf{V} & \stackrel{\text{def}}{=} & \begin{bmatrix} V_{1,1} & -V_{1,2}^* & \cdots & V_{n_R,1} & -V_{n_R,2}^* \end{bmatrix}^T, \end{cases} \tag{3.16}$$

where subscripts of $R$ and $V$ denote the receive antenna and OFDM symbol indices in each Alamouti STBC code word, $H_{i,j}$ and $X_j$ are defined the same as in a VBLAST-OFDM system.

In order to support linear detection, the number of receive antennas needs to satisfy the following relation

$$\min(n_R) \geq \frac{n_T}{2}.$$

**Hybrid of GSTBC and VBLAST** As we briefly mentioned, when the number of transmit antennas is not even, a hybrid of VBLAST and GSTBC precoding scheme can be applied. In this case, the first $n_T-1$ antennas transmit $\frac{n_T-1}{2}$ streams of GSTBC signals , and the last antenna transmits one independent stream of signal. The number of receive antennas needs to satisfy

$$\min(n_R) \geq \frac{n_T+1}{2}.$$

in order to support linear receivers.

The transmitted signal in the hybrid GSTBC-VBLAST system is

$$\begin{bmatrix} X_1 & -X_2^* \\ X_2 & X_1^* \\ \vdots & \vdots \\ X_{n_T-2} & -X_{n_T-1}^* \\ X_{n_T-1} & X_{n_T-2}^* \\ X_{n_T} & X_{n_T+1}^* \end{bmatrix},$$

i.e., the first $(n_T - 1)$ rows denote the Alamouti STBC-coded signals, and the last row denotes the spatial multiplexing signal.

Following the same way of manipulation as in GSTBC [44][73], we can develop the corresponding linear signal model. The signals $\mathbf{R}$, $\mathbf{X}$ and $\mathbf{V}$ are defined the same as in a GSTBC system, but the dimension of $\mathbf{X}$ changes to $n_T + 1$, and the precoded channel $\mathbf{H}$ is expressed as

$$\mathbf{H} \stackrel{\text{def}}{=} \begin{bmatrix} H_{1,1} & H_{1,2} & \cdots & H_{1,n_T-2} & H_{1,n_T-1} & H_{1,n_T} & 0 \\ -H_{1,2}^* & H_{1,1}^* & \cdots & -H_{1,n_T-1}^* & H_{1,n_T-2}^* & 0 & -H_{1,n_T}^* \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ H_{n_R,1} & H_{n_R,2} & \cdots & H_{n_R,n_T-2} & H_{n_R,n_T-1} & H_{n_R,n_T} & 0 \\ -H_{n_R,2}^* & H_{n_R,1}^* & \cdots & -H_{n_R,n_T-1}^* & H_{n_R,n_T-2}^* & 0 & -H_{n_R,n_T}^* \end{bmatrix}. \quad (3.17)$$

Obviously, due to lack of transmit diversity, the data rate supported by the last stream is lower than the other $\frac{n_T-1}{2}$ STBC-ed streams. One way to compensate for the diversity loss is higher power allocation to the non-STBC coded stream, as indicated in [43].

### 3.2.3 QSTBC-OFDM

For QSTBC, we focus on the $n_T = 4$ cases and consider the rate-1 code in (2.33) by Jafarkhani [38] and the rate-2 code by Yuen *et. al.* in [41].

For the rate-1 QSTBC coded $(n_T = 4) \times n_R$ system, the signals $\mathbf{R}$, $\mathbf{H}$, $\mathbf{X}$ and $\mathbf{V}$ are defined as

$$\left\{\begin{array}{lcl} \mathbf{R} & \stackrel{\text{def}}{=} & \begin{bmatrix} R_{1,1} & -R^*_{1,2} & R_{1,3} & -R^*_{1,4} & \cdots & R_{n_R,1} & -R^*_{n_R,2} & R_{n_R,3} & -R^*_{n_R,4} \end{bmatrix}^T, \\ \mathbf{H} & \stackrel{\text{def}}{=} & \begin{bmatrix} H_{1,1} & H_{1,2} & H_{1,3} & H_{1,4} \\ -H^*_{1,2} & H^*_{1,1} & -H^*_{1,4} & H^*_{1,3} \\ H_{1,3} & H_{1,4} & H_{1,1} & H_{1,2} \\ -H^*_{1,4} & H^*_{1,3} & -H^*_{1,2} & H^*_{1,1} \\ \vdots & \vdots & \vdots & \vdots \\ H_{n_R,1} & H_{n_R,2} & H_{n_R,3} & H_{n_R,4} \\ -H^*_{n_R,2} & H^*_{n_R,1} & -H^*_{n_R,4} & H^*_{n_R,3} \\ H_{n_R,3} & H_{n_R,4} & H_{n_R,1} & H_{n_R,2} \\ -H^*_{n_R,4} & H^*_{n_R,3} & -H^*_{n_R,2} & H^*_{n_R,1} \end{bmatrix}, \\ \mathbf{X} & \stackrel{\text{def}}{=} & \begin{bmatrix} X_1 & X_2 & X_3 & X_4 \end{bmatrix}^T, \\ \mathbf{V} & \stackrel{\text{def}}{=} & \begin{bmatrix} V_{1,1} & -V^*_{1,2} & V_{1,3} & -V^*_{1,4} & \cdots & V_{n_R,1} & -V^*_{n_R,2} & V_{n_R,3} & -V^*_{n_R,4} \end{bmatrix}^T, \end{array}\right. \tag{3.18}$$

where subscripts of $R$ and $V$ denote the receive antenna and OFDM symbol indices in each QSTBC code word, $H_{i,j}$ and $X_j$ are defined the same as in a VBLAST-OFDM system.

The rate-2 QSTBC in (2.34) can be generated by applying a $2 \times 2$ Walsh-Hadamard spatial-spreading matrix onto the GSTBC codeword, as

$$\mathbf{x}_{\text{QSTBC}} = \mathbf{W}_2 \mathbf{x}_{\text{GSTBC}} = \mathbf{W}_2 \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \end{bmatrix}, \tag{3.19}$$

where

$$\mathbf{W}_2 = \sqrt{\frac{1}{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

is the $2 \times 2$ Walsh-Hadamard matrix, $\mathbf{A}_i$, $i = 1,\ 2$ denotes the $i$th Alamouti codeword.

We therefore have

$$\left\{\begin{array}{lcl}
\mathbf{R} & \overset{\text{def}}{=} & \begin{bmatrix} R_{1,1} & -R_{1,2}^* & \cdots & R_{n_R,1} & -R_{n_R,2}^* \end{bmatrix}^T, \\
\mathbf{H} & \overset{\text{def}}{=} & \sqrt{\frac{1}{2}} \begin{bmatrix} H_{1,1}+H_{1,3} & H_{1,2}+H_{1,4} & H_{1,1}-H_{1,3} & H_{1,2}-H_{1,4} \\ -H_{1,2}^*-H_{1,4}^* & H_{1,1}^*+H_{1,3}^* & -H_{1,2}^*+H_{1,4}^* & H_{1,1}^*-H_{1,3}^* \\ \vdots & \vdots & \vdots & \vdots \\ H_{n_R,1}+H_{n_R,3} & H_{n_R,2}+H_{n_R,4} & H_{n_R,1}-H_{n_R,3} & H_{n_R,2}-H_{n_R,4} \\ -H_{n_R,2}^*-H_{n_R,4}^* & H_{n_R,1}^*+H_{n_R,3}^* & -H_{n_R,2}^*+H_{n_R,4}^* & H_{n_R,1}^*-H_{n_R,3}^* \end{bmatrix}, \\
\mathbf{X} & \overset{\text{def}}{=} & \begin{bmatrix} X_1 & X_2 & X_3 & X_4 \end{bmatrix}^T, \\
\mathbf{V} & \overset{\text{def}}{=} & \begin{bmatrix} V_{1,1} & -V_{1,2}^* & \cdots & V_{n_R,1} & -V_{n_R,2}^* \end{bmatrix}^T,
\end{array}\right. \tag{3.20}$$

where subscripts of $R$ and $V$ denote the receive antenna and OFDM symbol indices in each Alamouti STBC code word, $H_{i,j}$ and $X_j$ are defined the same as in a VBLAST-OFDM system.

**Remark** Comparing the GSTBC signal model in (3.16) and the rate-2 QSTBC model in (3.20), we can see that $\mathbf{R}$, $\mathbf{X}$ and $\mathbf{V}$ are defined the same way in the two systems. A closer look at the channel matrix $\mathbf{H}$ definition for the two systems leads to the following linear relation:

$$\mathbf{H}_{\text{QSTBC}} = \mathbf{H}_{\text{GSTBC}}\mathbf{T}, \tag{3.21}$$

where $\mathbf{T}$ is a $4 \times 4$ matrix defined as

$$\mathbf{T} = \sqrt{\frac{1}{2}} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix} = \mathbf{W}_2 \otimes \mathbf{I}_2, \tag{3.22}$$

where $\mathbf{W}_2$ is the $2\times2$ Walsh-Hadamard matrix, $\mathbf{I}_2$ the $2\times2$ identity matrix, and $\otimes$, the Kronecker product.

Taking note that $\mathbf{T}\mathbf{T}^H = \mathbf{T}^H\mathbf{T} = \mathbf{I}$, we therefore have

$$\mathbf{H}_{\text{QSTBC}}\mathbf{H}_{\text{QSTBC}}^H = \mathbf{H}_{\text{GSTBC}}\mathbf{H}_{\text{GSTBC}}^H. \tag{3.23}$$

### 3.2.4 LDC-OFDM

We consider the LDC-OFDM system in which each subcarrier is independently encoded by the same linear dispersion matrices and the channel is quasi-static within the LDC codeword. We hence have a real signal model in which the equivalent signals in (3.14) are defined as follows (*cf. Eqns. (22) - (25)* in [66])

$$\begin{cases} \mathbf{R} \overset{\text{def}}{=} \begin{bmatrix} R_{Re,1} & R_{Im,1} & \cdots & R_{Re,n_RT} & R_{Im,n_RT} \end{bmatrix}^T, \\ \mathbf{H} \overset{\text{def}}{=} \begin{bmatrix} \mathcal{A}_1\underline{\mathbf{H}}_1 & \mathcal{B}_1\underline{\mathbf{H}}_1 & \cdots & \mathcal{A}_Q\underline{\mathbf{H}}_1 & \mathcal{B}_Q\underline{\mathbf{H}}_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathcal{A}_1\underline{\mathbf{H}}_{n_R} & \mathcal{B}_1\underline{\mathbf{H}}_{n_R} & \cdots & \mathcal{A}_Q\underline{\mathbf{H}}_{n_R} & \mathcal{B}_Q\underline{\mathbf{H}}_{n_R} \end{bmatrix}, \\ \mathbf{X} \overset{\text{def}}{=} \begin{bmatrix} X_{Re,1} & X_{Im,1} & \cdots & X_{Re,Q} & X_{Im,Q} \end{bmatrix}^T, \\ \mathbf{V} \overset{\text{def}}{=} \begin{bmatrix} V_{Re,1} & V_{Im,1} & \cdots & V_{Re,n_RT} & V_{Im,n_RT} \end{bmatrix}^T, \end{cases} \tag{3.24}$$

where subscripts "$Re$" and "$Im$" denote real and imaginary parts of the signal, $T$ is the LDC codeword interval, $Q$ is the total number of symbols transmitted by one LDC codeword, and $\mathcal{A}_q$ and $\mathcal{B}_q$, $q = 1,\ 2,\ \cdots,\ Q$ are generated from the LDC encoding matrices $\mathbf{A}_q$ and $\mathbf{B}_q$ as

$$\mathcal{A}_q = \begin{bmatrix} \mathbf{A}_{Re,q} & -\mathbf{A}_{Im,q} \\ \mathbf{A}_{Im,q} & \mathbf{A}_{Re,q} \end{bmatrix},$$

$$\mathcal{B}_q = \begin{bmatrix} -\mathbf{B}_{Im,q} & -\mathbf{B}_{Re,q} \\ \mathbf{B}_{Re,q} & -\mathbf{B}_{Im,q} \end{bmatrix},$$

that is, $\mathcal{A}_q$ and $\mathcal{B}_q$ are both of dimension of $2n_RT \times 2n_T$, and $\underline{\mathbf{H}}_n$ is a column vector of dimension $2n_T$ generated from the channel response corresponding to the $n$th receive antenna as

$$\underline{\mathbf{H}}_n = \begin{bmatrix} H_{Re,n,1} & H_{Re,n,2} & \cdots & H_{Re,n,n_T} & H_{Im,n,1} & H_{Im,n,2} & \cdots & H_{Im,n,n_T} \end{bmatrix}^T.$$

The precoding rate of LDC is $R_{P,\text{LDC}} = Q/T$.

### 3.2.5 CDDSS-OFDM

Cyclic-delay diversity spatial spreading (CDDSS) is an open loop precoding scheme which can map $n_S$ streams of data to $n_T$ transmit antennas, $n_T \geq n_S$. It is a combination of SS and CDD. CDD was first proposed by Kaiser in [62] as a transmit diversity scheme for OFDM systems, as an extension of the delay diversity scheme proposed for single-carrier modulation systems. Kaiser also proved in [62] that CDD is equivalent to phase diversity (PD) and in [61] that the operation of CDD is transparent to the receiver.

Before proceeding to develop the signal model for CDDSS-OFDM, we first develop the CDD-OFDM signal model for $n_T = 2$ and $n_R = 1$ system.

**CDD-OFDM**

Denoting the IFFT output of the original transmitted signal as

$$\mathbf{x}_t = \mathbf{F}^H \mathbf{X} = \begin{bmatrix} x_t^0 & x_t^1 & \cdots & x_t^{N-1} \end{bmatrix}^T,$$

and the cyclic delay value as $s$, we have the cyclic delayed signal sequence as

$$\begin{aligned} \mathbf{x}_t^{\mathrm{CDD}} &= \begin{bmatrix} x_t^s & x_t^{s+1} & \cdots & x_t^{N-1} & x_t^0 & x_t^1 & \cdots & x_t^{s-1} \end{bmatrix}^T \\ &= \mathbf{P}^{\mathrm{CDD}} \mathbf{x}_t, \end{aligned} \tag{3.25}$$

where $\mathbf{P}^{\mathrm{CDD}}$ is a circulant matrix with its first row as

$$\begin{pmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{pmatrix}$$

i.e., only its $s$th element is “1” and all the other elements are “0”.

Then CP is appended to both the original time domain sequence $\mathbf{x}_t$ and the cyclic delayed signal sequence $\mathbf{x}_t^{\mathrm{CDD}}$ as

$$\begin{aligned} \mathbf{x} &= \mathbf{T}_{\mathrm{CP}} \mathbf{x}_t = \mathbf{T}_{\mathrm{CP}} \mathbf{F}^H \mathbf{X}, \\ \mathbf{x}^{\mathrm{CDD}} &= \mathbf{T}_{\mathrm{CP}} \mathbf{x}_t^{\mathrm{CDD}} = \mathbf{T}_{\mathrm{CP}} \mathbf{P}^{\mathrm{CDD}} \mathbf{F}^H \mathbf{X}. \end{aligned}$$

The corresponding received signal after removing the CP portion is therefore written as

$$\mathbf{y}^{\mathrm{CDD}} = \mathbf{R}_{\mathrm{CP}} \mathbf{h}_{0,1,1} \mathbf{x} + \mathbf{R}_{\mathrm{CP}} \mathbf{h}_{0,1,2} \mathbf{x}^{\mathrm{CDD}} + \mathbf{v} \tag{3.26}$$

$$= \mathbf{R}_{\mathrm{CP}} \mathbf{h}_{0,1,1} \mathbf{T}_{\mathrm{CP}} \mathbf{F}^H \mathbf{X} + \mathbf{R}_{\mathrm{CP}} \mathbf{h}_{0,1,2} \mathbf{T}_{\mathrm{CP}} \mathbf{P}^{\mathrm{CDD}} \mathbf{F}^H \mathbf{X} + \mathbf{v} \tag{3.27}$$

following (3.4). Here $\mathbf{h}_{0,1,1}$ and $\mathbf{h}_{0,1,2}$ are the Toeplitz channel matrices corresponding to receive antenna 1, and transmit antenna 1 and 2, respectively.

From (3.26) and (3.27), it is obvious that CDD does not incur any ISI.

From Section 3.1, $(\mathbf{R}_{\mathrm{CP}}\mathbf{h}_{0,1,1}\mathbf{T}_{\mathrm{CP}})$ and $(\mathbf{R}_{\mathrm{CP}}\mathbf{h}_{0,1,2}\mathbf{T}_{\mathrm{CP}})$ are both $N \times N$ circulant matrices with their first column respectively written as

$$\mathbf{h}_{t,1,1} = [h_{0,1,1},\ h_{1,1,1}, \cdots,\ h_{L-1,1,1},\ 0, \cdots,\ 0]^T$$

and

$$\mathbf{h}_{t,1,2} = [h_{0,1,2},\ h_{1,1,2}, \cdots,\ h_{L-1,1,2},\ 0, \cdots,\ 0]^T.$$

Matrix $\left(\mathbf{R}_{\mathrm{CP}}\mathbf{h}_{0,1,2}\mathbf{T}_{\mathrm{CP}}\mathbf{P}^{\mathrm{CDD}}\right)$ is therefore also a circulant matrix and its first column is the $s$'th column of $(\mathbf{R}_{\mathrm{CP}}\mathbf{h}_{0,1,2}\mathbf{T}_{\mathrm{CP}})$, written as

$$\mathbf{h}_{t,1,2}^{\mathrm{CDD}} = \begin{cases} \Big[\underbrace{0,\ \cdots,\ 0}_{s}\ h_{0,1,2},\ h_{1,1,2},\ \cdots,\ h_{L-1,1,2},\ \underbrace{0, \cdots,\ 0}_{N-L-s}\Big]^T, & s \leq (N-L), \\ \Big[\underbrace{h_{N-s,1,2},\ \cdots,\ h_{L,1,2}}_{L+s-N}, \underbrace{0,\ \cdots,\ 0}_{N-L}\ \underbrace{h_{0,1,2},\ h_{1,1,2},\ \cdots,\ h_{N-s-1,1,2}}_{N-s}\Big]^T, & s > (N-L), \end{cases} \tag{3.28}$$

i.e., elements of the first column of circulant matrix $\left(\mathbf{R}_{\mathrm{CP}}\mathbf{h}_{0,1,2}\mathbf{T}_{\mathrm{CP}}\mathbf{P}^{\mathrm{CDD}}\right)$ are written as

$$h_{t,1,2}^{\mathrm{CDD}}(n) = h_{t,1,2}\left[(n-s) \mod N\right], \quad n = 0,\ 1, \cdots,\ N-1. \tag{3.29}$$

We can therefore re-write (3.27) as

$$\mathbf{y}^{\mathrm{CDD}} = \mathbf{h}_{\mathrm{equiv}}^{\mathrm{CDD}}\mathbf{F}^H\mathbf{X} + \mathbf{v} \tag{3.30}$$

where $\mathbf{h}_{\mathrm{equiv}}^{\mathrm{CDD}}$ is $N \times N$ circulant matrix with its first column written as

$$\mathbf{h}_{t,\mathrm{equiv}}^{\mathrm{CDD}} = \begin{cases} [h_{0,1,1}, \cdots, h_{s-1,1,1},\ h_{s,1,1} + h_{0,1,2}, \cdots, h_{L-1,1,1} + h_{L-1-s,1,2}, \\ \qquad h_{L-s,1,2}, \cdots, h_{L-1,1,2},\ 0, \cdots, 0]^T, & s < L, \\ [h_{0,1,1}, h_{1,1,1}, \cdots, h_{L-1,1,1}, 0, \cdots, 0, h_{0,1,2},\ h_{1,1,2}, \cdots,\ h_{L-1,1,2},\ 0, \cdots,\ 0]^T, & L \leq s \leq (N-L), \\ [h_{0,1,1} + h_{N-s,1,2}, \cdots, h_{L+s-N-1,1,1} + h_{L-1,1,2},\ h_{L+s-N,1,1}, \cdots, \\ \qquad h_{L-1,1,1}, 0, \cdots, 0, h_{0,1,2},\ h_{1,1,2},\ \cdots,\ h_{N-s-1,1,2}]^T, & s > (N-L), \end{cases} \tag{3.31}$$

i.e., for $0 < s < N$, *additional multipaths* have been created through CDD. Therefore, higher frequency diversity is made possible through CDD.

Making use of (3.29), we can alternatively write the elements in the first column of circulant matrix $\mathbf{h}_{\text{equiv}}^{\text{CDD}}$ as

$$\begin{aligned} h_{t,\text{equiv}}^{\text{CDD}}(n) &= h_{t,1,1}(n) + h_{t,1,2}^{\text{CDD}}(n) \\ &= h_{t,1,1}(n) + h_{t,1,2}\left[(n-s) \mod N\right] \end{aligned} \tag{3.32}$$

where $n = 0,\ 1, \cdots,\ N-1$.

For $n_T = 2$ system, we can see that to fully exploit the frequency diversity potential in CDD, the cyclic delay $s$ is preferably in the range $L \geq s < N - L$.

Performing FFT on (3.27), we obtain the frequency domain signal model, as

$$\mathbf{Y}^{\text{CDD}} = \mathbf{H}_{11}\mathbf{X} + \mathbf{H}_{12}^{\text{CDD}}\mathbf{X} + \mathbf{V}, \tag{3.33}$$

where $\mathbf{H}_{11}$ and $\mathbf{H}_{12}^{\text{CDD}}$ are $N \times N$ diagonal matrices, with the $k$th diagonal element for $\mathbf{H}_{11}$ as

$$H_{11,k} = \sum_{l=0}^{L} h_{11,l} e^{-j\frac{2\pi}{N}kl}$$

and the $k$th diagonal element for $\mathbf{H}_{12}^{\text{CDD}}$ as

$$H_{12,k}^{\text{CDD}} = \begin{cases} \sum_{l=0}^{L-1} h_{12,l} e^{-j\frac{2\pi}{N}k(l+s)} = e^{-j\frac{2\pi}{N}ks} \sum_{l=0}^{L-1} h_{12,l} e^{-j\frac{2\pi}{N}kl} = e^{-j\frac{2\pi}{N}ks} H_{12,k}, & s \leq (N-L), \\ \sum_{l=0}^{N-s-1} h_{12,l} e^{-j\frac{2\pi}{N}k(l+s)} + \sum_{l=N-s}^{L-1} h_{12,l} e^{-j\frac{2\pi}{N}k(l+s-N)} & \\ \qquad = \sum_{l=0}^{L-1} h_{12,l} e^{-j\frac{2\pi}{N}k(l+s)} = e^{-j\frac{2\pi}{N}ks} H_{12,k}, & s > (N-L). \end{cases}$$

We hence have the frequency domain channel expressed as

$$\mathbf{H}_{12}^{\text{CDD}} = \mathbf{\Psi}\mathbf{H}_{12} \tag{3.34}$$

where

$$\mathbf{\Psi} = \text{diag}\left(\ 1,\quad e^{-j\frac{2\pi}{N}s},\quad \ldots,\quad e^{-j\frac{2\pi}{N}ks},\quad \ldots,\quad e^{-j\frac{2\pi}{N}(N-1)s}\ \right). \tag{3.35}$$

(3.33) can therefore be re-written as

$$\mathbf{Y}^{\mathrm{CDD}} = \left(\mathbf{H}_{11} + \mathbf{\Psi}\mathbf{H}_{12}\right)\mathbf{X} + \mathbf{V}, \tag{3.36}$$

For each subcarrier, we have the received signal expressed as

$$Y_k = \left(H_{11,k} + e^{-j\frac{2\pi}{N}ks}H_{12,k}\right)X_k + V_k = \mathbf{H}_k\mathbf{\Phi}_k\mathbf{W}X_k + V_k, \tag{3.37}$$

where

$$\begin{aligned}
\mathbf{H}_k &= \begin{bmatrix} H_{11,k} & H_{12,k} \end{bmatrix}, \\
\mathbf{\Phi}_k &= \mathrm{diag}\begin{pmatrix} 1, & e^{-j\frac{2\pi}{N}ks} \end{pmatrix}, \\
\mathbf{W} &= \begin{bmatrix} 1 \\ 1 \end{bmatrix}.
\end{aligned}$$

**Remark** Matrix $\mathbf{W}$ can be seen as a spatial spreading matrix, which spreads the single-stream transmitted signal $X_k$ to $n_T = 2$ antennas.

The $n_T = 2$ time-domain channel model in (3.32) and frequency domain model in (3.37) can be extended to $n_T > 2$ CDD transmit diversity system in a straightforward manner. Assuming the incremental cyclic delay per antenna is $\tau$, then the cyclic delay applied to the $n_t$'th antenna is $(n_t - 1)\tau$. We hence have the time-domain equivalent channel as

$$h_{t,\mathrm{equiv}}^{\mathrm{CDD}}(n) = \sum_{n_t=1}^{n_T} h_{t,1,n_t}\left[(n - (n_t - 1)\tau) \mod N\right], \tag{3.38}$$

and the frequency domain channel as

$$\mathbf{Y}^{\mathrm{CDD}} = \sum_{n_t=1}^{n_T} \mathbf{\Psi}_{n_t}\mathbf{H}_{1n_t}\mathbf{X} + \mathbf{V}, \tag{3.39}$$

where $\mathbf{\Psi}_{n_t}$'s are diagonal matrices with the $k$th diagonal element $\psi_{n_t}(k)$ written as

$$\psi_{n_t}(k) = e^{-j\frac{2\pi}{N}(n_t-1)\tau k}.$$

We hence have the signal model at each subcarrier as

$$Y_k = \mathbf{H}_k\mathbf{\Phi}_k\mathbf{W}X_k + V_k, \tag{3.40}$$

where

$$
\begin{aligned}
\mathbf{H}_k &= \begin{bmatrix} H_{11,k} & H_{12,k} & \cdots & H_{1,n_T,k} \end{bmatrix}, \\
\mathbf{\Phi}_k &= \mathrm{diag}\left( 1, \quad e^{-j\frac{2\pi}{N}k\tau}, \quad \cdots, \quad e^{-j\frac{2\pi}{N}k(n_T-1)\tau} \right), \\
\mathbf{W} &= \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}.
\end{aligned}
$$

**CDDSS-OFDM**

As we have shown, in a CDD system, the single-stream signal is first spread to $n_T$ antennas with equal energy. After that, a cyclic shift is applied to the time domain signals transmitted over $n_t > 1$ antennas to transfer the spatial domain freedom to frequency diversity. When the number of transmitted spatial streams of signals $n_S$ is larger than one but smaller than the number of transmit antennas $n_T$, orthonormal spatial spreading with cyclic delay diversity, i.e., CDDSS, can be used to map the $n_S$ streams of signals to $n_T$ transmit antennas and to transfer the additional spatial domain freedom to frequency diversity. In this case, the signals in (3.14) are written as

$$
\begin{cases}
\mathbf{R} \overset{\text{def}}{=} \begin{bmatrix} R_1 & R_2 & \cdots & R_{n_R} \end{bmatrix}^T, \\
\mathbf{H} \overset{\text{def}}{=} \begin{bmatrix} H_{1,1} & H_{1,2} & \cdots & H_{1,n_T} \\ \vdots & \vdots & \vdots & \vdots \\ H_{n_R,1} & H_{n_R,2} & \cdots & H_{n_R,n_T} \end{bmatrix} \mathbf{\Phi W}, \\
\mathbf{X} \overset{\text{def}}{=} \begin{bmatrix} X_1 & X_2 & \cdots & X_{n_S} \end{bmatrix}^T, \\
\mathbf{V} \overset{\text{def}}{=} \begin{bmatrix} V_1 & V_2 & \cdots & V_{n_R} \end{bmatrix}^T,
\end{cases}
\tag{3.41}
$$

where $\mathbf{\Phi}$ denotes the $n_T \times n_T$ diagonal phase rotation matrix introduced by CDD with the $n_t$'th diagonal element

$$
\phi_{n_t} = \exp\left[ -j\frac{2\pi}{N}(n_t - 1)k\tau \right],
$$

with $\tau$ being the incremental cyclic delay value, N the FFT size, and $k$ the subcarrier index.

Matrix $\mathbf{W}$ is the spatial spreading matrix of dimension $n_T \times n_S$, corresponding to the first $n_S$ columns of a $n_T \times n_T$ unitary matrix.

**Theorem 4.** *In a CDDSS-OFDM system, each individual data stream is transmitted through CDD.*

*Proof:* From (3.41), the received signal corresponding to spatial stream $n_s$, $1 \le n_s \le n_S$, is written as

$$\mathbf{R}_k^{n_s} = \mathbf{H}_k \mathbf{\Phi} \mathbf{W}_{n_s} X_k^{n_s} + \mathbf{V}_k$$

where $\mathbf{W}_{n_s}$ denotes the $n_s$th column of $\mathbf{W}$, and $(\mathbf{W}_{n_s})^H \mathbf{W}_{n_s} = 1$.

Comparing with (3.40), the conclusion is straightforward. ■

**Theorem 5.** *The maximum frequency diversity order for each spatial stream in the CDDSS MIMO-OFDM system is* $\min(Ln_T,\ N)$*, achieved when the cyclic delay* $\tau = L$*, with* $L$ *being the number of multipaths in each transmit-receive antenna pair channel,* $n_T$ *the number of transmit antennas,* $N$ *the number of subcarriers in OFDM.*

*Proof:* From Theorem 4, we only need to prove that the maximum frequency diversity order for CDD-OFDM is $\min(Ln_T,\ N)$ when $\tau = L$.

From (3.38), when $\tau = L$, and $n_T L \le N$, we have the equivalent time-domain channel as

$$\begin{aligned} h_{t,\text{equiv}}^{\text{CDD}}(n) &= \sum_{n_t=1}^{n_T} h_{t,1,n_t}\left[(n-(n_t-1)L) \mod N\right] \\ &= \begin{cases} h_{t,1,1}(n) & 0 \le n \le L-1 \\ h_{t,1,2}(n-L) & L \le n \le 2L-1 \\ \dots & \\ h_{t,1,n_T}(n-(n_T-1)L) & (n_T-1)L \le n \le n_T L - 1 \\ 0 & n_T L < n \end{cases} \end{aligned}$$

i.e., there are $n_T L$ multipaths in the equivalent time-domain channel.

When $n_T L > N$, the number of multipaths is $N$.

Therefore, the maximum achievable frequency diversity order is $\min(n_T L, N)$. ■

### 3.2.6 RAS-OFDM

When the receiver has a lot more antennas than the transmitter, receive antenna selection (RAS) can be performed. Due to the fact that the extra $(n_R - n_T)$ receive antennas provide only higher diversity, and that when the diversity order is getting higher, the additional SNR gain becomes smaller [9], RAS can significantly reduce the hardware cost yet maintain a negligible performance loss. In [74], Molisch *et. al.* further looked into the ergodic capacity of antenna selection system and showed that the achieved capacity of the $n_T \times n_T$ RAS system is close to the full $n_T \times n_R$ system where $n_R > n_T$.

For RAS spatial multiplexing systems selecting $L_R$ "best" out of $n_R$ available antennas, the corresponding signals in (3.14) are written as

$$\begin{cases} \mathbf{R} \stackrel{\text{def}}{=} \begin{bmatrix} R_1 & R_2 & \cdots & R_{L_R} \end{bmatrix}^T, \\ \mathbf{H} \stackrel{\text{def}}{=} \mathbf{S} \begin{bmatrix} H_{1,1} & H_{1,2} & \cdots & H_{1,n_T} \\ \vdots & \vdots & \vdots & \vdots \\ H_{n_R,1} & H_{n_R,2} & \cdots & H_{n_R,n_T} \end{bmatrix}, \\ \mathbf{X} \stackrel{\text{def}}{=} \begin{bmatrix} X_1 & X_2 & \cdots & X_{n_T} \end{bmatrix}^T, \\ \mathbf{V} \stackrel{\text{def}}{=} \begin{bmatrix} V_1 & V_2 & \cdots & V_{L_R} \end{bmatrix}^T, \end{cases} \tag{3.42}$$

where $\mathbf{S}$ denotes the receive antenna selection matrix of size $L_R \times n_R$. $\mathbf{S}$ is constructed from the rows of the $n_R \times n_R$ identity matrix $\mathbf{I}_{n_R}$, i.e., only those rows corresponding to the selected antenna indices will be taken to form $\mathbf{S}$.

### 3.2.7 TAS-OFDM

Similar to RAS, transmit antenna selection (TAS) can be performed at the transmitter. Different from RAS, TAS is a closed loop precoding scheme. Feedback information is needed from the receiver as to which are the "best" antennas for transmission, based on the pre-determined antenna selection criteria. For TAS in spatial multiplexing systems [75, 76], $n_R$ out of $n_T$ antennas will be selected, and the corresponding signals in (3.14) are written as

$$\begin{cases} \mathbf{R} \overset{\text{def}}{=} \begin{bmatrix} R_1 & R_2 & \cdots & R_{n_R} \end{bmatrix}^T, \\ \mathbf{H} \overset{\text{def}}{=} \begin{bmatrix} H_{1,1} & H_{1,2} & \cdots & H_{1,n_T} \\ \vdots & \vdots & \vdots & \vdots \\ H_{n_R,1} & H_{n_R,2} & \cdots & H_{n_R,n_T} \end{bmatrix} \mathbf{S}, \\ \mathbf{X} \overset{\text{def}}{=} \begin{bmatrix} X_1 & X_2 & \cdots & X_{n_R} \end{bmatrix}^T, \\ \mathbf{V} \overset{\text{def}}{=} \begin{bmatrix} V_1 & V_2 & \cdots & V_{n_R} \end{bmatrix}^T, \end{cases} \tag{3.43}$$

where $\mathbf{S}$ denotes the transmit antenna selection matrix of size $n_T \times n_R$, and $\mathbf{S}$ is constructed from the columns of the $n_R \times n_R$ identity matrix $\mathbf{I}_{n_R}$, i.e., only those columns corresponding to the selected antenna indices will be taken to form $\mathbf{S}$.

For asymmetric downlink MIMO channels with more transmit than receive antennas due to the size limitation and power consumption constraint of the terminal, TAS reduces the hardware cost yet maintaining the same diversity order as a full-antenna system [76][77].

### 3.2.8 SVD-OFDM

When CSI is perfectly known at the transmitter, SVD can be applied to fully decouple the MIMO channels. For MIMO-OFDM systems, the suboptimal sub-channel grouping and statistical water-filling technique proposed in [67] was proved to asymptotically achieve the ergodic channel capacity. For SVD-MIMO-OFDM, the equivalent signals in (3.14) are written as

$$\begin{cases} \mathbf{R} \overset{\text{def}}{=} \begin{bmatrix} R_1 & R_2 & \cdots & R_{n_R} \end{bmatrix}^T, \\ \mathbf{H} \overset{\text{def}}{=} \begin{bmatrix} H_{1,1} & H_{1,2} & \cdots & H_{1,n_T} \\ \vdots & \vdots & \vdots & \vdots \\ H_{n_R,1} & H_{n_R,2} & \cdots & H_{n_R,n_T} \end{bmatrix} \mathbf{V}, \\ \mathbf{X} \overset{\text{def}}{=} \begin{bmatrix} X_1 & X_2 & \cdots & X_{n_T} \end{bmatrix}^T, \\ \mathbf{V} \overset{\text{def}}{=} \begin{bmatrix} V_1 & V_2 & \cdots & V_{n_R} \end{bmatrix}^T, \end{cases} \tag{3.44}$$

where $\mathbf{H} = \mathbf{U}\mathbf{D}\mathbf{V}^H$ is the SVD of the channel matrix $\mathbf{H}$.

For ordered MIMO-OFDM channels, the coded modulation scheme, coding rate and modulation order for each of the ordered channels needs to be changed adaptively to maximize the achievable throughput for a given power budget and bit error rate performance requirement. The system depicted in Fig. 3.2 can be taken as a multiple codebook variable rate (MCVR) BICM solution in which each grouped channel may use different coding scheme with different coding rate, and different modulation order to adapt to the corresponding available diversity order and SNR in the channel. However, MCVR BICM solution may suffer some performance loss from the highest order grouped channel, as suggested in [68]. This is due to the fact that the highest order grouped channel has the highest diversity order among all the groups, making it close to a Gaussian channel. In this case, trellis coded modulation (TCM) which is optimized for the AWGN channel will have better performance than BICM. Besides MCVR, another system structure, called single codebook constant rate (SCCR) [78], can also be used. In SCCR SVD-OFDM, the various grouped-channels make use of the same coding and modulation order, and the achievable throughput is optimized through power loading.

## 3.3 Summary of the Chapter

A generalized linear system model has been developed for space-time-frequency coded MIMO-OFDM systems. This generalized linear signal model will facilitate the capacity analysis in Chapter 4 and the receiver design in Chapter 5.

# Chapter 4

# Precoding in Asymmetric MIMO-OFDM Channels

Based on the generalized linear signal model for the various STFP precoded MIMO-OFDM systems in Chapter 3, we will study their capacity, diversity and bit/frame error rate (BER/FER) performance in this chapter. We will put our special focus on the asymmetric channels with more transmit than receive antennas. They are typically created for downlink transmission when the terminal station can not accommodate as many antennas as the base station or access point (AP) due to size limitation and power consumption constraints. Among such asymmetric channels, it has been proven in [66] that the Alamouti STBC is both capacity and diversity optimal for $2 \times 1$ configurations. However, for other MIMO configurations, we will show that the known precoding schemes are either capacity lossy, or diversity lossy, or both capacity and diversity lossy. We then propose a 2DLPT MIMO-OFDM system which can fully exploit the spatial and frequency diversities available in the MIMO-OFDM channels. We will also prove that the proposed 2DLPT achieves full capacity when the number of spatial streams is set equal to the number of transmit antennas.

## 4.1 The Ergodic Capacity of MIMO-OFDM Systems

For an $n_T \times n_R$ direct mapping MIMO-OFDM system modeled in (3.10), we can have its capacity written as

$$C \stackrel{\text{def}}{=} \frac{1}{N + L_{\text{CP}}} \log_2 \left( \mathbf{I} + \frac{\rho}{n_T} \boldsymbol{\mathcal{H}} \boldsymbol{\mathcal{H}}^H \right), \tag{4.1}$$

when the channel input $\boldsymbol{\mathcal{X}}$ is i.i.d. Gaussian. In the above equation, the capacity loss due to the CP has been taken into account and it has been assumed that all the $N$ subcarriers are active. Here $\rho$ is the receive SNR at each subcarrier and each antenna.

Alternatively, we can use the signal model in (3.14) to derive capacity of the DM and other precoded MIMO-OFDM channels. Assuming

- no spatial correlation at both the transmitter and the receiver[1];
- zero-mean WSSUS CSCG multipath MIMO channels, i.e., Rayleigh fading channels;
- perfect CSI at receiver but no CSI at transmitter;
- i.i.d. Gaussian transmitted signals;
- all $N$ subcarriers used to transmit information;
- capacity loss due to CP not taken into account;

we have the MIMO-OFDM channel capacity written as [52]

$$\begin{aligned} C &= \frac{1}{N} \sum_{k=0}^{N-1} C_k \\ &= \frac{1}{N} \sum_{k=0}^{N-1} \frac{1}{\beta} \log_2 \det \left( \mathbf{I} + \frac{\rho}{n_T} \mathbf{H}_k \mathbf{H}_k^H \right) \\ &= \frac{1}{N\beta} \sum_{k=0}^{N-1} \sum_{i=1}^{M} \log_2 (1 + \frac{\rho}{n_T} \lambda_{k,i}) \ \text{bits/channel use}, \end{aligned}$$

where $k$ is the subcarrier index, $\lambda_{k,i}$ is the $i$th eigenvalue of $\mathbf{H}_k \mathbf{H}_k^H$, $\rho$ is the receive SNR at each subcarrier and each antenna, $M$ is the rank of $\mathbf{H}_k$, and $\beta$ is the channel use per precoding interval. For the open-loop

[1] This is possible for indoor wireless channels, which has rich local scatterers at both the transmitter and receiver. To enable this condition, the antenna separation needs to be wide enough as well.

precoding schemes discussed in Chapter 3, we have

$$\beta = \begin{cases} 1, & \text{DM and CDDSS,} \\ 2, & \text{GSTBC and QSTBC,} \\ 2T, & \text{LDC,} \end{cases}$$

with $T$ being the number of symbol intervals per LDC code word.

From Theorem 3, the MIMO channels at different subcarriers are statistically the same, hence the capacity at each subcarriers are statistically the same. We therefore have the ergodic capacity of the MIMO-OFDM channel as [52]

$$C_{\mathrm{E}} = \frac{1}{\beta}\sum_{i=1}^{M}\mathcal{E}\left[\log_2(1+\frac{\rho\lambda_i}{n_T})\right] \text{ bits/channel use.}$$

From Corollary 1 in Chapter 3, we have independent CSCG channels at each subcarrier when the time-domain multipaths are WSSUS zero-mean complex Gaussian and spatially uncorrelated. Therefore, the ergodic capacity can be easily computed for the DM-OFDM systems with CSCG channels by making use of the results from Appendix 2A.

Next we will prove that results in Appendix 2A can also be directly applied to compute the ergodic capacity of CDDSS CSCG MIMO channels. We will prove that the ergodic capacity of $n_T \times n_R \times n_S$ ($n_T > n_R \geq n_S$) CDDSS-MIMO channel is equal to that of $n_S \times n_R$ DM channel, when the MIMO channel coefficients are CSCG with i.i.d. elements of $CN(0,1)$.

### 4.1.1 Ergodic Capacity of CDDSS MIMO-OFDM Channels

Denoting the $n_T \times n_R$ ($n_R < n_T$) propagation channel as $\mathbf{H}^{\mathrm{p}}$ where superscript $(\cdot)^{\mathrm{p}}$ stands for propagation, the $n_S \times n_R$ CDDSS MIMO channel as $\mathbf{H}^{\mathrm{s}} = \mathbf{H}^{\mathrm{p}}\mathbf{S}$ with $\mathbf{S}$ being the $n_T \times n_S$ orthonormal CDDSS matrix, i.e., $\mathbf{S}^H\mathbf{S} = \mathbf{I}_{n_S}$, and the $n_S \times n_R$ DM channel as $\mathbf{H}^{\mathrm{d}}$, and assuming that $\mathbf{H}^{\mathrm{p}}$ is CSCG with i.i.d. elements of $CN(0,1)$, then $\mathbf{H}^{\mathrm{d}} = \mathbf{H}^{\mathrm{p}}(:, 1 : n_S)$[2] is also CSCG with i.i.d. elements of $CN(0,1)$.

[2]For $N \times M$ matrix $\mathbf{A}$, notation $\mathbf{A}(:, P_1 : P_2)$ with $P_1 \geq 1$ and $P_2 \leq M$ denotes submatrix of $\mathbf{A}$ with all its rows but only columns of $P_1$ to $P_2$, and notation $\mathbf{A}(Q_1 : Q_2, :)$ with $Q_1 \geq 1$ and $Q_2 \leq N$ denotes submatrix of $\mathbf{A}$ with all its columns but only rows of $Q_1$ to $Q_2$.

When $n_S = n_T$, the CDDSS matrix $\mathbf{S} = \mathbf{\Phi W}$ is unitary, i.e., $\mathbf{S}^H\mathbf{S} = \mathbf{S}\mathbf{S}^H = \mathbf{I}_{n_T}$. Therefore, matrices $\mathbf{H}^{\mathrm{p}}$ and $\mathbf{H}^{\mathrm{s}}$ are statistically the same. It is hence straightforward that the CDDSS has the same ergodic capacity as $n_T \times n_R$ DM, i.e., CDDSS achieves full capacity.

When $n_S < n_T$, $\mathbf{S}$ is orthonormal but not unitary matrix. So we have to look at the statistical properties of individual elements $H^{\mathrm{s}}(m,n)$ in $\mathbf{H}^{\mathrm{s}}$, with $m = 1, \cdots, n_R$, and $n = 1, \cdots, n_S$. As

$$H^{\mathrm{s}}(m,n) = \sum_{i=1}^{n_T} H^{\mathrm{p}}(m,i)S(i,n),$$

$H^{\mathrm{s}}(m,n)$ is still complex Gaussian. Its statistical properties are determined by its first- and second-order moments, which are derived as

$$\begin{aligned}
\mathcal{E}\left\{H^{\mathrm{s}}(m,n)\right\} &= \mathcal{E}\left\{\sum_{i=1}^{n_T} H^{\mathrm{p}}(m,i)S(i,n)\right\} = \sum_{i=1}^{n_T}\mathcal{E}\left\{H^{\mathrm{p}}(m,i)\right\}S(i,n) = 0,\\
\mathcal{E}\left\{|H^{\mathrm{s}}(m,n)|^2\right\} &= \mathcal{E}\left\{\left[\sum_{i=1}^{n_T} H^{\mathrm{p}}(m,i)S(i,n)\right]\left[\sum_{i=1}^{n_T} H^{\mathrm{p}}(m,i)S(i,n)\right]^*\right\}\\
&= \mathcal{E}\left\{\sum_{i=1}^{n_T}\sum_{j=1}^{n_T} H^{\mathrm{p}}(m,i)S(i,n)\left(H^{\mathrm{p}}(m,j)\right)^* S^*(j,n)\right\}\\
&= \sum_{i=1}^{n_T}\sum_{j=1}^{n_T}\mathcal{E}\left\{H^{\mathrm{p}}(m,i)\left(H^{\mathrm{p}}(m,j)\right)^*\right\}S(i,n)S^*(j,n)\\
&= \sum_{i=1}^{n_T}\sum_{j=1}^{n_T}\delta(i-j)S(i,n)S^*(j,n)\\
&= \sum_{i=1}^{n_T}|S(i,n)|^2 = 1,
\end{aligned}$$

and

$$\begin{aligned}
\mathcal{E}\left\{Re(H^{\mathrm{s}}(m,n))Im(H^{\mathrm{s}}(m,n))\right\} &= \mathcal{E}\left\{\left[\sum_{i=1}^{n_T} Re(H^{\mathrm{p}}(m,i))Re(S(i,n)) - Im(H^{\mathrm{p}}(m,i))Im(S(i,n))\right]\right.\\
&\quad\left.\left[\sum_{j=1}^{n_T} Re(H^{\mathrm{p}}(m,j))Im(S(j,n)) + Im(H^{\mathrm{p}}(m,j))Re(S(j,n))\right]\right\}\\
&= \frac{1}{2}\sum_{i=1}^{n_T}\left(Re(S(i,n))Im(S(i,n)) - Im(S(i,n))Re(S(i,n))\right)\\
&= 0.
\end{aligned}$$

Therefore, $H^{\mathrm{s}}(m,n)$ is CSCG with mean zero and variance of 1, i.e., $H^{\mathrm{s}}(m,n) \sim CN(0,1)$.

We now look into the cross correlation between the different elements of $\mathbf{H}^{\mathrm{s}}$:

$$\begin{aligned}
\mathcal{E}\left\{H^{\mathrm{s}}(m,n)\left(H^{\mathrm{s}}(k,l)\right)^*\right\} &= \mathcal{E}\left\{\left[\sum_{i=1}^{n_T} H^{\mathrm{p}}(m,i)S(i,n)\right]\left[\sum_{i=1}^{n_T} H^{\mathrm{p}}(k,i)S(i,l)\right]^*\right\} \\
&= \mathcal{E}\left\{\sum_{i=1}^{n_T}\sum_{j=1}^{n_T} H^{\mathrm{p}}(m,i)S(i,n)\left(H^{\mathrm{p}}(k,j)\right)^* S^*(j,l)\right\} \\
&= \sum_{i=1}^{n_T}\sum_{j=1}^{n_T} \mathcal{E}\left\{H^{\mathrm{p}}(m,i)\left(H^{\mathrm{p}}(k,j)\right)^*\right\} S(i,n)S^*(j,l) \\
&= \sum_{i=1}^{n_T}\sum_{j=1}^{n_T} \delta(m-k)\delta(i-j)S(i,n)S^*(j,l) \\
&= \delta(m-k)\sum_{i=1}^{n_T} S(i,n)S^*(i,l) \\
&= \delta(m-k)\delta(n-l),
\end{aligned}$$

i.e., elements in $\mathbf{H}^{\mathrm{s}}$ are independent.

Till now, we have proved that $\mathbf{H}^{\mathrm{s}}$ is $n_R \times n_S$ CSCG with i.i.d. elements of $CN(0,1)$, i.e., $\mathbf{H}^{\mathrm{s}}$ is *statistically* the same as the $n_S \times n_R$ DM channel matrix $\mathbf{H}^{\mathrm{d}}$. The eigenvalues of $\mathbf{H}^{\mathrm{s}}\left(\mathbf{H}^{\mathrm{s}}\right)^H$ and $\mathbf{H}^{\mathrm{d}}\left(\mathbf{H}^{\mathrm{d}}\right)^H$ are hence also *statistically* the same, following the distribution given in (2.14).

Therefore, the $n_T \times n_R \times n_S$ CDDSS channel has the same ergodic capacity as the $n_S \times n_R$ DM channel. For $n_S \le n_R \le n_T$, we we hence have

$$C_{E,CDDSS,n_T \times n_R \times n_S} \le C_{E,VBLAST,n_S \times n_R},$$

the equality is true when $n_S = n_T$.

Making use of (2.15), the ergodic capacity of $n_T \times n_R \times n_S$, $n_S \le n_R \le n_T$, CDDSS-MIMO CSCG channels can be evaluated as

$$C_{E,CDDSS} = \int_0^\infty \log_2\left(1 + \frac{P\lambda}{n_S N_0}\right)\sum_{k=0}^{n_S-1}\left[L_k^{n_R-n_S}(\lambda)\right]^2 \lambda^{n_R-n_S} e^{-\lambda} d\lambda. \tag{4.2}$$

### 4.1.2 Ergodic Capacity of GSTBC, QSTBC, and LDC Asymmetric MIMO-OFDM Channels

For GSTBC, QSTBC, and LDC precoded MIMO-OFDM channels, $\mathbf{H}$ is no longer i.i.d. CSCG. To obtain the ergodic capacity, we need to derive the statistical distribution of the eigenvalues in $\mathbf{H}\mathbf{H}^H$. This is, however, non-trivial. So in this thesis, we use Monte Carlo simulations to obtain their ergodic capacities.

### 4.1.3 Numerical Results

We now present the numerical results. FFT size of $N = 64$ is used, and all the 64 subcarriers are used to transmit data. For CDDSS, we set $n_S = n_R$.

In Fig. 4.1, we depict the ergodic capacity versus SNR of the various precoded $4 \times 2$ MIMO-OFDM channels. Eight i.i.d. zero mean complex Gaussian multipath channels are used, and perfect spatial uncorrelation is assumed. The $2 \times 2$ DM-OFDM capacity is also included for comparison. The LDC in [66] was used in generating the result. From the channel matrix relation in (3.23), it is straightforward that $4 \times 2$ GSTBC and QSTBC have the same eigenvalues for each channel realization and hence the same ergodic capacity. But more interestingly, the figure shows that LDC also has the same ergodic capacity as GSTBC and QSTBC. This can be implicitly explained by the fact that the three schemes have the same precoding rate, i.e., $R_{P,\mathrm{LDC}} = R_{P,\mathrm{GSTBC}} = R_{P,\mathrm{QSTBC}} = 2$, and that the dispersion matrices $\mathbf{A}_q$ and $\mathbf{B}_q$ satisfy the constraint of $\mathbf{A}_q^H \mathbf{A}_q = \mathbf{B}_q^H \mathbf{B}_q = \frac{T}{Q} \mathbf{I}_{n_T}$, hence dispersing the transmitted symbols with equal energy in all spatial and temporal directions. This is exactly the same as what the rate-2 QSTBC does.

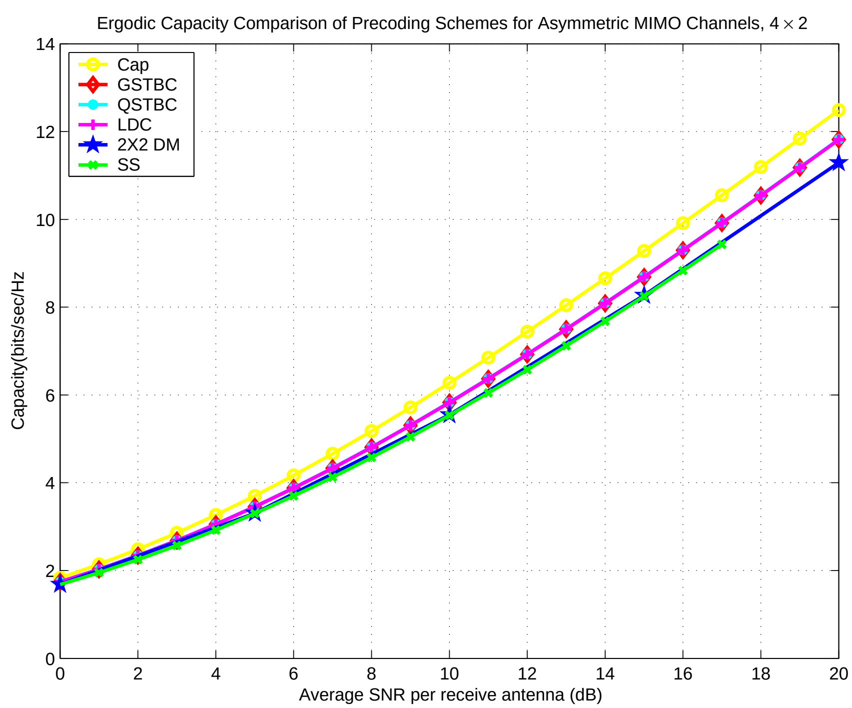


**Figure 4.1: Ergodic capacity comparison for a $4 \times 2$ system.**

Comparing the implementation complexity of GSTBC, QSTBC, and LDC, however, LDC will be less preferable due to its higher implementation complexity in both the encoding and decoding processes. In terms of design flexibility, both LDC and QSTBC lose to GSTBC due to the fact that with every additional pair of transmit antennas, we can add one more group of Alamouti STBC if the number of receive antennas satisfy $n_R \geq \frac{n_T}{2}$. The coding rate is always $\frac{n_T}{2}$. A rate-$\frac{n_T}{2}$ QSTBC and LDC, on the other hand, may have to be derived for different $n_T$'s.

From Fig. 4.1, we can also see that CDDSS has the same ergodic capacity as the $2 \times 2$ DM-OFDM channel, as proved in Section 4.1.1.

The ergodic capacity for $8 \times 4$ channels is depicted in Fig. 4.2. Only GSTBC, LDC, CDDSS, and $4 \times 4$ DM are considered in this case. Studying the results in the figure, we can draw the same conclusion as the $4 \times 2$ channel, that is, carefully designed precoding such as GSTBC and LDC can make use of the additional transmit antennas to introduce capacity improvement over $n_R \times n_R$ DM channel, and CDDSS has the same ergodic capacity as the $n_S \times n_R$ DM channels. All these precoding schemes, however, introduce capacity loss.

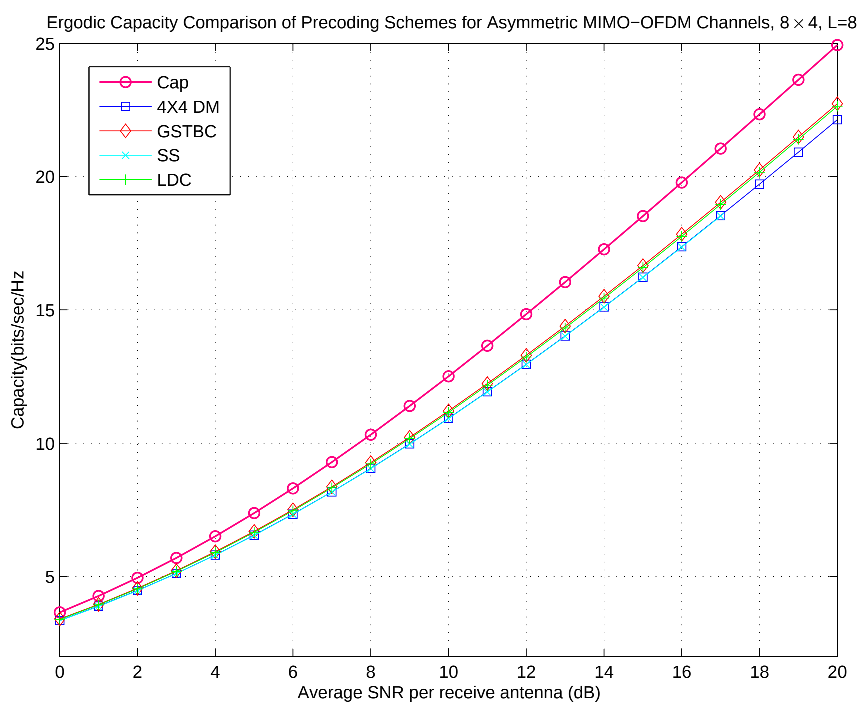


**Figure 4.2: Ergodic capacity comparison for a $8 \times 4$ system.**

## 4.2 Outage Capacity

Having compared the ergodic capacity, we now look at the outage capacity of the precoded MIMO-OFDM channels. As pointed out in [52], the outage properties are determined by the number of spatial and frequency diversity degrees of freedom in the channel. Assuming that a codeword spans only one OFDM symbol, the outage probability $P_{\text{out}}$ for a given rate $R$ is defined as

$$P_{\text{out}} = prob\left(I = \frac{1}{N}\sum_{k=0}^{N-1} I_k < R\right), \tag{4.3}$$

where $I$ is the mutual information of the MIMO-OFDM channel. The outage capacity can be obtained analytically if the statistical property of $I$ is known.

From Theorem 3 in Chapter 3, the MIMO channel $\mathbf{H}_k$ at each subcarrier $k$ is statistically the same. Hence, the mutual information of each subcarrier, $I_k$, is statistically the same. The mean mutual information of the MIMO-OFDM channel is thus written as

$$\mathcal{E}\left\{I\right\} = \mathcal{E}\left\{\frac{1}{N}\sum_{k=0}^{N-1} I_k\right\} = \frac{1}{N}\sum_{k=0}^{N-1}\mathcal{E}\left\{I_k\right\} = \mathcal{E}\left\{I_k\right\} = \bar{I}, \tag{4.4}$$

which is independent of the multipath channel characteristics. The correlation of $I_k$ at different subcarriers, however, is dependent on the frequency domain correlation of $\mathbf{H}_k$'s, which depends on the dispersiveness of the multipath channels. We now look at two extreme cases - the single-path flat fading channel, and the highly dispersive frequency independent fading. The later case is obtained with i.i.d. zero-mean complex Gaussian N-path channel and $L_{\text{CP}} = N$.

For the flat fading channel, each subcarrier has the same $\mathbf{H}_k$, hence

$$\begin{aligned} I_k &= I_{\text{flat}}, \qquad k = 0,\ 1,\ \cdots,\ N-1, \\ I &= I_{\text{flat}}, \end{aligned}$$

we therefore have

$$\begin{aligned} \bar{I} &= \bar{I}_{\text{flat}}, \\ Var(I) &= \mathcal{E}\left\{I^2\right\} - \left(\mathcal{E}\left\{I\right\}\right)^2 = \mathcal{E}\left\{I_{\text{flat}}^2\right\} - \bar{I}_{\text{flat}}^2 = Var(I_{\text{flat}}). \end{aligned}$$

For the highly dispersive frequency independent fading channel, the MIMO channels at different subcarriers are statistically the same but are independent of each other, i.e., i.i.d. The mutual information of each subcarrier is therefore also i.i.d. We hence have

$$
\begin{aligned}
\mathcal{E}\left\{I\right\} &= \frac{1}{N}\sum_{k=0}^{N-1}\mathcal{E}\left\{I_k\right\} = \frac{1}{N}\sum_{k=0}^{N-1}\bar{I}_k = \bar{I}_k, \\
Var(I) &= \mathcal{E}\left\{I^2\right\} - \left(\mathcal{E}\left\{I\right\}\right)^2 = \mathcal{E}\left\{\frac{1}{N}\sum_{k=0}^{N-1}I_k\,\frac{1}{N}\sum_{m=0}^{N-1}I_m\right\} - \left(\frac{1}{N}\sum_{k=0}^{N-1}\bar{I}_k\right)^2 \\
&= \frac{1}{N^2}\left[\sum_{k=0}^{N-1}\sum_{m=0}^{N-1}\left(\mathcal{E}\left\{I_k I_m\right\} - \bar{I}_k\bar{I}_m\right)\right] \\
&= \frac{1}{N^2}\left[\sum_{k=0}^{N-1}\left(\mathcal{E}\left\{I_k^2\right\} - \bar{I}_k^2\right)\right] \\
&= \frac{1}{N^2}\left(\sum_{k=0}^{N-1}Var(I_k)\right) \\
&= \frac{1}{N}Var(I_k)
\end{aligned}
$$

From the above two extreme cases, we can make a qualitative conclusion that the more dispersive the multipath channel, the smaller the variance of the mutual information, hence the lower the outage probability and the higher the outage capacity.

We next look at some numerical results with i.i.d. zero mean complex Gaussian multipath channels. Same as the ergodic capacity study, 64 subcarriers are used in the OFDM channels, and $L_{\mathrm{CP}}$ is assumed to be long enough to achieve perfect ISI mitigation.

### 4.2.1 Numerical Results for Frequency-Domain Correlated Channels

We now present the numerical results in frequency-domain correlated channels. Again spatial independence is assumed among the channel coefficients for all the subcarriers. We first verify our qualitative analysis that frequency correlation introduces degradation to the outage capacity. Presented in Fig. 4.3 is the outage probability of $4\times4$ DM-OFDM channels with the number of i.i.d. multipaths of $L=3$, $L=8$, $L=16$, and $L=64$, respectively. The SNR is set to 10dB. It clearly shows that the richer the multipaths in the channel, the higher the outage capacity. Comparing the outage capacity of the four channels at $P_{\mathrm{out}}=0.01$, we can see that from the most correlated channel of $L=3$ to the fully uncorrelated channel of $L = 64$, the outage capacity can increase by over 1 bit/channel use!

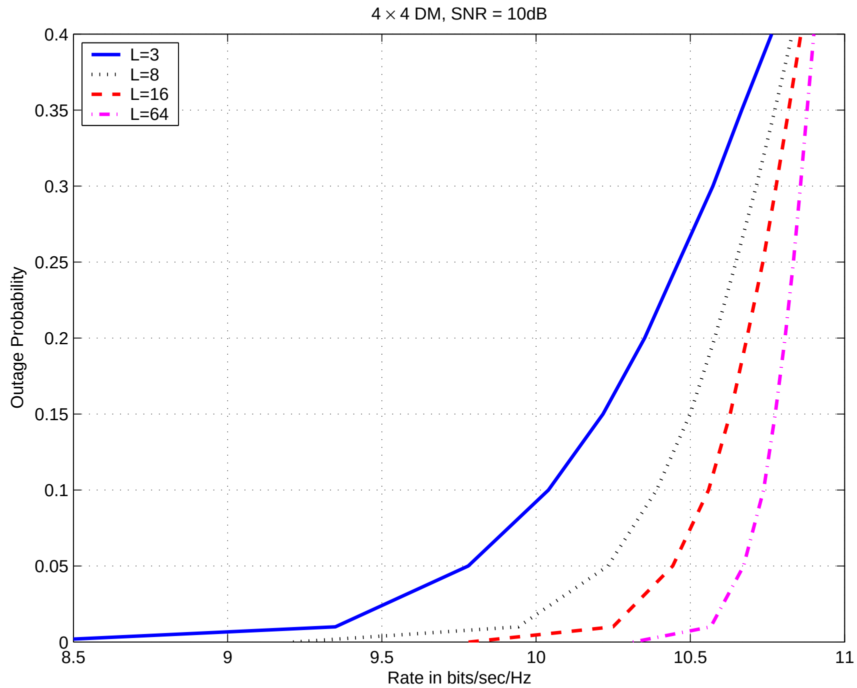


**Figure 4.3: Outage Capacity of $4 \times 4$ Direct Mapping MIMO-OFDM. SNR = 10 dB.**

Same observation can be made for the DM-OFDM outage capacity performance in asymmetric MIMO channels, as shown in Fig. 4.4 for the $4 \times 2$ setup at SNR of 10dB. Three multipath channels are used for the comparison, namely, the i.i.d. multipath channel with $L = 3$, $L = 8$, and $L = 16$, respectively. At $P_{\text{out}} = 0.01$, the outage capacity increases from 5.05 bits/channel use of the $L = 3$ channel, to 5.50 bits/channel use of the $L = 8$ channel, and 5.7 bits/channel use of the $L = 16$ channel.

We next look at the outage capacity of pre-coded asymmetric MIMO-OFDM channels. Depicted in Fig. 4.5 are the outage probability versus mutual information curves for the $4 \times 2$ GSTBC at SNR = 10dB. Again three multipath channels are used for the comparison, i.e., $L = 3$, $L = 8$, and $L = 16$. Similarly, the richer the multipath components, the higher the outage capacity. For example, at $P_{\text{out}} = 0.01$, the outage capacity increases from 4.65 bits/channel use of the $L = 3$ channel, to 5.08 bits/channel use of the $L = 8$ channel, and 5.3 bits/channel use of the $L = 16$ channel.

We summarize the outage capacity performance of all the precoded $4 \times 2$ MIMO-OFDM channels in Fig. 4.6 for $L = 8$ i.i.d. zero mean complex Gaussian multipath channels. Two incremental cyclic delay

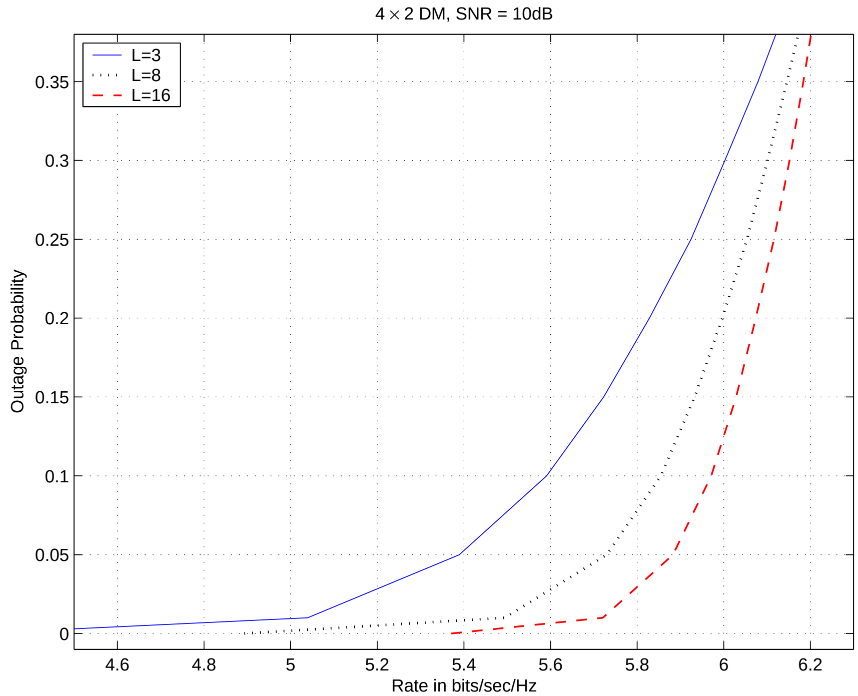


**Figure 4.4: Outage Capacity of $4 \times 2$ Direct Mapping MIMO-OFDM. SNR = 10 dB.**

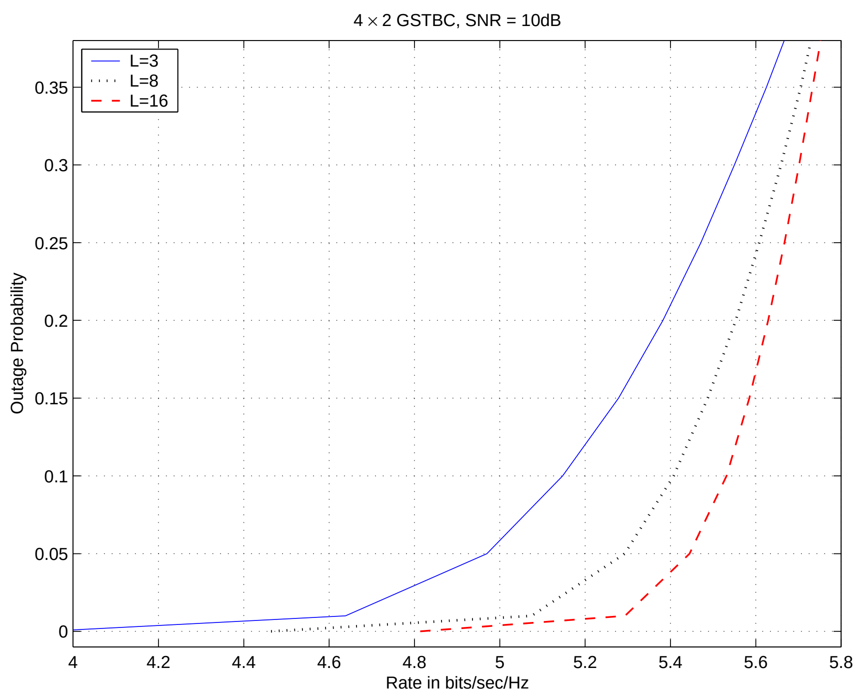


**Figure 4.5: Outage Capacity of $4 \times 2$ GSTBC MIMO-OFDM. SNR = 10 dB.**

values are used in CDDSS, i.e., $\tau = 3$ and $\tau = 8$. It can seen that different from the ergodic capacity case, CDDSS has higher outage capacity than the $2 \times 2$ DM channels, which is attributed to the additional frequency diversity introduced by CDD. As different cyclic delay values introduce different additional frequency diversity, they also result in different outage capacity. The figure also shows that GSTBC and LDC have almost the same outage capacity.

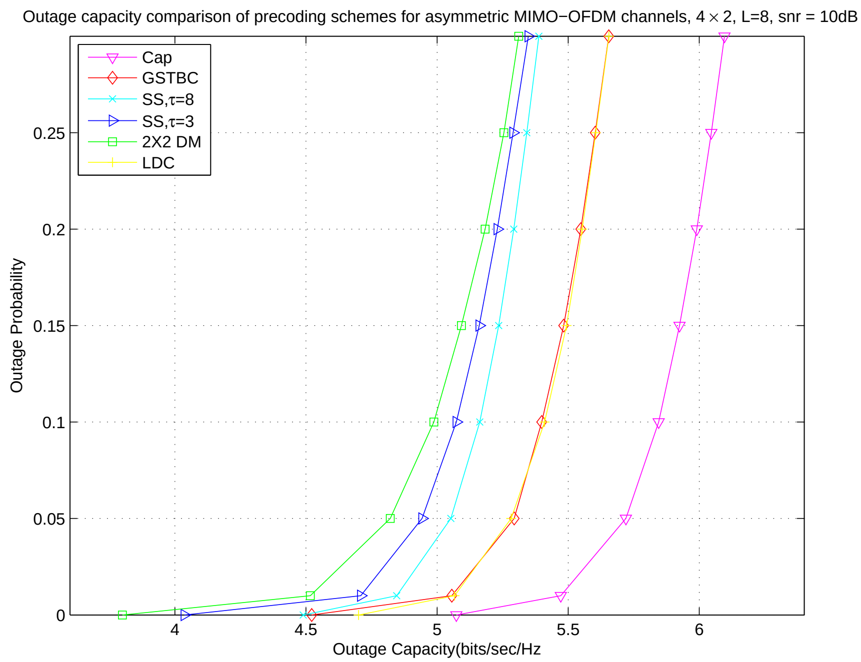


**Figure 4.6: Outage Capacity of $4 \times 2$ Precoded MIMO-OFDM. $L = 8$.**

Fig. 4.8 summarizes the outage capacity of the precoded $8 \times 4$ CDDSS MIMO-OFDM channels with $L = 8$ i.i.d. zero-mean complex Gaussian multipaths, and $\tau = 1,\ 3,\ 5$ and $\tau = 8$. It again proves that different cyclic delay values result in different outage capacity.

Fig. 4.8 summarizes the outage capacity of the precoded $8 \times 4$ MIMO-OFDM channels with $L = 16$ i.i.d. zero-mean complex Gaussian multipaths, and $P_{\text{out}} = 1\%$. $\tau = 16$ is used for CDDSS channels. From this figure, we can see that for GSTBC, the required SNR is 1.52dB to achieve 4 bits/channel use, 7.07dB to achieve 8 bits/channel use; for CDDSS, the required SNR is respectively 1.66dB and 7.29dB to achieve 4 bits/channel use and 8 bits/channel use. Therefore, the outage capacity difference is getting much smaller between GSTBC and CDDSS.

In Fig. 4.9, we compare the influence of power delay profiles (PDF) on the outage behavior. $8 \times 4$

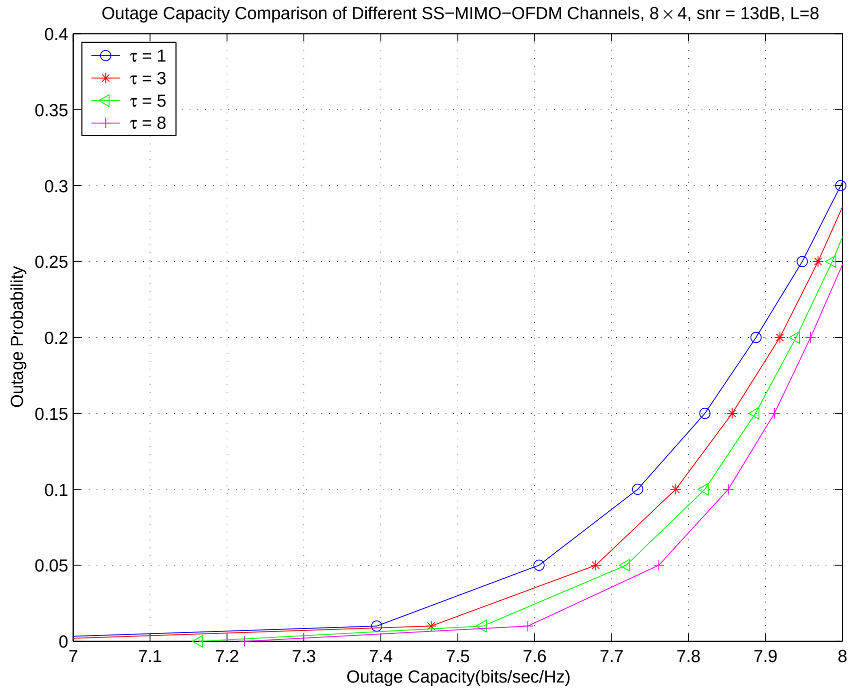


**Figure 4.7: Outage Capacity versus SNR of $8 \times 4$ CDDSS MIMO-OFDM. $L = 8$, $\tau = 1,\ 3,\ 5$ and $\tau = 8$. Uniform power delay profiles.**

GSTBC MIMO-OFDM channels with $L = 16$ are used. Two PDF's are considered, namely, the uniform power delay profile (UPDP) with i.i.d. zero mean complex Gaussian multipaths, and exponential power delay profile (EPDP) with zero mean complex Gaussian multipaths and incremental power loss of 3dB per multipath component. Obviously, UPDP leads to better outage performance.

From the qualitative analysis and the numerical results presented in this section, we have shown that for MIMO-OFDM, the richer the diversity in the precoded channel, the higher the outage capacity. Therefore, the precoding scheme should be designed to exploit all the frequency and space diversities in the channel.

## 4.3 The Mutual Information With Fixed-Order Modulation

Ergodic capacity is obtained when the channel input $\mathbf{X}$ is i.i.d. Gaussian. For practical systems, symbols with fixed constellation, e.g., M-PSK or M-QAM, have to be transmitted. Therefore, the ergodic capacity in Section 4.1 is not achievable, and a more realistic indication of the precoding optimality will be the

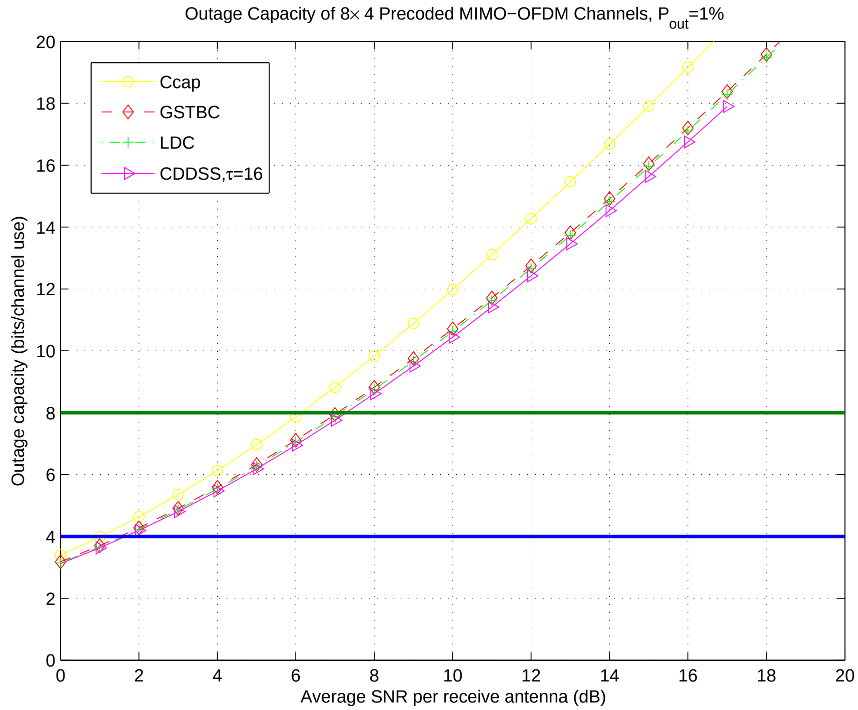


**Figure 4.8: Outage Capacity versus SNR of $8 \times 4$ Precoded MIMO-OFDM at $P_{\text{out}} = 1\%$. $L = 16$, Uniform power delay profiles.**

mutual information between the channel output $\mathbf{R}$ and the channel input $\mathbf{X}$, assuming that the elements in $\mathbf{X}$ are i.i.d. from the modulation constellation. The mutual information is computed as [26]

$$I(\mathbf{X};\ \mathbf{R}) = \left[H(\mathbf{R}) - H(\mathbf{R}|\mathbf{X})\right]/T, \tag{4.5}$$

where $H(\cdot) = -\mathcal{E}\log p(\cdot)$ is the entropy function, and $T$ is the precoding interval. From the generalized channel model in (3.14), we have

$$p(\mathbf{R}|\mathbf{X}) = \frac{1}{(2\pi\sigma)^{\mathrm{d_m}}}\exp\left(-\frac{|\mathbf{R}-\mathbf{H}\mathbf{X}|^2}{2\sigma^2}\right),$$

where '$\mathrm{d_m}$' denotes the dimension of the complex AWGN in (3.14). Hence

$$H(\mathbf{R}|\mathbf{X}) = \mathrm{d_m}\log 2\pi\sigma^2 e.$$

Calculation of $H(\mathbf{R})$ needs to take expectation over three random variables, i.e., $\mathbf{H}$, $\mathbf{X}$, and $\mathbf{V}$. Here we obtain the numerical results through Monte Carlo simulations.

We depict the mutual information of a $4 \times 2$ precoded channel with QPSK modulation in Fig. 4.10, and 16QAM modulation in Fig. 4.11. From the two figures, we see that same as the ergodic capacity,

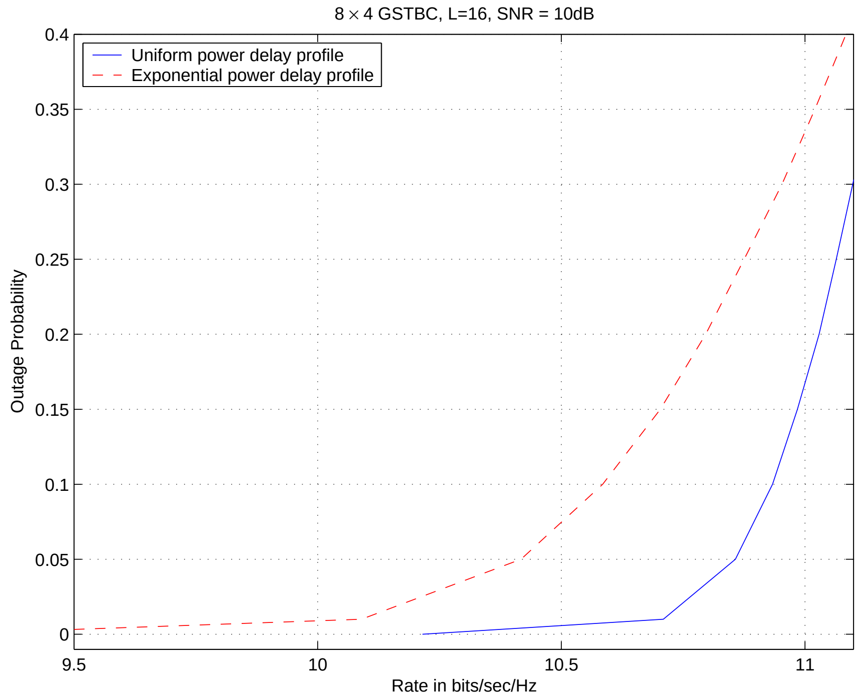


**Figure 4.9: Outage Capacity of** $8 \times 4$ **GSTBC MIMO-OFDM.** $L = 16$**, Uniform and exponential power delay profiles, SNR = 10dB.**

CDDSS has the same mutual information as the $2 \times 2$ DM for both the QPSK and 16QAM modulation channel inputs, and LDC, GSTBC, and QSTBC have the same mutual information.

## 4.4 The Diversity Gain

As shown in Section 4.2, the richer the diversity in the precoded channels, the higher the outage capacity that can be achieved. The diversity potential should therefore be fully exploited. In this section, we will study the diversity performance of the precoding schemes based on two assumptions:

**Maximum receive antenna diversity of** $n_R$**.** This can be achieved when the interference from the other spatial streams is completely cancelled by advanced receivers, e.g., the MLD receivers, the iterative turbo receivers to be discussed in Chapter 5, the BI-GDFE in [79][80][81], etc..

**Full exploitation of frequency diversity.** This can be achieved when a very powerful FEC is deployed.

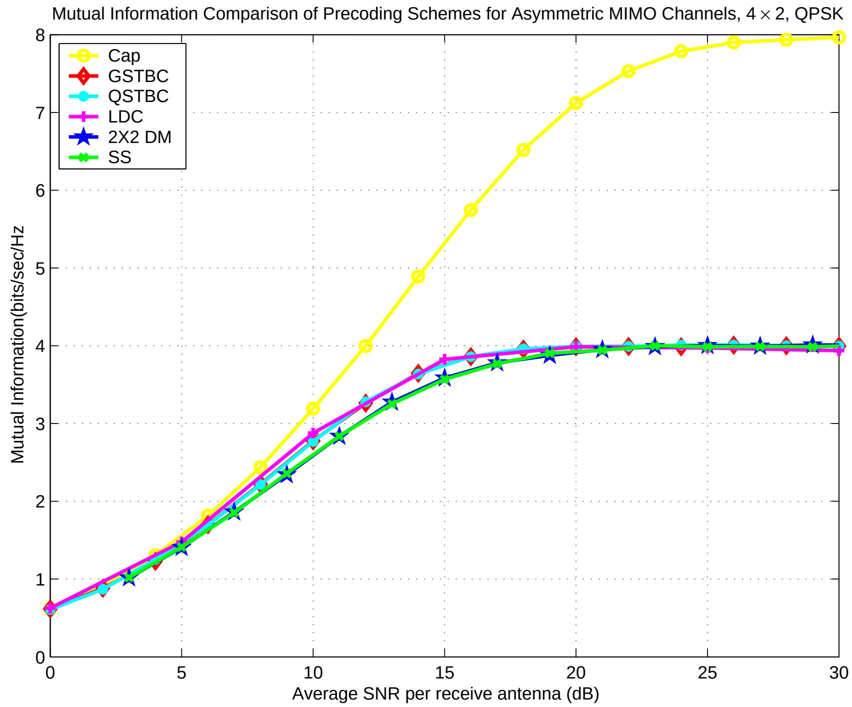


**Figure 4.10: Mutual information comparison for a $4 \times 2$ system, QPSK.**

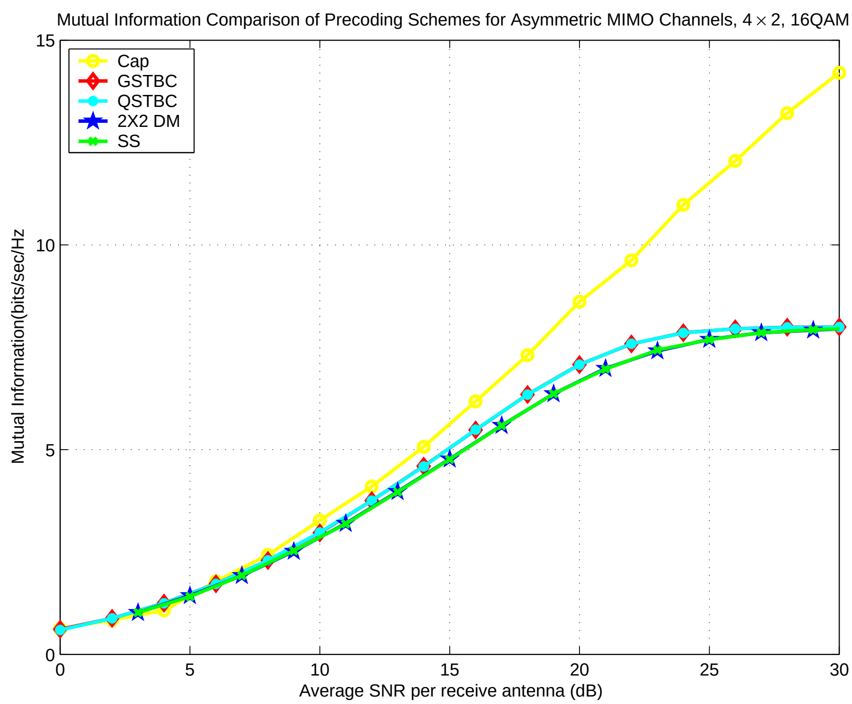


**Figure 4.11: Mutual information comparison for a $4 \times 2$ system, 16QAM.**

With the two assumptions, it has been proven in [44] that the maximum diversity order of GSTBC is $2n_R$, achieved with MLD or with perfect interference cancellation among the different groups. Same diversity order is achieved by the $n_T = 4$ QSTBC, as proven in [41]. In fact, [41] has shown that for spatially uncorrelated channels, QSTBC and GSTBC have the same BER performance.

Similarly, we can have $n_R$ the maximum space-diversity order of CDDSS and DM. However, the maximum frequency diversity order of DM is $L$, the number of multipaths in the channel, whereas additional frequency diversity is made available in CDDSS whose order is dependent on the delay value $\tau$, the power delay profile of the fading channel, and the total number of transmit antennas [63]. The maximum achievable frequency domain diversity is

$$\min(n_T L, N).$$

As for LDC, although the BER union bound has been derived in [66], no explicit diversity order can be obtained from it. We therefore depict in Fig. 4.12 the MLD performance for the 16QAM-modulated $4 \times 2$ precoded system. No FEC coding is applied, so only spatial diversity contributes to the BER versus SNR slope. Naturally, SS has the lowest spatial diversity when compared with LDC and GSTBC. Between LDC and GSTBC, we can see from the figure that LDC has slightly higher gain than GSTBC.

The frequency diversity gain effect of CDDSS is illustrated in Fig. 4.13 for an $8 \times 4$ convolutionally coded (CC) CDDSS-OFDM system with 16QAM modulation. The CC has rate $R_c = \frac{1}{2}$, and minimum free distance of $d_{\text{free}} = 5$. $L$-path UPDF is deployed in each MIMO multipath channel, and $\tau$ is the CDD value. At the receiver, turbo processing is employed. From the figure, it can be seen that for low-order multipath channel($L = 3$, $L = 8$), larger $\tau$ will lead to higher diversity gain. For high-order multipath channel, e.g., $L = 16$, the different CDD values do not have much impact on the BER performance. This is because the frequency diversity is realized by the soft decision decoding of the CC whose performance is limited by its $d_{\text{free}}$. To maximize the frequency diversity gain of CDDSS-OFDM, either a stronger code or a linear frequency domain transform [82] has to be used. More details will be given in Section 4.6.

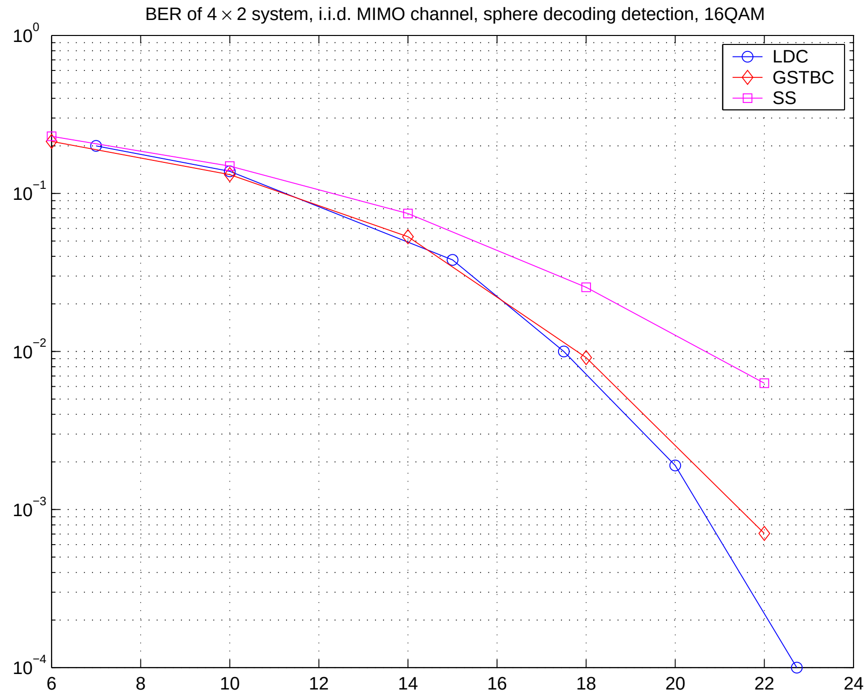


**Figure 4.12: BER performance of the different precoding schemes for $4 \times 2$ channels, ML detection, 16QAM.**

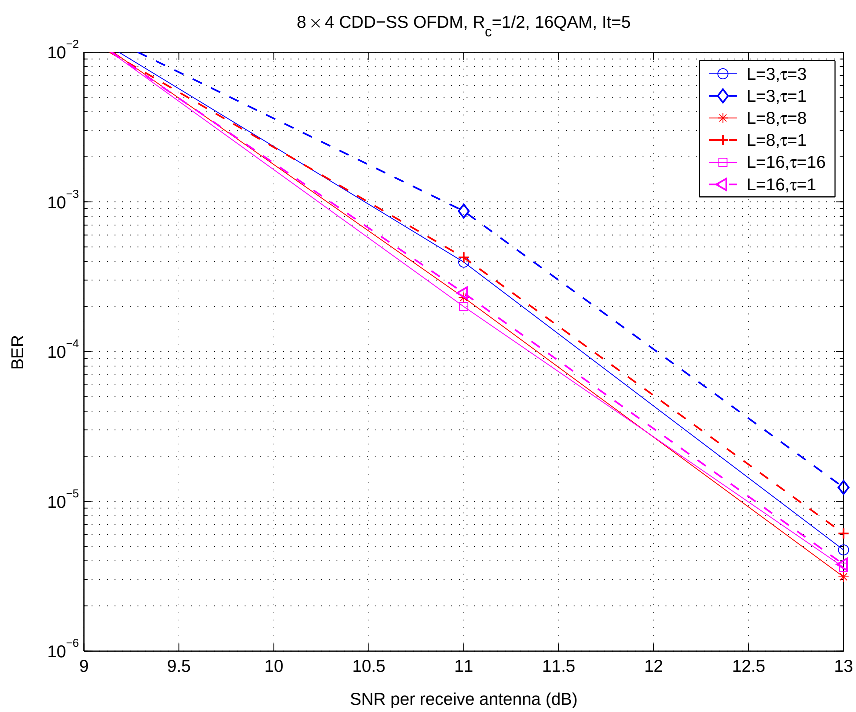


**Figure 4.13: BER performance of the $8 \times 4$ CDD-CDDSS MIMO-OFDM with different channel order and delay values. $R_c = \frac{1}{2}$, $d_{\text{free}} = 5$ CC, turbo receiver, 16QAM.**

## 4.5 Bit Error Rate

The BER performance is related to the diversity $d$ that the system can exploit as:

$$\mathrm{BER} \propto \mathrm{SNR}^{d},$$

where $d$ is determined by the precoding scheme deployed at the transmitter, the richness of multipath in the channel, and optimality of the receiver algorithms. In this section, we look into the BER performance of the precoded $8 \times 4$ systems with the $R_c = \frac{1}{2}$ $d_{\mathrm{free}} = 5$ CC. For receiver, we use a five-iteration turbo receiver whose details are given in Chapter 5. 16-tap multipath channel with UPDF is used in the simulations. For CDDSS, we set $\tau = 16$.

Depicted in Fig. 4.14 is the BER for QPSK, and Fig. 4.15 for 16QAM modulated systems. From the figures, it can be seen that GSTBC has the best BER performance among the three precoding schemes. Compared with LDC, it is about 0.25dB better with QPSK, and 0.5dB better with 16QAM, at BER = $10^{-5}$. This should be due to the same turbo receiver structure being used for all the three precoded systems. As recommended by [83], a widely linear filter can be used for LDC systems to make use of both the original signal and its conjugate in the turbo receiver detection. Then the performance advantage of LDC will be better realized. But this also means additional implementation complexity.

Although CDDSS still loses to LDC and GSTBC, the performance gap, however, becomes smaller than the capacity performance and the uncoded performance. This is due to the additional frequency diversity from CDDSS. For QPSK signals, CDDSS loses to GSTBC by only about 0.5dB at BER= $10^{-5}$, and for 16QAM, 0.7dB. However, it has to be pointed out that the detection complexity of CDDSS is lower than GSTBC, mainly due to its smaller $\mathbf{H}$ dimension than the GSTBC systems.

Presented in Fig. 4.16 and Fig. 4.17 are the FER performance of the $8 \times 4$ $R_c = \frac{1}{2}$ $K = 3$ convolutionally coded MIMO-OFDM systems with QPSK and 16QAM modulation signals, respectively. The frame length is set to one OFDM symbol with 64 subcarriers. $L = 16$ i.i.d. complex Gaussian multipath channels are used, and $\tau = 16$ for CDDSS. The $P_{\mathrm{out}} = 1\%$ outage capacity is also included in the figure for comparison. For QPSK modulated signals corresponding to 4 bits/channel use spectral efficiency, we can see from Fig. 4.16 that the GSTBC is 7.34 dB away from the outage capacity, and

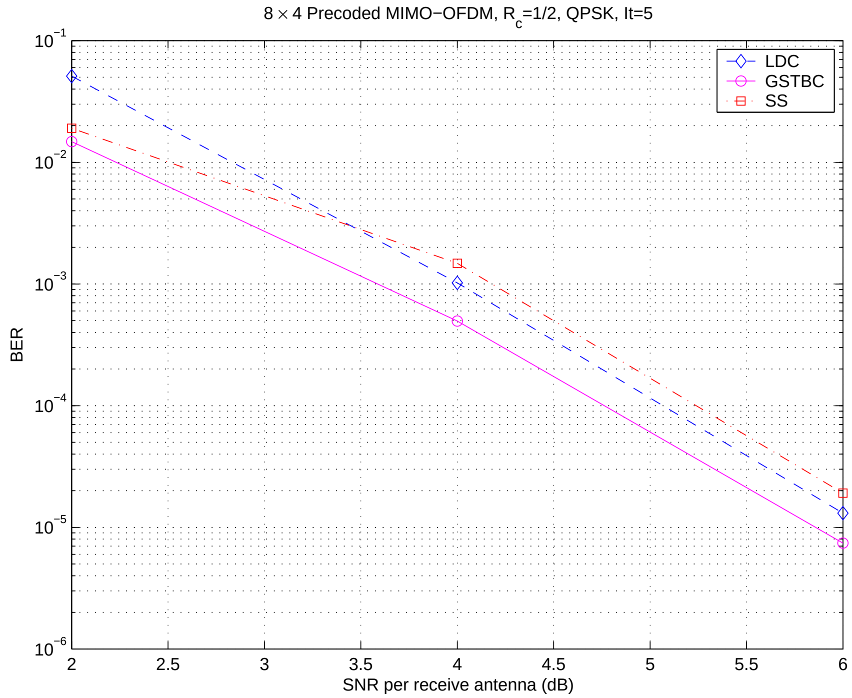


**Figure 4.14: BER performance of the different precoding schemes for $8 \times 4$ MIMO-OFDM channels. QPSK.**

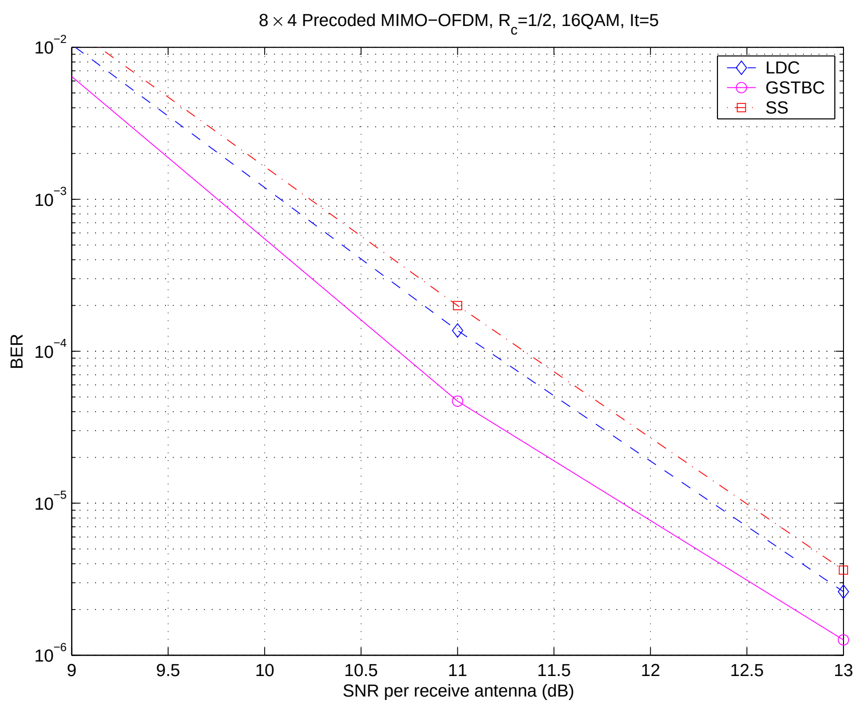


**Figure 4.15: BER performance of the different precoding schemes for $8 \times 4$ MIMO-OFDM channels. 16QAM.**

CDDSS is about 7.30dB away from its outage capacity. For 16QAM modulated signals corresponding to 8 bits/channel use spectral efficiency, Fig. 4.17 shows that GSTBC and CDDSS are respectively 5.2dB and 5.1dB away from their outage capacities. Therefore, CDDSS is slightly better than GSTBC from the view point of approaching the outage capacity limit.

Lastly, there is one point that needs to be pointed out. Results in Fig. 4.16 and Fig. 4.17 are generated with a very simple CC, which is selected to just show optimality of the precoding scheme. To approach the capacity limit, a more powerful code, e.g., turbo code or LDPC code, should be used.

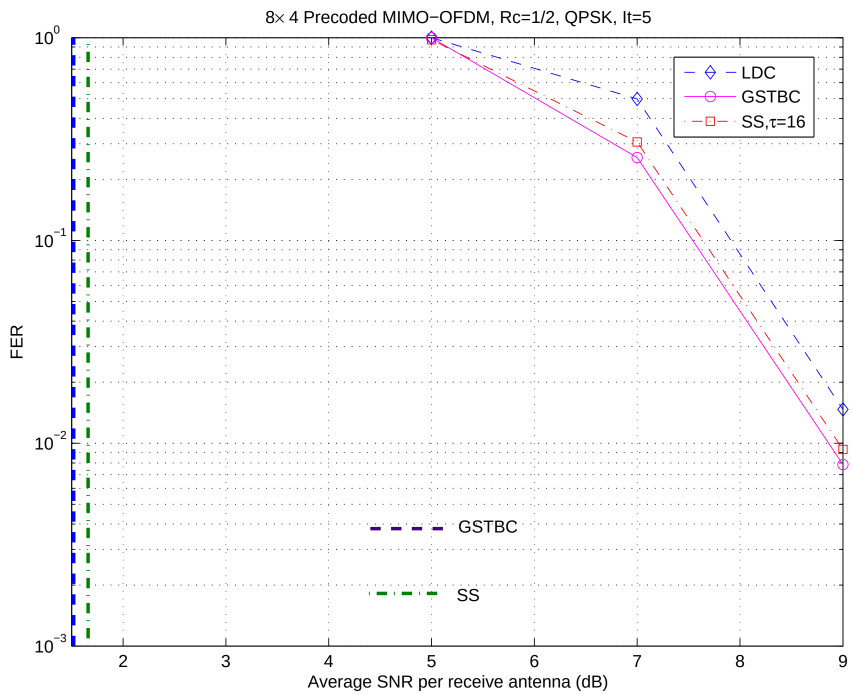


**Figure 4.16: FER performance of the different precoding schemes for $8 \times 4$ MIMO-OFDM channels. QPSK.**

## 4.6 Two-dimensional Linear Pre-transformed MIMO-OFDM

We have shown that CDDSS is capacity lossless when the number of spatial streams is equal the number of transmit antennas. Otherwise, CDDSS and the other precoding schemes we have studied in the previous sections are all capacity lossy. We have also proved that CDD-OFDM and CDDSS-OFDM transfer the transmit diversity from the "extra" antennas to frequency diversity, and frequency diversity can be realized by the FEC code. However, this also means that the realizable diversity order is limited by the free distance

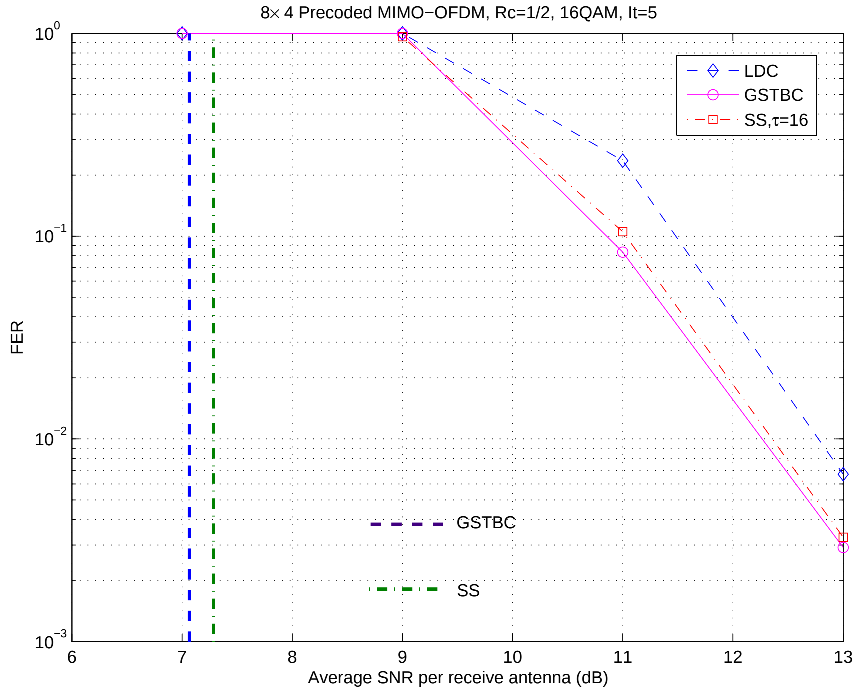


**Figure 4.17: FER performance of the different precoding schemes for $8 \times 4$ MIMO-OFDM channels. 16QAM.**

of the FEC code. A precoding scheme which is able to achieve full capacity and full diversity is therefore desired.

Motivated by the works in [84] and [82] for single-transmit single-receive OFDM systems, here we propose a two-dimensional linear pre-transformed MIMO-OFDM system, i.e., 2DLPT MIMO-OFDM, as depicted in Fig. 4.18. Linear pre-transforms (LPT) are applied in both frequency and spatial domains, as shown in the figure. The $N \times N$ linear unitary frequency domain LPT (FD-LPT) is applied to each individual spatial streams independently before the spatial domain LPT (SD-LPT). The SD-LPT is applied to the FD-LPT output, subcarrier by subcarrier, independently. When the number of spatial streams $n_S$ is equal to the number of transmit antennas $n_T$, i.e., $n_S = n_T$, the SD-LPT is $n_T \times n_T$ unitary transform. When $n_S < n_T$, the SD-LPT is orthonormal matrix. We will show that for the case of $n_S = n_T$, the 2DLPT MIMO-OFDM can achieve full capacity and full diversity simultaneously; when $n_S < n_T$, 2DLPT is capacity lossy but achieves full diversity.

Following (3.10), the 2D-PT MIMO-OFDM system can still be modeled as

$$\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{T}}\boldsymbol{\mathcal{X}} + \boldsymbol{\mathcal{V}}, \tag{4.6}$$

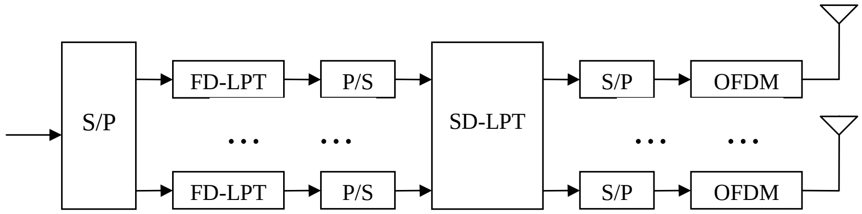


**Figure 4.18: Transmitter block diagram of 2DLPT MIMO-OFDM.**

where $\boldsymbol{\mathcal{Y}}$, $\boldsymbol{\mathcal{X}}$, $\boldsymbol{\mathcal{H}}$, and $\boldsymbol{\mathcal{V}}$ are defined the same as in (3.10), and $\mathcal{T}$ denotes the 2DLPT which is written as

$$\begin{aligned}
\mathcal{T} &= \mathbf{T}^S\mathbf{T}^F && (4.7)\\
&= \mathbf{Q}\,\mathrm{diag}\left(\mathbf{T}_1^S,\ \mathbf{T}_2^S,\ \cdots,\mathbf{T}_N^S\right)\mathbf{P}\mathrm{diag}\left(\mathbf{T}_1^F,\ \mathbf{T}_2^F,\ \cdots,\mathbf{T}_{n_T}^F\right) && (4.8)\\
&= \mathbf{Q}\begin{bmatrix} \mathbf{T}_1^S & 0 & \cdots & 0\\ 0 & \mathbf{T}_2^S & \cdots & 0\\ \vdots & \vdots & \vdots & 0\\ 0 & \cdots & 0 & \mathbf{T}_N^S \end{bmatrix}\mathbf{P}\begin{bmatrix} \mathbf{T}_1^F & 0 & \cdots & 0\\ 0 & \mathbf{T}_2^F & \cdots & 0\\ \vdots & \vdots & \vdots & 0\\ 0 & \cdots & 0 & \mathbf{T}_{n_S}^F \end{bmatrix} && (4.9)
\end{aligned}$$

where $\mathbf{Q}$ and $\mathbf{P}$ are respectively row and column permutation matrices, with their elements defined as

$$Q_{i,j} = \begin{cases} 1, & \text{when } i = \lfloor \frac{j}{n_T} \rfloor + (j \mod n_T) \times N, \\ 0, & \text{otherwise} \end{cases} \quad (4.10)$$

$$P_{i,j} = \begin{cases} 1, & \text{when } j = \lfloor \frac{i}{n_T} \rfloor + (i \mod n_T) \times N, \\ 0, & \text{otherwise} \end{cases} \quad (4.11)$$

for $i,\ j = 0,\ 1,\cdots,Nn_T - 1$.

$\mathbf{T}_k^S$ denotes the $n_T \times n_S$ SD-LPT matrix at subcarrier $k$, $\mathbf{T}_{n_s}^S$ denotes the FD-LPT for spatial stream $n_s$ with the total number of spatial streams defined as $n_S$.

**SD-LPT** The CDDSS precoding matrices can be used for SD-LPT. From Theorem 5, the maximum achievable frequency diversity for each spatial stream is $\min(n_T L,\ N)$, when the incremental cyclic delay $\tau$ is set to $\tau = L$.

**FD-LPT** The linear transform proposed in [82] is used as FD-LPT $\mathbf{T}^F$ to exploit the frequency diversity.

### 4.6.1 Ergodic Capacity

As we have proved in Section 4.1.1, the CDDSS MIMO-OFDM has the same ergodic capacity as the $n_S \times n_T$ DM MIMO-OFDM channels. For 2DLPT MIMO-OFDM, the FD-LPT matrix $\mathbf{T}^F$ is unitary, hence $\mathbf{T}^F \boldsymbol{\mathcal{X}}$ remains i.i.d. Gaussian if the input signal $\boldsymbol{\mathcal{X}}$ is i.i.d. Gaussian. Therefore, 2DLPT MIMO-OFDM achieves the same ergodic capacity as the CDDSS MIMO-OFDM, i.e., when $n_S < n_T$, 2DLPT has the same ergodic capacity as the $n_S \times n_T$ DM MIMO-OFDM channels, and when $n_S = n_T$, 2DLPT has the same ergodic capacity as the $n_T \times n_T$ DM MIMO-OFDM channels.

Below we provide a direct proof that when $n_S = n_T$, 2DLPT is capacity lossless, by making use of the signal model in (4.6), and the fact that $\mathbf{T}^S$ is unitary:

$$\begin{aligned} C &= \mathcal{E}\left\{\log_2\left[\det\left(\mathbf{I} + \frac{\rho}{n_T}(\boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{T}})(\boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{T}})^H\right)\right]\right\} \\ &= \mathcal{E}\left\{\log_2\left[\det\left(\mathbf{I} + \frac{\rho}{n_T}\boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{T}}\boldsymbol{\mathcal{T}}^H\boldsymbol{\mathcal{H}}^H\right)\right]\right\} \\ &= \mathcal{E}\left\{\log_2\left[\det\left(\mathbf{I} + \frac{\rho}{n_T}\boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{H}}^H\right)\right]\right\}, \end{aligned}$$

i.e., 2DLPT MIMO-OFDM channel has the same ergodic capacity as the SDM MIMO-OFDM channel defined by $\boldsymbol{\mathcal{H}}$. Here $\rho$ is defined as the average signal to noise ratio (SNR) per receive antenna.

### 4.6.2 Diversity

**Theorem 6.** *The maximum diversity order of the 2DLPT MIMO-OFDM system is* $n_R \min(n_L n_T, N)$.

*Proof:* From (4.6), assuming perfect CSI at the receiver, we have the PEP based on MLD as

$$\begin{aligned} P(\mathbf{X} \to \mathbf{X}_e) &= P\left(p(\boldsymbol{\mathcal{Y}}|\mathbf{X}, \boldsymbol{\mathcal{H}}) < p(\boldsymbol{\mathcal{Y}}|\mathbf{X}_e, \boldsymbol{\mathcal{H}})\right) \\ &= P\left(|\boldsymbol{\mathcal{Y}} - \boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{T}}\mathbf{X}|^2 > |\boldsymbol{\mathcal{Y}} - \boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{T}}\mathbf{X}_e|^2\right) \\ &= Q\left(\sqrt{\frac{d^2(\boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{T}}\mathbf{X}, \boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{T}}\mathbf{X}_e)}{2N_0}}\right) \\ &\leq \exp\left(-\frac{\|\boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{T}}\mathbf{X} - \boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{T}}\mathbf{X}_e\|^2}{4N_0}\right), \end{aligned}$$

where $\mathbf{X}$ is the transmitted sequence, and $\mathbf{X}_e$ is the erroneously detected sequence.

Define

$$\mathbf{e} = \mathbf{T}^F (\mathbf{X} - \mathbf{X}_e) = \boldsymbol{\Upsilon}\boldsymbol{\Lambda}_e^{\frac{1}{2}}\boldsymbol{\Omega}^H,$$

$$\mathbf{H}_{n_r}^S = \boldsymbol{\mathcal{H}}_{n_r}\mathbf{T}^S = \boldsymbol{\Psi}\boldsymbol{\Gamma}^{\frac{1}{2}}\boldsymbol{\Xi}^H,$$

where $\boldsymbol{\mathcal{H}}_{n_r}$ denotes the submatrix of $\boldsymbol{\mathcal{H}}$ corresponding to the $n_r$th receive antenna, i.e., $\boldsymbol{\mathcal{H}}_{n_r} = \boldsymbol{\mathcal{H}}((n_r - 1)N + 1 : n_r N, :)$, and $\boldsymbol{\Lambda}_e^{\frac{1}{2}}$ and $\boldsymbol{\Gamma}^{\frac{1}{2}}$ are respectively the singular value matrices of $\mathbf{e}$ and $\mathbf{H}_{n_r}^S$.

The number of non-zero singular values in $\boldsymbol{\Lambda}_e^{\frac{1}{2}}$ is determined by the distance property of $\mathbf{T}^F$, $d_{\mathbf{T}^F}$. As $\mathbf{T}^F$ is $N \times N$ unitary, we have $d_{\mathbf{T}^F} = N$, i.e., the number of non-zero singular values in $\boldsymbol{\Lambda}_e^{\frac{1}{2}}$ is $N$. For $\boldsymbol{\Gamma}^{\frac{1}{2}}$, the number of non-zero elements is $\min(n_T L, N)$, from Theorem 4.

We therefore have

$$\mathbf{H}_{n_r}^S \mathbf{e} = \boldsymbol{\Psi}\boldsymbol{\Gamma}^{\frac{1}{2}}\boldsymbol{\Xi}^H\boldsymbol{\Upsilon}\boldsymbol{\Lambda}_e^{\frac{1}{2}}\boldsymbol{\Omega}^H = \boldsymbol{\Phi}\boldsymbol{\Delta}^{\frac{1}{2}}\boldsymbol{\Theta}^H,$$

and $\boldsymbol{\Delta}^{\frac{1}{2}}$ has $d_o = \min(n_T L, N)$ non-zero singular values.

Hence

$$\begin{aligned}
\|\boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{T}}\mathbf{X} - \boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{T}}\mathbf{X}_e\|^2 &= \sum_{n_r=1}^{n_R} \|\mathbf{H}_{n_r}^S \mathbf{e}\|^2 \\
&= \sum_{n_r=1}^{n_R} \mathrm{trace}\left(\mathbf{H}_{n_r}^S \mathbf{e}\mathbf{e}^H \left(\mathbf{H}_{n_r}^S\right)^H\right) \\
&= \sum_{n_r=1}^{n_R} \mathrm{trace}\left(\boldsymbol{\Phi}\boldsymbol{\Delta}\boldsymbol{\Phi}^H\right) = \sum_{n_r=1}^{n_R} \mathrm{trace}\left(\boldsymbol{\Delta}\right) \\
&= \sum_{n_r=1}^{n_R}\sum_{m=1}^{d_o} \delta_{m,n_r},
\end{aligned}$$

and the PEP as

$$P(\mathbf{X} \to \mathbf{X}_e) = \prod_{n_r=1}^{n_R}\prod_{m=1}^{d_o} \exp\left(-\frac{\delta_{m,n_r}}{4N_0}\right),$$

When the multipath components are zero-mean i.i.d. complex Gaussian, we have the average PEP written as [12]

$$P_e(\mathbf{X} \to \mathbf{X}_e) \propto \left(\prod_{l=0}^{d_l - 1} \delta_l\right)^{-1}, \tag{4.12}$$

where $d_l = n_R d_o = n_R \min(L n_T, N)$ is the *maximum achievable diversity order of the 2DLPT MIMO-OFDM system*.

When $L n_T \leq N$, the 2DLPT achieves full diversity of $d_l = n_R n_T L$.

■

### 4.6.3 Numerical Results

In this section, we present the simulation results. We first consider $2 \times 2$ and $2 \times 3$ flat fading MIMO-OFDM channels with FFT size of $N = 8$. As $(n_L n_T = 2) < N$, the overall diversity orders in the two channels are respectively $2 \times 2 \times 1 = 4$ and $2 \times 3 \times 1 = 6$. We will show that the diversity can be fully exploited by using the 2DLPT and MLD.

In addition to the 2DLPT system, we also consider the DM system without 2DLPT, the single dimensional frequency domain PT (1D FD-LPT) system, and single dimensional spatial domain PT (1D FD-LPT) system. The intention of the performance comparison is as follows.

For DM MIMO-OFDM without any PT, the frequency diversity in each single-input single-output channel is of order one (flat fading), and the spatial diversity order is two from the receive diversity achieved through MLD. When spatial domain transform is applied to the system, extra frequency diversity is made available to each single-input single-output channel. For the $n_T = 2$ flat fading channels considered, each single-input single-output channel after the SD-LPT has frequency diversity order of two. This extra frequency diversity, however, can only be exploited when a FD-LPT is applied in each layer. Hence, for the 1D SD-LPT system, the maximum diversity order is two for $2 \times 2$ and three for $2 \times 3$ systems, from the receive antennas and achieved through MLD, same as the DM system. For the 1D FD-LPT MIMO-OFDM systems, the frequency diversity in each layer is order one, hence the maximum diversity order of the systems remains as two for $2 \times 2$ and three for $2 \times 3$ systems. In summary, when MLD is used, only the 2DLPT MIMO-OFDM system can achieve the maximum diversity order. The other three schemes can only achieve the receive diversity, of order two for $2 \times 2$, and order three for $2 \times 3$ configuration.

Besides MLD, we also demonstrate the performance of the four schemes using zero-forcing (ZF) detection. In this case, all the four $2 \times 2$ systems have diversity order of one, and the four $2 \times 3$ systems have diversity order of two. The simulation scenarios and the corresponding achievable diversity order are summarized in Table 4.1 for $2 \times 2$, and in Table 4.2 for $2 \times 3$.

The simulated BER performance for $2 \times 2$ configuration with QPSK modulation is depicted in Fig. 4.19. As expected, the four systems with ZF detection has exactly the same performance, and the lowest diversity order. Then the DM system with no PT, as well as the 1D SD-LPT and FD-LPT systems with MLD have exactly the same performance, and higher diversity order than the ZF performance. The 2DLPT with MLD has the best performance and highest diversity order of all systems.

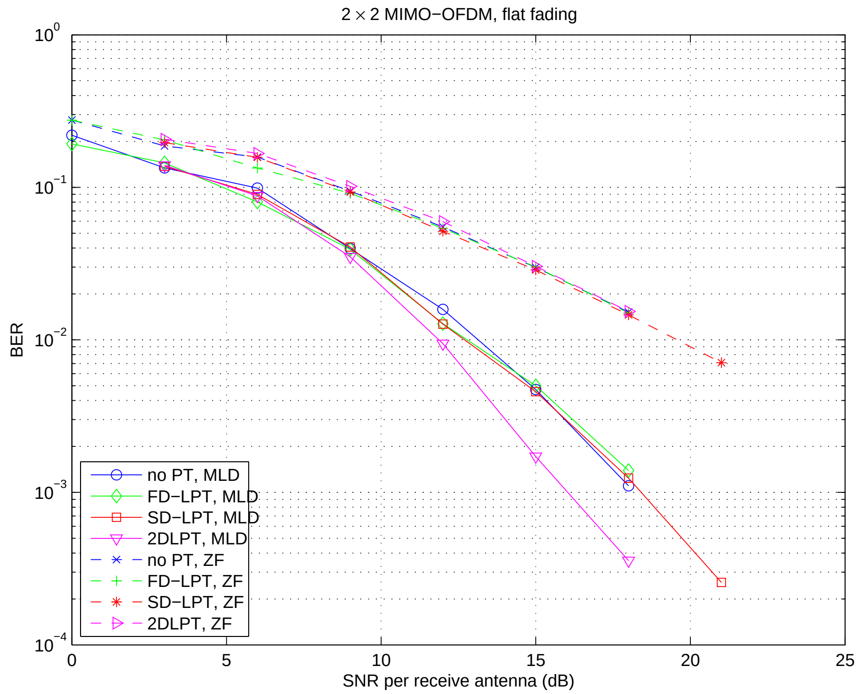


**Figure 4.19: BER performance of a $2 \times 2$ 2DLPT MIMO-OFDM system with MLD and ZF detection, flat-fading Rayleigh channel.**

The simulated BER performance for $2 \times 3$ configuration with QPSK modulation is depicted in Fig. 4.20. In this case, ZF detection results in diversity order of two for all the four schemes, which is the lowest diversity order. Then the DM system with no PT, the 1D SD-LPT and 1D FD-LPT systems with MLD have exactly the same performance, with diversity order of three. The 2DLPT with MLD has the best performance and highest diversity order of all the four schemes.

### 4.6.4 BICM-2DLPT MIMO-OFDM

Same as the other precoding schemes we studied, BICM can be applied to the 2DLPT MIMO-OFDM systems. In Fig. 4.21, a BICM-2DLPT MIMO-OFDM transmitter block diagram is presented.

At the receiver, interference-cancelation based iterative receiver as presented in Chapter 5 can be

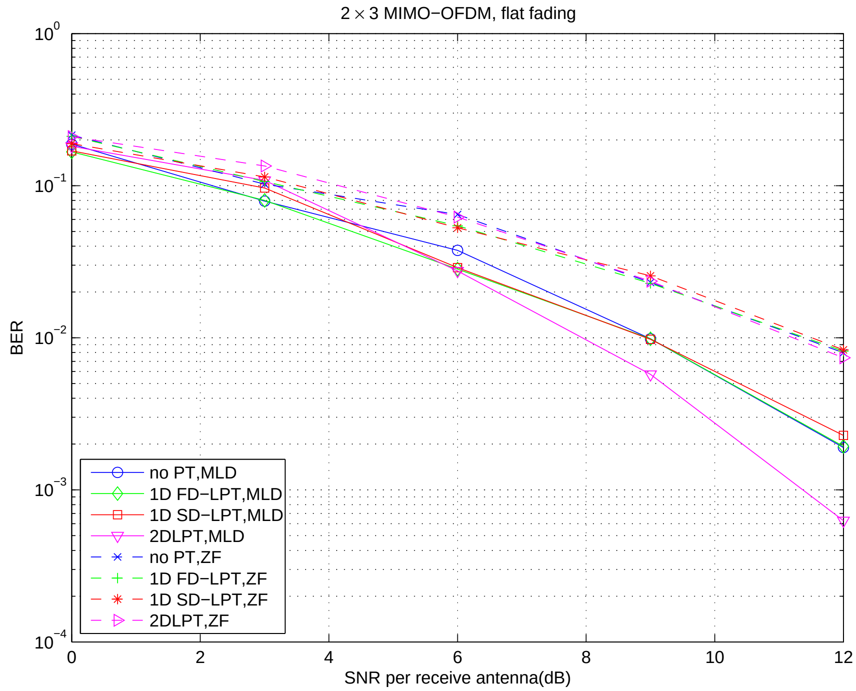


**Figure 4.20: BER performance of a $2 \times 3$ 2DLPT MIMO-OFDM system with MLD and ZF detection, flat-fading Rayleigh channel.**

**Table 4.1: Summary of the Simulation Setup, $2 \times 2$ Flat Fading Channel**

| Transmitter Setup | Receiver Setup | Achievable Diversity |
|---|---|---|
| 2DLPT | MLD | 4 |
| 1D FD-LPT | MLD | 2 |
| 1D SD-LPT | MLD | 2 |
| no PT (SDM) | MLD | 2 |
| 2DLPT | ZF | 1 |
| 1D FD-LPT | ZF | 1 |
| 1D SD-LPT | ZF | 1 |
| no PT (SDM) | ZF | 1 |

implemented for the BICM-2DLPT MIMO-OFDM systems. Presented in Fig. 4.22 is the simulated BER performance for $K = 3$ $R_c = \frac{1}{2}$ convolutional coded QPSK-modulated 2 2DLPT MIMO-OFDM system at iteration 5. The results of 2 1D-SDLPT-OFDM, and that of the Alamouti STBC without PT are also included in the figure for comparison. Obviously, the BICM-2DLPT system can exploit the frequency diversity much more effectively, resulting in steeper BER versus SNR slope of the BER curves. From

**Table 4.2: Summary of the Simulation Setup, $2 \times 3$ Flat Fading Channel**

| Transmitter Setup | Receiver Setup | Achievable Diversity |
|---|---|---|
| 2DLPT | MLD | 6 |
| 1D FD-LPT | MLD | 3 |
| 1D SD-LPT | MLD | 3 |
| no PT (SDM) | MLD | 3 |
| 2DLPT | ZF | 2 |
| 1D FD-LPT | ZF | 2 |
| 1D SD-LPT | ZF | 2 |
| no PT (SDM) | ZF | 2 |

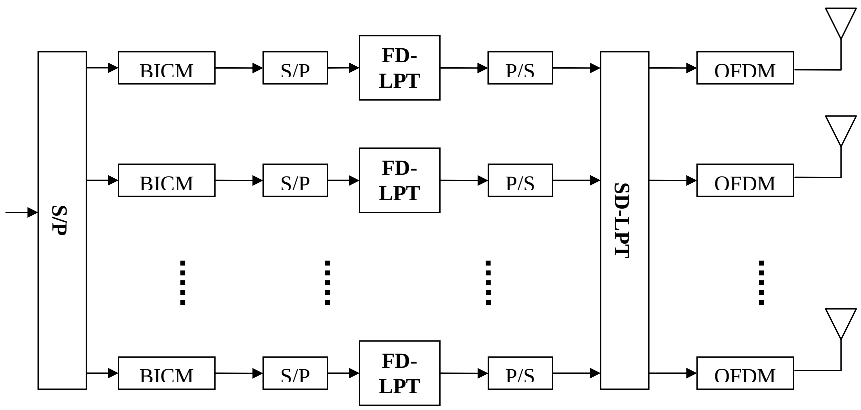


**Figure 4.21: Transmitter block diagram of 2DLPT MIMO-OFDM with BICM.**

the figure, we can also see that 2DLPT-OFDM has equivalent performance as the STBC system, with a performance difference of about 0.2dB. When we look at the FER performance as depicted in Fig. 4.23, we can see that the 2DLPT-OFDM and STBC have converged performance in the high SNR regions. For the 1D-SDLPT system, however, due to the large frequency diversity order available in the channel ($L = 16$), and the limited effective free distance of the FEC ($d_{\text{free}} = 5$), the frequency diversity can not be fully exploited, as exhibited by the slope of the BER/FER versus SNR performance curves.

## 4.7 Summary of the Chapter

In this chapter, we have studied the capacity and diversity performance of some precoded MIMO-OFDM channels and showed that none of these known precoding schemes achieved optimal capacity and diversity

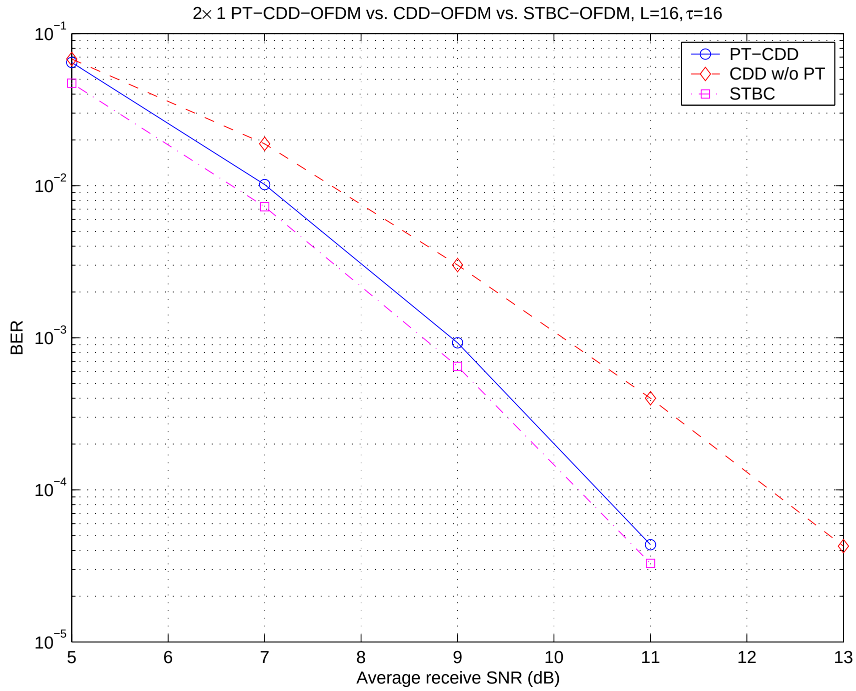


**Figure 4.22: BER performance of** $2 \times 1$ **PT-CDD-OFDM with** $K = 3$ $R_c = \frac{1}{2}$ **convolutional coded QPSK-modulated BICM.** $L = 16$**,** $\tau = 16$**.**

performance. We then proposed a two-dimensional linear pre-transformed MIMO-OFDM system which achieves simultaneously full capacity and full diversity when the number of spatial streams is equal to the number of transmit antennas. For the asymmetric MIMO-OFDM channel with more transmit than receive antennas, the proposed 2DLPT system achieves full diversity and maximum capacity of $n_R \times n_R$ channel.

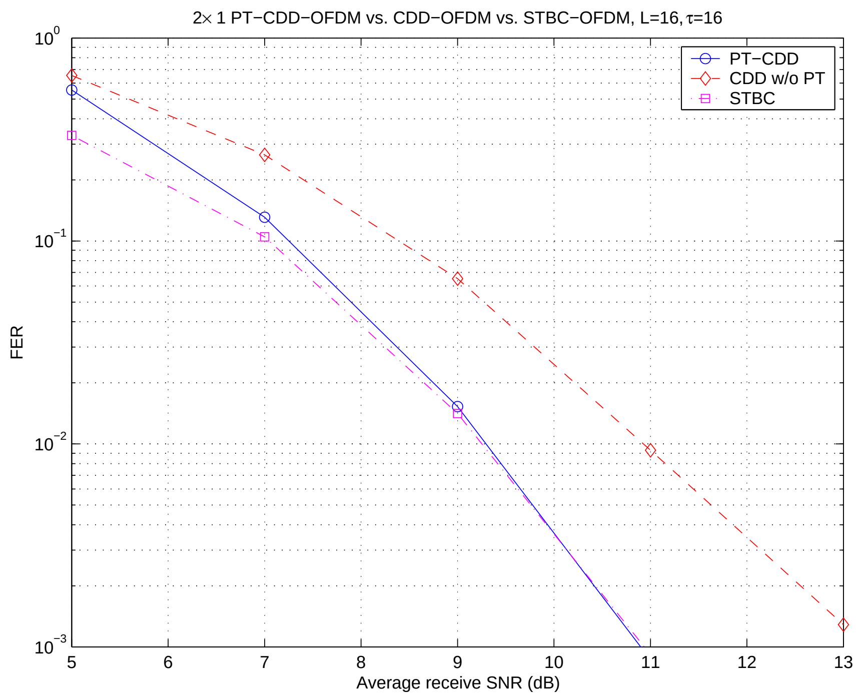


**Figure 4.23: FER performance of $2 \times 1$ PT-CDD-OFDM with $K = 3$ $R_c = \frac{1}{2}$ convolutional coded QPSK-modulated BICM. $L = 16$, $\tau = 16$.**

# Chapter 5

# Bayesian Iterative Turbo Receiver

## 5.1 Introduction

When FEC code is applied in a MIMO system, an iterative ("turbo") receiver in which soft information is exchanged between the detector and the decoder can significantly improve decoding performance. There have been many works on turbo receiver design. For example, the earliest literature on turbo receiver was in the field of code division multiple access (CDMA) systems, by Wang and Poor [85], and Alexander *et. al.* [86][87][88]. In the field of MIMO systems, Lu *et. al.* proposed a linear minimum mean squared error (LMMSE) interference cancellation (IC)-based turbo receiver in [89] for multiuser STC systems (referred as groupwise STC systems in this thesis), which was a straightforward extension of Wang's work in [85] to MIMO systems. Sellathurai and Haykin applied the LMMSE-IC turbo receiver to coded BLAST systems in [90] (the so called turbo BLAST, or T-BLAST), and further evaluated its performance in correlated Rayleigh fading channels [91]. The practical virtue of T-BLAST has also been verified in experiments [92]. In [93], Caire *et. al.* developed a generalized framework on iterative receivers for CDMA systems, which is applicable to MIMO systems as well. Based on the factor-graph representation and the sum-product algorithm (SPA) [94], they showed that the estimated interference at each iteration is a function of the decoders' extrinsic information (EXT), rather than of the decoders' *a posteriori* probability (APP). The EXT is used to obtain the *a priori* probability of the coded bits, from which the *statistical mean*, or the prior estimate [95], of the transmitted signals are calculated. IC is then performed and SISO decoding implemented. Different filtering schemes can be applied in the IC step, for example, matched filter (MF) or LMMSE. In a semi-tutorial paper [96], Biglieri *et. al.* discussed a class of iterative receivers for MIMO systems that combine soft decoder and spatial interference cancellers and analyzed their performances using extrinsic information transfer (EXIT) charts. The IC-based turbo receivers considered in these papers are referred as *"conventional"* turbo receivers in this thesis.

In these conventional turbo receivers, only Phase Shift Keying (PSK) modulation signals, e.g., BPSK, 8PSK, etc., were considered. However, to address the key concern of spectral efficiency, higher order modulation needs to be applied, e.g., Quadrature Amplitude Modulation (QAM). The interference statistical mean estimation using the SISO decoder EXT should be studied. Due to the increase of number of bits per modulation symbol, a number of exponential terms need to be computed in the soft decision functions (SDF's), complexity will therefore be a major concern in practical implementation. Simplified SDF's in which the exponential terms are converted to linear calculations are desired.

When exact *a priori* probability is available, statistical mean provides the best estimate of the interference signals. However, when the *a priori* probability obtained from the SISO decoders has some degree of inaccuracy, for example, at low to medium SNR values, or in the initial iterations in the turbo receiver, or when punctured code is used in the system, the accuracy of the estimated interference mean will be rather poor. Therefore, schemes have to be found to compensate for the estimation errors.

In this chapter, we study the above two problems in turbo receiver. We first present the iterative receiver design and derive the exact SDF's for two commonly used MQAM modulation signals, i.e., 16QAM and 64QAM, based on the estimated *a priori* probability. We then proceed to derive the simplified SDF's using Maclaurin series. Performance comparison of the simplified and the exact SDF's will show that the simplified SDF's introduce negligible performance degradations.

To improve the estimation accuracy of the interference signals, we propose a novel Bayesian MMSE (BMMSE) turbo receiver that exploits EXT in the *Bayesian estimation (BE)* of the interference signals. We start by deriving the BMMSE estimate [95] [97], which is the mean of the posterior probability density function (pdf) of the desired signal, and show that the BMMSE estimate conditioned on the received signal and the estimated interference (from the previous iteration) is a function of *both* EXT and the IC-MRC decision statistic. We refer this receiver as the *Bayesian IC-MRC turbo receiver* to differentiate it from the conventional IC-MRC turbo receiver in [93]. We show that the BMMSE estimate has much smaller mean squared error (MSE) than the statistical mean. The improved MSE leads to better bit error rate (BER) and frame error rate (FER) performance. When the same number of iterations are used, the Bayesian turbo receiver can achieve at least 1 dB's performance gain over the receiver of [93] at BER of $10^{-5}$.

Using the Gaussian model developed in [98] for the output of conventional LMMSE-IC detectors, we further apply the BMMSE estimation to LMMSE-IC receivers and refer this class of receiver as the *Bayesian LMMSE-IC turbo receiver*. Similar to IC-MRC receivers, the BMMSE estimate is a function of *both* EXT and the LMMSE-IC decision statistic. The Bayesian LMMSE-IC turbo receiver is desired for a system deploying punctured code and high order modulation to achieve high spectral efficiency, as both accurate interference estimation by BMMSE and effective interference suppression by LMMSE filter is required in order to guarantee convergence to the lower performance bound. Simulation results show performance gains over the conventional LMMSE-IC receivers.

The EXT and decision statistic represent two types of information from two different domains. The decision statistic is obtained from spatial domain, by making use of the received signal and the estimated interference from the interference layers. The EXT is obtained from time domain (when single-carrier transmission scheme is used) or frequency domain (when multi-carrier modulation is used) through the knowledge of the other symbols in the same layer and same domain, by exploiting their correlation produced by the encoder (and the modulation mapper). Therefore, the information is not repetitive but complementary to interference estimation. This leads to much more accurate IC and thus improves the turbo receiver performance, as will be demonstrated by the lower BER and FER values in the simulations to be presented in this chapter, as well as the much higher output mutual information in the EXIT chart analysis which will be presented in Chapter 6.

The rest of the chapter is organized as follows. In Section 5.2, we dedicate our study to the SDF simplification in conventional receivers. We will give a brief overview of the conventional turbo receiver design and derive both the exact and simplified SDF's based on the estimated *a priori* probability.

The simulated performance of the proposed SDF's will be compared with the exact SDF's to show the negligible degradation introduced in the proposed simplification. In Section 5.3, we derive the Bayesian IC-MRC turbo receiver, and extend it to the LMMSE-IC receiver in Section 5.4. The simulated BER and FER performances of the Bayesian IC-MRC and LMMSE-IC receivers are presented in Section 5.6. Finally in Section 5.7, we make our concluding remarks.

Throughout this chapter, we use the GSTBC OFDM system model in the derivations unless pointed out otherwise. The results can be extended to other systems in a very straightforward manner. We also use the terms maximal ratio combining (MRC) and MF interchangeably. For transmitted signal $X$, $\breve{X}$, $\hat{X}$, and $\tilde{X}$ represent its BMMSE estimate, statistical mean estimate, and decision statistic at the detector output, respectively.

## 5.2 SDF Simplification in Conventional Turbo Receivers

### 5.2.1 The Conventional Turbo Receiver

Referring to the system block diagram in Fig. 3.2, and making use of the signal model in (3.14), we will first give a brief overview of the conventional turbo receiver, making use of the IC-MRC receiver depicted in Fig. 5.1 as example.

In this figure, signal $\mathbf{R}$ denotes the received signal at each subcarrier, as defined in (3.14) and (3.16), $\tilde{X}_{n_t}$ and $\hat{X}_{n_t}$, $n_t = 1,\ 2,\ \cdots,\ n_T$ denotes the decision statistic from the IC-MRC detector, and the statistical mean estimate of signal $X_{n_t}$, respectively. $\lambda(\mathrm{I})$ and $\lambda(\mathrm{O})$ denote the input *a priori* and output extrinsic information of the SISO decoders.

At each iteration, an IC and MRC ("IC & MRC") unit is implemented for each subcarrier to cancel the estimated interference from other antenna groups, and then exploit the diversity from the spatial domain by MRC, generating the following decision statistic

$$\tilde{X}_{k,i} = \mathbf{H}_k^H \left( \mathbf{R} - \tilde{\mathbf{I}}_{k,i} \right), \tag{5.1}$$

where subscripts $k$ and $i$ denote the transmitted signal index and the iteration numbers, respectively, $\mathbf{H}_k$

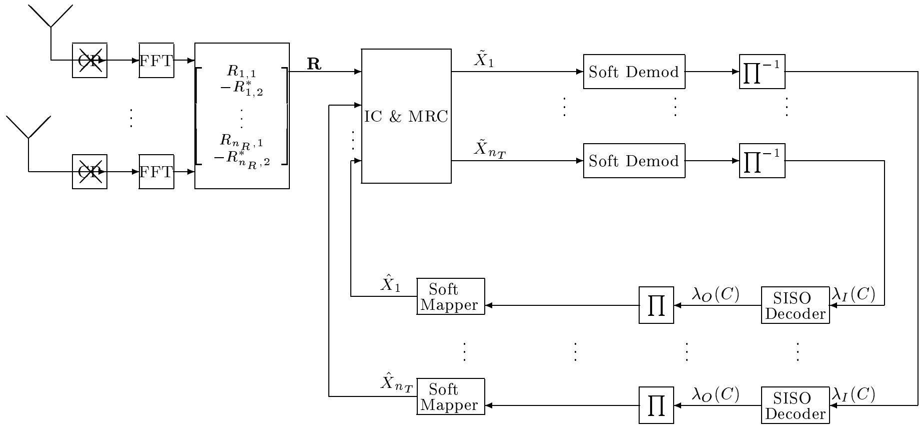


**Figure 5.1: The iterative receiver for BICM GSTBC-OFDM systems. $\prod$ and $\prod^{-1}$ stand for interleaver and deinterleaver, respectively.**

denotes the $k$th column of $\mathbf{H}$, and $\tilde{\mathbf{I}}_{k,i}$ represents the estimated interference which is calculated as

$$\tilde{\mathbf{I}}_{k,i} = \sum_{\substack{p=1 \\ p \neq k}}^{n_T} \mathbf{H}_p \hat{X}_{p,i-1},$$

where $\hat{X}_{p,i-1}$ is the estimated statistical mean of $X_p$ at iteration $(i-1)$, as given later in (5.7).

An initial estimate of $\tilde{X}$ can be obtained through ZF interference suppression (IS) [44], expressed as

$$\tilde{\mathbf{X}}_{\mathbf{0}} = \mathbf{H}^\dagger \mathbf{R}, \tag{5.2}$$

where subscript $0$ denotes $0$th iteration (initialization), $\mathbf{H}^\dagger$ denotes the pseudo-inverse matrix of $\mathbf{H}$.

Alternatively, an LMMSE filter can be used, expressed as

$$\tilde{\mathbf{X}}_{\mathbf{0}} = \left( \mathbf{H}^H \mathbf{H} + \frac{2\sigma^2}{\sigma_x^2} \mathbf{I} \right)^{-1} \mathbf{H}^H \mathbf{R}, \tag{5.3}$$

where $\sigma^2$ is the AWGN noise variance defined in (3.14), $\sigma_x^2$ is the signal power, and $\mathbf{I}$ is the identity matrix.

**The Soft Demodulator**

The soft demodulator calculates the updated log-likelihood ratio (LLR) of coded bits, i.e., the extrinsic metric values, using the detector output $\tilde{X}_{k,i}$, as

$$\begin{aligned}
\Lambda^i(k,l) &= \log \frac{p(c_k^l = 1|\tilde{X}_{k,i})}{p(c_k^l = 0|\tilde{X}_{k,i})} \\
&= \log \frac{p\left(\tilde{X}_{k,i}|c_k^l = 1\right)\, p(c_k^l = 1)}{p\left(\tilde{X}_{k,i}|c_k^l = 0\right)\, p(c_k^l = 0)} \\
&= \log \frac{p\left(\tilde{X}_{k,i}|c_k^l = 1\right)}{p\left(\tilde{X}_{k,i}|c_k^l = 0\right)} + \underbrace{\log \frac{p(c_k^l = 1)}{p(c_k^l = 0)}}_{\lambda_{\mathbf{a}}^{i-1}(k,l)} \\
&= \log \frac{\sum_{S_k \in \Omega^+} p\left(\tilde{X}_{k,i}|S_k\right) p\left(S_k|c_k^l = 1\right)}{\sum_{S_k \in \Omega^-} p\left(\tilde{X}_{k,i}|S_k\right) p\left(S_k|c_k^l = 0\right)} + \lambda_{\mathbf{a}}^{i-1}(k,l) \\
&= \underbrace{\log \frac{\sum_{S_k \in \Omega^+} p(\tilde{X}_{k,i}|S_k) \prod_{\substack{j=1 \\ j \neq l}}^{\log_2 M} p(c_k^j)}{\sum_{S_k \in \Omega^-} p(\tilde{X}_{k,i}|S_k) \prod_{\substack{j=1 \\ j \neq l}}^{\log_2 M} p(c_k^j)}}_{\lambda_{\mathbf{e}}^i(k,l)} + \lambda_{\mathbf{a}}^{i-1}(k,l),
\end{aligned} \tag{5.4}$$

where $l = 1,\ 2,\ \cdots,\ \log_2 M$, $\Omega^+ = \left\{S_k :\ c_k^l = 1\right\}$, i.e., the subset of modulation symbols whose $l$th bit is 1, and $\Omega^- = \left\{S_k :\ c_k^l = 0\right\}$ represents the signal subset whose $l$th bit is 0. $p(c_k^l)$ is the estimated *a priori* probability of $c_k^l$, as given later in (5.5). At the initialization (iteration 0), $p(c_k^l) = \frac{1}{2}$, and $\lambda_{\mathbf{a}}^0(k,l) = 0$. It can be seen that the LLR is composed of two parts - the updated bit metric value $\lambda_{\mathbf{e}}^i(k,l)$ and the *a priori* information $\lambda_{\mathbf{a}}^{i-1}(k,l)$.

$\{\lambda_{\mathbf{e}}^i(k,l)\}$ are de-interleaved to generate the input $\{\lambda_I(C)\}$ for the SISO decoder which produces the *a posteriori* and extrinsic LLR information $\{\lambda_O(C)\}$ (EXT) for the coded bits.

The extrinsic information is then interleaved to generate the *a priori* information $\lambda_{\mathbf{a}}^i(k,l)$ for the "Soft Mapper" in which the *a priori* probability of the coded bits is first calculated as

$$\left\{ \begin{aligned}
P(c_k^l = 1) &= \frac{\exp(\lambda_{\mathbf{a}}^i(k,l))}{1 + \exp(\lambda_{\mathbf{a}}^i(k,l))} \\
P(c_k^l = 0) &= \frac{1}{1 + \exp(\lambda_{\mathbf{a}}^i(k,l))}
\end{aligned} \right. \tag{5.5}$$

from which the symbol *a priori* probability is computed as

$$P\left\{S_m\right\} = \prod_{l=1}^{\log_2 M} P\left(c_m^l\right), \tag{5.6}$$

under the assumption that $S_m$ is mapped from bits $[c_m^1,\ c_m^2, \cdots, c_m^{\log_2 M}]$, and these bits are uncorrelated with sufficient interleaving.

Finally, the modulated symbols are estimated as its statistical mean:

$$\hat{X}_{k,i} = E\{X_{k,i}\} = \sum_{S_m \in \Omega} S_m P(S_m), \tag{5.7}$$

where $\Omega$ denotes the signal set for all the modulation symbols.

### 5.2.2 Exact SDF's

Applying the gray mapping rules as given in Table 5.1-5.5 for BPSK, QPSK, 8PSK, 16QAM, and 64QAM signals, respectively, we can obtain the exact SDF's in the conventional turbo receivers, as

- **BPSK**

$$\hat{X}_{m,i} = (+1) \times P(c_m = 1) + (-1) \times P(c_m = 0) = \tanh\left(\frac{\lambda_{\mathbf{a}}^i(m)}{2}\right) \tag{5.8}$$

- **QPSK**

$$\hat{X}_{m,i} = \tanh\left(\frac{\lambda_{\mathbf{a}}^i(m,1)}{2}\right) + j\tanh\left(\frac{\lambda_{\mathbf{a}}^i(m,2)}{2}\right) \tag{5.9}$$

- **8PSK**

$$\hat{X}_{m,i} = \tanh\left(-\frac{\lambda_{\mathbf{a}}^i(m,2)}{2}\right)\frac{a + be^{\lambda_{\mathbf{a}}^i(m,3)}}{1 + e^{\lambda_{\mathbf{a}}^i(m,3)}} + j\tanh\left(-\frac{\lambda_{\mathbf{a}}^i(m,1)}{2}\right)\frac{ae^{\lambda_{\mathbf{a}}^i(m,3)} + b}{1 + e^{\lambda_{\mathbf{a}}^i(m,3)}}, \tag{5.10}$$

  where $a = \frac{\sqrt{2+\sqrt{2}}}{2}$, and $b = \frac{\sqrt{2-\sqrt{2}}}{2}$.

- **16QAM**

$$\hat{X}_{m,i} = \tanh\left(\frac{\lambda_{\mathbf{a}}^i(m,1)}{2}\right)\frac{3 + e^{\lambda_{\mathbf{a}}^i(m,2)}}{1 + e^{\lambda_{\mathbf{a}}^i(m,2)}} + j\tanh\left(\frac{\lambda_{\mathbf{a}}^i(m,3)}{2}\right)\frac{3 + e^{\lambda_{\mathbf{a}}^i(m,4)}}{1 + e^{\lambda_{\mathbf{a}}^i(m,4)}} \tag{5.11}$$

- **64QAM**

$$\begin{aligned}\hat{X}_{m,i} &= \tanh\left(\frac{\lambda_{\mathbf{a}}^{i}(m,1)}{2}\right)\frac{7+5e^{\lambda_{\mathbf{a}}^{i}(m,3)}+3e^{\lambda_{\mathbf{a}}^{i}(m,2)+\lambda_{\mathbf{a}}^{i}(m,3)}+e^{\lambda_{\mathbf{a}}^{i}(m,2)}}{\left(1+e^{\lambda_{\mathbf{a}}^{i}(m,2)}\right)\left(1+e^{\lambda_{\mathbf{a}}^{i}(m,3)}\right)} \\ &\quad +j\tanh\left(\frac{\lambda_{\mathbf{a}}^{i}(m,4)}{2}\right)\frac{7+5e^{\lambda_{\mathbf{a}}^{i}(m,6)}+3e^{\lambda_{\mathbf{a}}^{i}(m,5)+\lambda_{\mathbf{a}}^{i}(m,6)}+e^{\lambda_{\mathbf{a}}^{i}(m,5)}}{\left(1+e^{\lambda_{\mathbf{a}}^{i}(m,5)}\right)\left(1+e^{\lambda_{\mathbf{a}}^{i}(m,6)}\right)} \\ &= \tanh\left(\frac{\lambda_{\mathbf{a}}^{i}(m,1)}{2}\right)\mathcal{F}\left(\lambda_{\mathbf{a}}^{i}(m,2),\lambda_{\mathbf{a}}^{i}(m,3)\right) \\ &\quad +j\tanh\left(\frac{\lambda_{\mathbf{a}}^{i}(m,4)}{2}\right)\mathcal{F}\left(\lambda_{\mathbf{a}}^{i}(m,5),\lambda_{\mathbf{a}}^{i}(m,6)\right), \end{aligned} \tag{5.12}$$

where $\mathcal{F}(x,y)$ in (5.12) is defined as

$$\mathcal{F}(x,y)=\frac{7+5e^{y}+3e^{x+y}+e^{x}}{(1+e^{x})(1+e^{y})}. \tag{5.13}$$

The detailed derivation of (5.10) and (5.11) can be found in Appendix A. Following the same procedure, the results in (5.8), (5.9), and (5.12) can be obtained.

The SDF's in (5.8) to (5.12) consist of a number of exponential terms. They are computationally expensive in practical implementations. Simplified SDF's are therefore desired.

**Table 5.1: BPSK Gray Mapping Table.**

| INPUT BIT($b_1$) | I-OUT | Q-OUT |
|---|---|---|
| 0 | -1 | 0 |
| 1 | 1 | 0 |

**Table 5.2: QPSK Gray Mapping Table.**

| INPUT BITS($b_1 b_2$) | I-OUT | Q-OUT |
|---|---|---|
| 00 | -1 | -1 |
| 01 | -1 | 1 |
| 11 | 1 | 1 |
| 10 | 1 | -1 |

**Table 5.3: 8PSK Gray Mapping Table.**

| INPUT BITS($b_1b_2$) | I-OUT | Q-OUT |
|---|---|---|
| 000 | a | b |
| 001 | b | a |
| 011 | -b | a |
| 010 | -a | b |
| 110 | -a | -b |
| 111 | -b | -a |
| 101 | b | -a |
| 100 | a | -b |

$a = \frac{\sqrt{2+\sqrt{2}}}{2}$, $b = \frac{\sqrt{2-\sqrt{2}}}{2}$

**Table 5.4: 16QAM Gray Mapping Table.**

| INPUT BITS($b_1b_2$) | I-OUT |
|---|---|
| 00 | -3 |
| 01 | -1 |
| 11 | 1 |
| 10 | 3 |

| INPUT BITS($b_3b_4$) | Q-OUT |
|---|---|
| 00 | -3 |
| 01 | -1 |
| 11 | 1 |
| 10 | 3 |

**Table 5.5: 64QAM Gray Mapping Table.**

| INPUT BITS($b_1b_2$) | I-OUT |
|---|---|
| 000 | -7 |
| 001 | -5 |
| 011 | -3 |
| 010 | -1 |
| 110 | 1 |
| 111 | 3 |
| 101 | 5 |
| 100 | 7 |

| INPUT BITS($b_3b_4$) | Q-OUT |
|---|---|
| 000 | -7 |
| 001 | -5 |
| 011 | -3 |
| 010 | -1 |
| 110 | 1 |
| 111 | 3 |
| 101 | 5 |
| 100 | 7 |

### 5.2.3 Simplified SDF's

#### Simplified SDF's for BPSK and QPSK

As

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}},$$

we have

$$\lim_{x\to+\infty} \tanh(x) = 1,$$
$$\lim_{x\to-\infty} \tanh(x) = -1,$$

and its Maclaurin series as

$$\tanh(x) = x - \frac{1}{3}x^3 + \frac{2}{15}x^5 - \cdots, \qquad |x| < \frac{\pi}{2},$$

we therefore can approximate $\tanh(x)$ as

$$\tanh(x) \approx \tilde{f}_1(x) = \begin{cases} +1, & x \geq 1 \\ x, & |x| < 1 \\ -1, & x \leq -1 \end{cases} . \tag{5.14}$$

The simplified SDF's for BPSK and QPSK are hence

- **BPSK**

$$\breve{X}_{m,i} = \tilde{f}_1\left(\frac{\lambda_{\mathbf{a}}^i(m)}{2}\right); \tag{5.15}$$

- **QPSK**

$$\breve{X}_{m,i} = \tilde{f}_1\left(\frac{\lambda_{\mathbf{a}}^i(m,1)}{2}\right) + j\tilde{f}_1\left(\frac{\lambda_{\mathbf{a}}^i(m,2)}{2}\right). \tag{5.16}$$

From our investigation on soft decision based iterative interference cancelation for uncoded GSTBC OFDM systems [44], the above approximation introduces nearly no performance degradation.

**Simplified SDF for 8PSK**

In 8PSK modulation, there are two constant values $a = \frac{\sqrt{2+\sqrt{2}}}{2}$, $b = \frac{\sqrt{2-\sqrt{2}}}{2}$, and $a = \left(\sqrt{2}+1\right)b$. We can therefore write its exact SDF as

$$\hat{X}_{m,i} = b\tanh\left(-\frac{\lambda_{\mathbf{a}}^i(m,2)}{2}\right)\left(\frac{\sqrt{2}}{1+e^{\lambda_{\mathbf{a}}^i(m,3)}}+1\right) + jb\tanh\left(-\frac{\lambda_{\mathbf{a}}^i(m,1)}{2}\right)\left(\frac{\sqrt{2}}{1+e^{-\lambda_{\mathbf{a}}^i(m,3)}}+1\right). \tag{5.17}$$

Defining

$$f_2(x) = \frac{1}{1+e^x}, \tag{5.18}$$

we have

$$\lim_{x \to +\infty} f_2(x) = 0, \quad \text{and} \quad \lim_{x \to -\infty} f_2(x) = 1.$$

Making use of the Maclaurin series of $f_2(x)$, we can have the following linear function to approximate $f_2(x)$

$$f_2(x) \approx \tilde{f}_2(x) = \begin{cases} 0, & x \geq 2 \\ \frac{1}{2} - \frac{x}{4}, & |x| < 2 \\ 1, & x \leq -2 \end{cases} . \tag{5.19}$$

Therefore, we have the simplified SDF for 8PSK signals as

$$\breve{X}_{m,i} = \tilde{f}_1\left(-\frac{\lambda_{\mathbf{a}}^i(m,2)}{2}\right)\left(\sqrt{2}\tilde{f}_2(\lambda_{\mathbf{a}}^i(m,3)) + 1\right) + j\tilde{f}_1\left(-\frac{\lambda_{\mathbf{a}}^i(m,1)}{2}\right)\left(\sqrt{2}\tilde{f}_2(-\lambda_{\mathbf{a}}^i(m,3)) + 1\right), \tag{5.20}$$

where $\tilde{f}_1(x)$ is defined in (5.14), and $\tilde{f}_2(x)$ is defined in (5.19).

**Simplified SDF for 16QAM**

Defining

$$f_3(x) = \frac{3 + e^x}{1 + e^x},$$

we have

$$\lim_{x \to +\infty} f_3(x) = 1, \quad \text{and} \quad \lim_{x \to -\infty} f_3(x) = 3.$$

Making use of the Maclaurin series of $f_3(x)$, we can have its linear approximation as

$$f_3(x) \approx \tilde{f}_3(x) = \begin{cases} 1, & x \geq 2 \\ 2 - \frac{x}{2}, & -2 < x < 2 \\ 3, & x \leq -2 \end{cases} . \tag{5.21}$$

and the following approximated linear SDF for 16QAM

$$\breve{X}_{m,i} = \tilde{f}_1\left(\frac{\lambda_{\mathbf{a}}^i(m,1)}{2}\right)\tilde{f}_3\left(\lambda_{\mathbf{a}}^i(m,2)\right) + j\tilde{f}_1\left(\frac{\lambda_{\mathbf{a}}^i(m,3)}{2}\right)\tilde{f}_3\left(\lambda_{\mathbf{a}}^i(m,4)\right), \tag{5.22}$$

where $\tilde{f}_1(x)$ is defined in (5.14).

**Simplified SDF for 64QAM**

In order to obtain the simplified linear SDF for 64QAM, we first decompose $\mathcal{F}(x, y)$ as

$$\mathcal{F}(x, y) \quad = \quad 3 + 2f_2(x) - 2f_2(y) + 4f_2(x)f_2(y),$$

with $f_2(x)$ is defined in (5.18).

Making use of the result in (5.19), we have the following approximation for $\mathcal{F}(x, y)$

$$\widetilde{\mathcal{F}}(x, y) = 3 + 2\tilde{f}_2(x) - 2\tilde{f}_2(y) + 4\tilde{f}_2(x)\tilde{f}_2(y), \tag{5.23}$$

The approximated linear SDF for 64QAM is then

$$\breve{X}_{m,i} = \tilde{f}_1\left(\frac{\lambda^i_{\mathbf{a}}(m,1)}{2}\right)\widetilde{\mathcal{F}}\left(\lambda^i_{\mathbf{a}}(m,2), \lambda^i_{\mathbf{a}}(m,3)\right) + j\tilde{f}_1\left(\frac{\lambda^i_{\mathbf{a}}(m,4)}{2}\right)\widetilde{\mathcal{F}}\left(\lambda^i_{\mathbf{a}}(m,5), \lambda^i_{\mathbf{a}}(m,6)\right). \tag{5.24}$$

Fig. 5.2 and Fig. 5.3 depict the comparison of the exact and simplified SDF's for 16QAM and 64QAM signals, respectively. From these two figures, we can see a very accurate approximation for the simplified SDF.

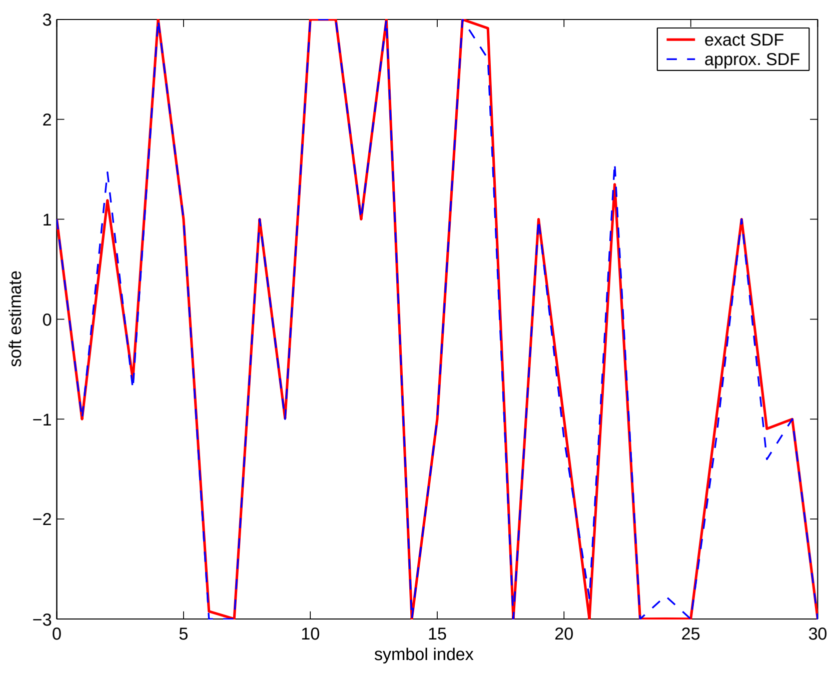


**Figure 5.2: Comparison of the exact and approximated SDF's for 16QAM signals.**

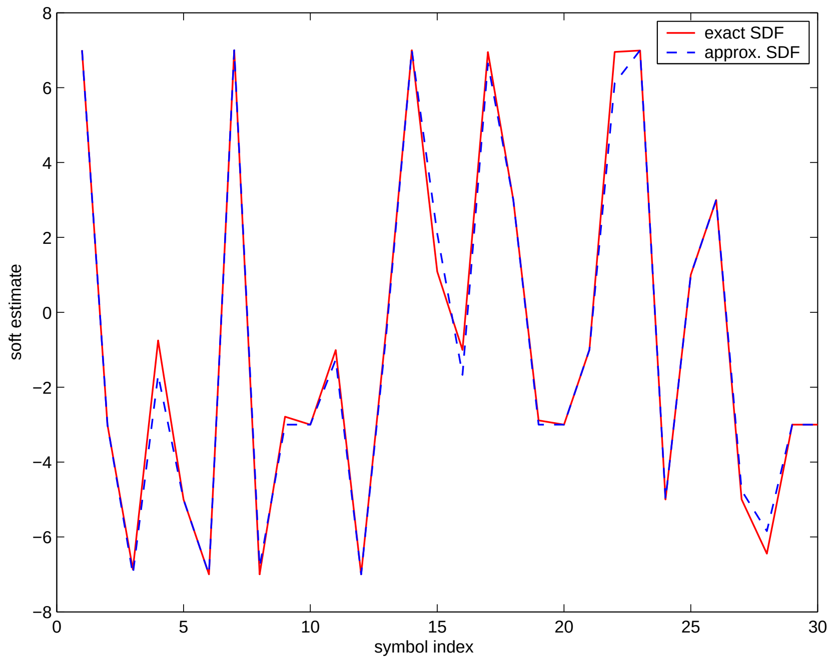


**Figure 5.3: Comparison of the exact and approximated SDF's for 64QAM signals.**

### 5.2.4 Simulation Results

In this section, we present the simulated BER performance of both the exact and the approximated SDF's for an $8 \times 4$ GSTBC-OFDM system. Rate $R_c = \frac{1}{2}$ CC with constraint length 3 and generation function of $(5, 7)_{\mathbf{octal}}$ is used. Rate $R_c = \frac{3}{4}$ CC was obtained through puncturing according to the puncturing pattern given in [99] and [100]. Bit interleavers proposed in IEEE 802.11a standard are adopted [2] in the simulations. Three modulation schemes are considered, namely QPSK, 16QAM and 64QAM. 64 subcarriers and 16 symbols of CP are assumed in OFDM modulation. Of the 64 subcarriers, only 48 are used to transmit data, as specified in IEEE 802.11a standard [2]. Sixteen independent multipaths are generated in the channel, with each multipath having i.i.d. zero mean complex Gaussian coefficient. Unless otherwise stated, the IC-MRC filtering is used in the iterations.

In Fig. 5.4, we present the performance of conventional IC-MRC receiver for $R_c = \frac{1}{2}$ QPSK performance. ZFIS is used in initialization. The simulated performance of soft decision Viterbi decoded single group GSTBC-OFDM with four receive antennas is also included in the figure. It gives the lower bound of the $8 \times 4$ GSTBC-OFDM system, which is achieved with perfect IC. From the figure, we can see that the BER improves from iteration to iteration, and more importantly, the slope of the BER versus SNR curves gets steeper and steeper, suggesting that the diversity order is getting higher and higher. However, its convergence speed is rather slow. From iteration 5 onward, the improvement from iteration to iteration is within 0.2dB, and at iteration 8, it is about 0.8dB away from the lower bound at BER = $10^{-5}$.

LMMSE IS can improve the convergence performance and hence bridge the gap between the turbo receiver and the lower bound, as shown in Fig. 5.5. The turbo receiver converges at iteration 4, and it approaches the lower bound at SNR = 6dB, BER of $4 \times 10^{-5}$.

When puncturing is applied to the FEC for higher spectral efficiency, the EXT accuracy will be degraded, which will lead to accuracy degradation in interference estimation and cancelation. In this case, even with the better LMMSE IS initialization, the conventional IC-MRC turbo receiver may not achieve convergence, instead, divergence can be observed in the BER in various iterations, as shown in Fig. 5.6.

One way to solve this problem is to use LMMSE-IC filtering scheme in the iterations. As the LMMSE filtering can mitigate the residual interference effectively, good convergence can be obtained, as shown in Fig. 5.7. With this scheme, however, we need to compute matrix inversion for each signal stream at each subcarrier and in each iteration, therefore it is very complicated. Other schemes are therefore desired. We will propose a Bayesian interference estimation scheme in Section 5.3 as an alternative solution. The proposed scheme can also work with LMMSE-IC, as we will show in Section 5.4.

In Fig. 5.8, we depict the simulation results for 16QAM systems with the exact SDF. The lower bound is also included in the figure to benchmark the performance. LMMSE IS is used in initialization, and IC-MRC filtering is used in the iterations. From iteration to iteration, we can see higher and higher diversity gains achieved, which is a result of improved accuracy of interference estimation and cancelation. As no puncturing is applied in the CC, the iterative receiver converges to the lower bound at BER of $10^{-4}$.

We next present in Fig. 5.9 the simulated performance for 64QAM with the exact SDF. Same as in Fig. 5.8, we depict the BER curves at the LMMSE IS, iterations 1-5, and the lower bound. We can see that same as the 16QAM system, the iterative receiver converges at fourth iteration, and it is approaching the lower bound within the presented BER and SNR range. The touching point to the lower bound, however, lies at BER values lower than $10^{-5}$. This may be due to the denser constellation of 64QAM which will

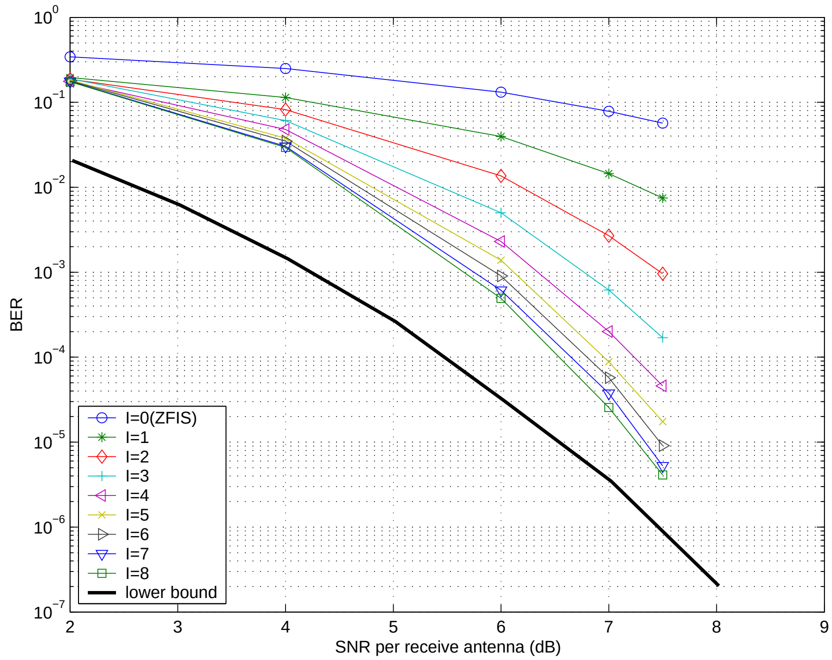


**Figure 5.4: Conventional IC-MRC turbo receiver performance for $8 \times 4$ GSTBC OFDM system. $R_c = \frac{1}{2}$ $K = 3$ CC, QPSK modulation, exact SDF, ZFIS initialization.**

requires higher SNR for accurate interference estimation.

The receiver performance with the simplified SDF for QPSK signals is depicted in Fig. 5.10 and compared with that with the exact SDF. We show only the results of iteration 1, 3, and 4 in the figure which prove that the simplification in the SDF introduces only negligible degradation.

We then present in Fig. 5.11 the receiver performance with the simplified SDF for 16QAM system. Results at iteration 1, 3 and 4 are depicted in the figure and are compared with those with the exact SDF. It can be seen that very little degradation is introduced by the SDF simplification.

In Fig. 5.12, we depict the receiver performance with the simplified SDF for 64QAM at iterations 1, 3 and 4 and compare with those with the exact SDF. Similar to 16QAM, only marginal degradation is introduced by the SDF simplification.

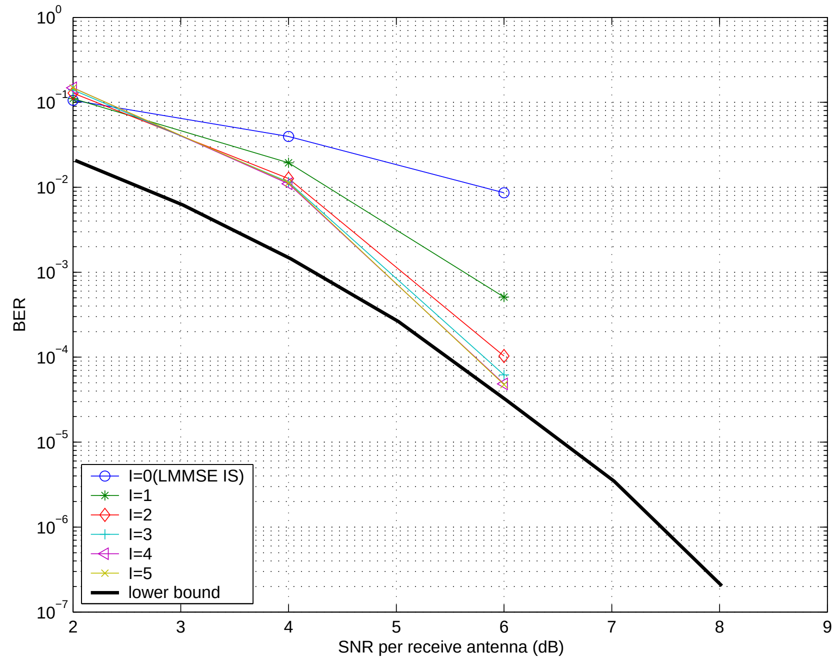


**Figure 5.5: Conventional IC-MRC turbo receiver performance for $8 \times 4$ GSTBC OFDM system. $R_c = \frac{1}{2}$ $K = 3$ CC, QPSK modulation, exact SDF, LMMSEIS initialization.**

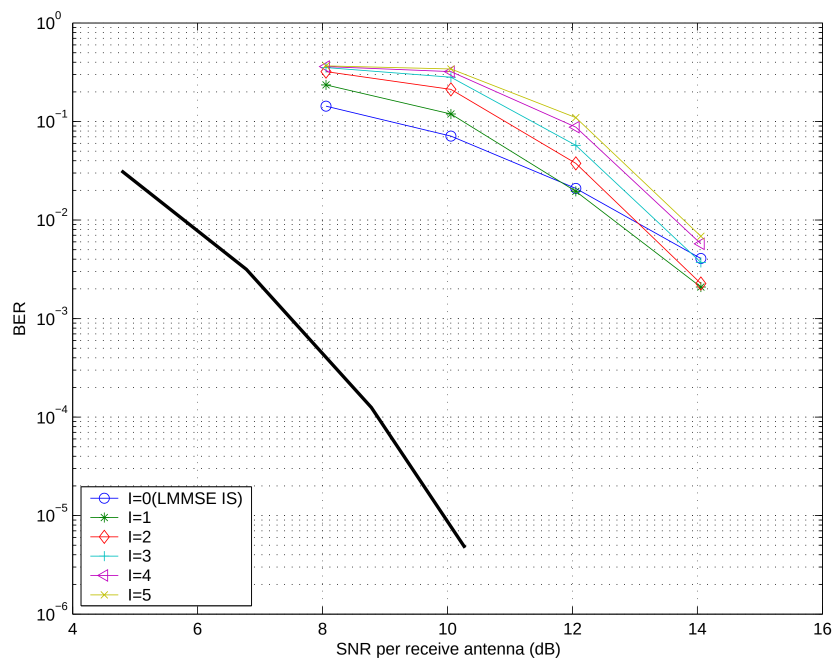


**Figure 5.6: Conventional IC-MRC turbo receiver performance for $8 \times 4$ GSTBC OFDM system. $R_c = \frac{3}{4}$ $K = 3$ CC, QPSK modulation, exact SDF, LMMSE IS initialization.**

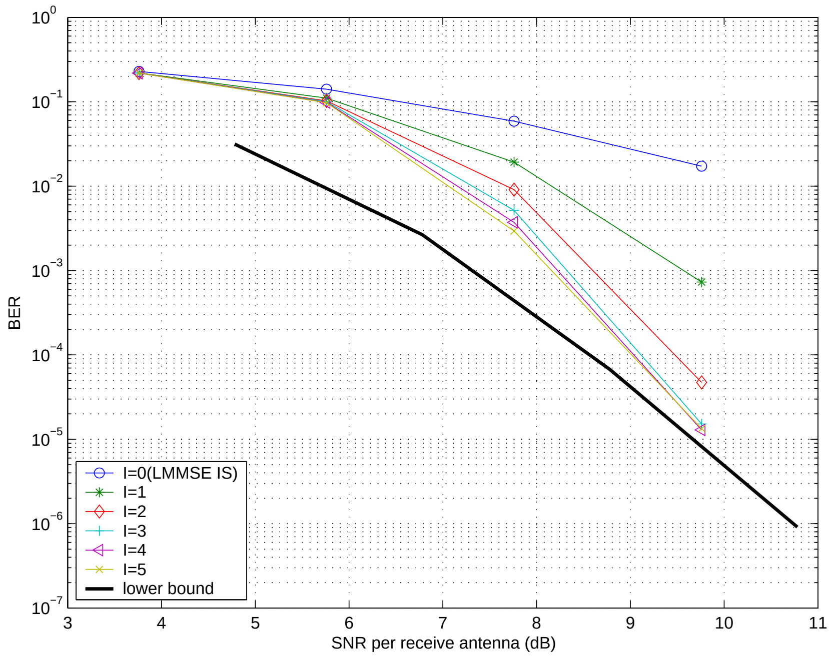


**Figure 5.7: Conventional LMMSE-IC turbo receiver performance for $8 \times 4$ GSTBC OFDM system. $R_c = \frac{3}{4}$ $K = 3$ CC, QPSK modulation, exact SDF.**

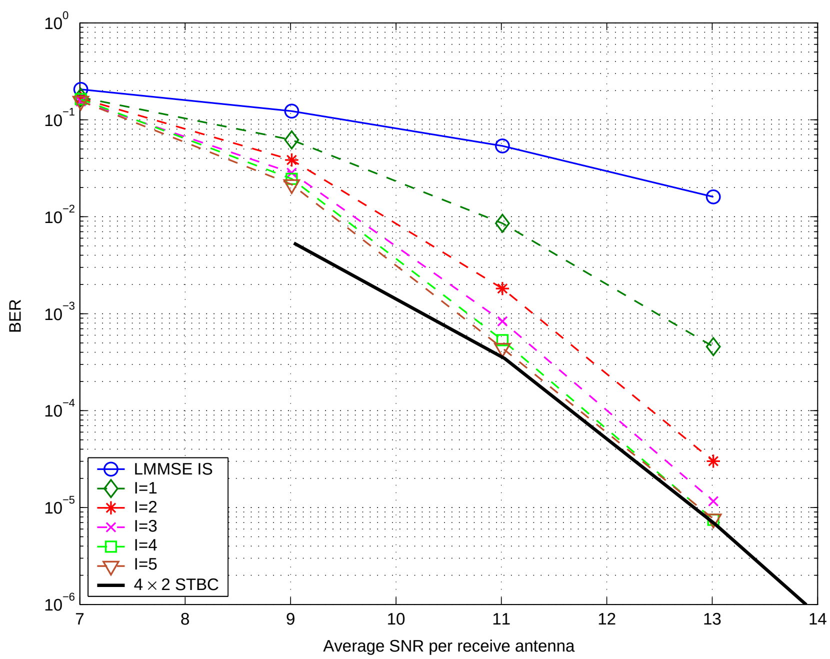


**Figure 5.8: Conventional IC-MRC turbo receiver performance for $8 \times 4$ GSTBC-OFDM. $R_c = \frac{1}{2}$ $K = 3$ CC, 16QAM modulation, exact SDF. LMMSEIS initialization.**

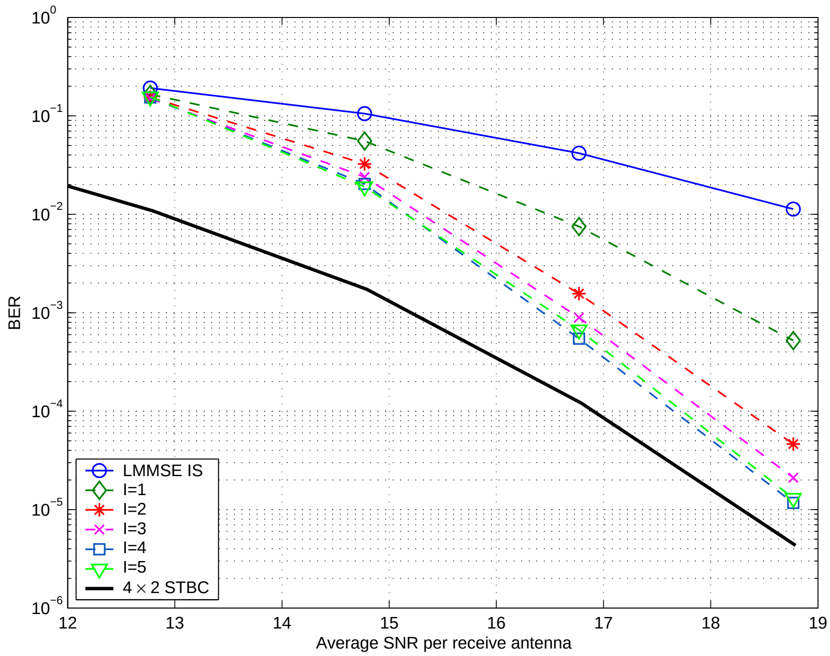


**Figure 5.9: Conventional IC-MRC turbo receiver performance for $8 \times 4$ GSTBC-OFDM. $R_c = \frac{1}{2}$ $K = 3$ CC, 64QAM modulation, exact SDF. LMMSEIS initialization.**

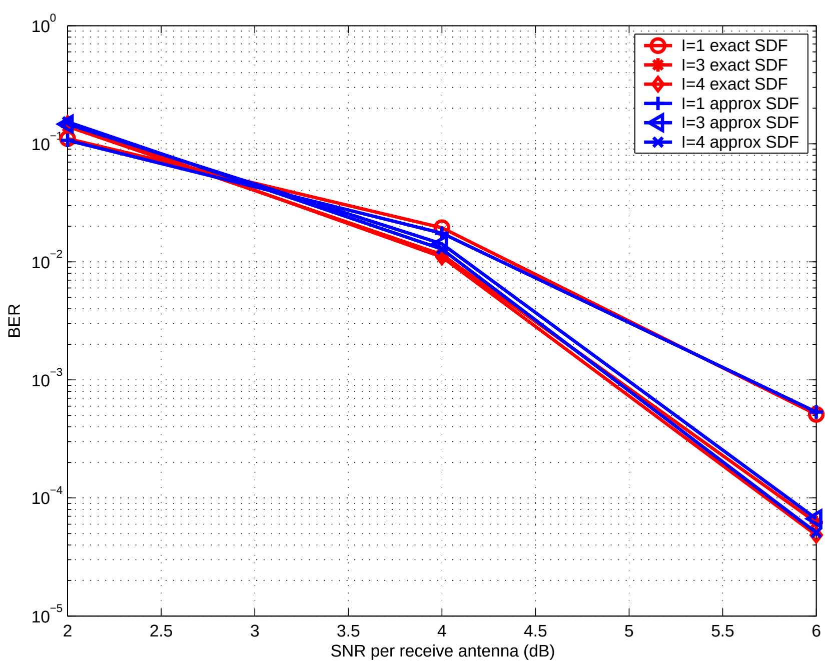


**Figure 5.10: Conventional IC-MRC turbo receiver performance for $8 \times 4$ GSTBC-OFDM. $R_c = \frac{1}{2}$ $K = 3$ CC, QPSK modulation, approximated linear SDF. LMMSEIS initialization.**

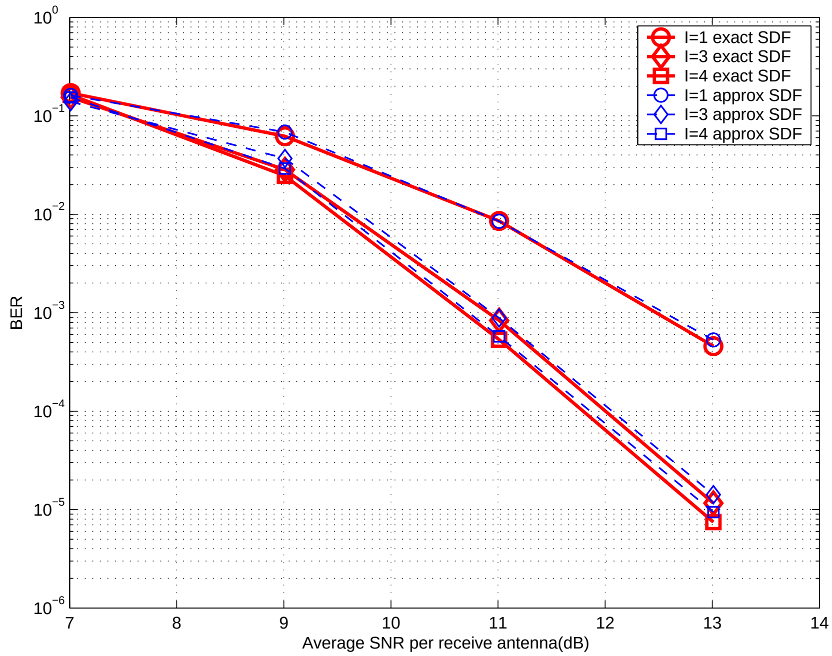


**Figure 5.11: Conventional IC-MRC turbo receiver performance for** $8 \times 4$ **GSTBC-OFDM.** $R_c = \frac{1}{2}$ $K = 3$ **CC, 16QAM modulation, approximated linear SDF. LMMSEIS initialization.**

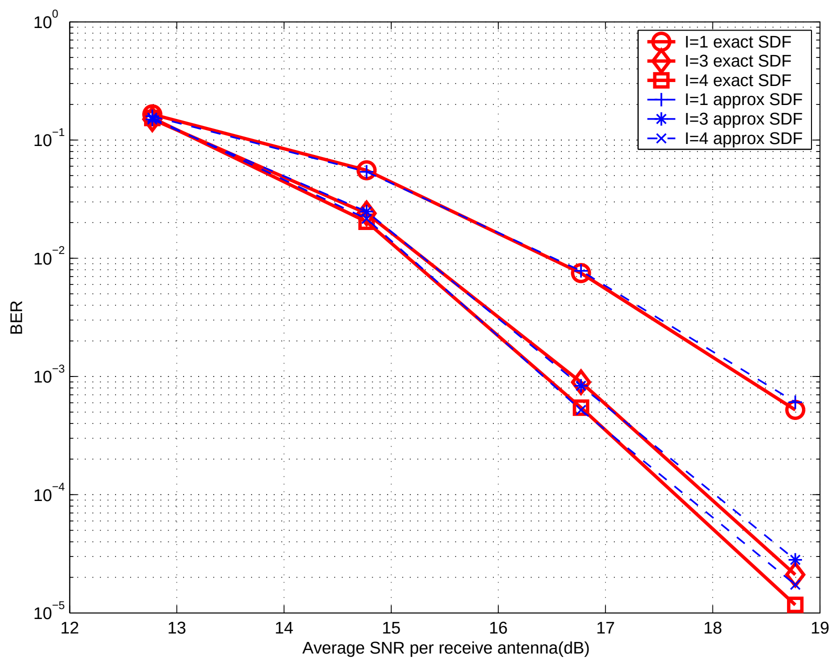


**Figure 5.12: Conventional IC-MRC turbo receiver performance for** $8 \times 4$ **GSTBC-OFDM.** $R_c = \frac{1}{2}$ $K = 3$ **CC, 64QAM modulation, approximated linear SDF. LMMSEIS initialization.**

## 5.3 The Bayesian IC-MRC Turbo Receiver

### 5.3.1 Motivation

Conventional turbo receiver makes use of the EXT from previous iteration to calculate the statistical mean as the interference estimate. When puncturing is applied to obtain higher code rate and higher spectral efficiency, the accuracy of statistical mean will be degraded, and the turbo receiver may not be able to converge to the lower bound, as exhibited in Fig. 5.6 for a $R_c = \frac{3}{4}$ QPSK GSTBC OFDM system. Therefore, a better interference estimator is desired. In this section, we propose a novel BMMSE interference estimator which employs both the EXT from the SISO decoder and the decision statistic from the soft output detector in the interference estimation. The application of complementary information improves the MSE of the estimated signals, and as a result, improves the BER and FER performance.

### 5.3.2 The Detector

The proposed Bayesian turbo receiver is depicted in Fig. 5.13. Similar to the conventional turbo receiver depicted in Fig. 5.1, a soft-output STFP detector is implemented at each iteration. The detector uses either IC-MRC or LMMSE-IC scheme. The decision statistic of the transmitted signals is then delivered to the soft-demodulator and the extrinsic bit metric values are calculated and de-interleaved, using which the SISO decoders compute the updated EXT. Different from the conventional turbo receiver in Fig. 5.1, both the EXT from the SISO decoder and the decision statistic from the detector is needed to compute the Bayesian MMSE estimate of the transmitted signals, and the Bayesian MMSE estimate is used in the STFP detector for the next iteration of IC.

The IC output for the $k$th signal at $i$th iteration is then

$$\tilde{\mathbf{R}}_{k,i} = \mathbf{R} - \sum_{\substack{p=1 \\ p\neq k}}^{n_T} \mathbf{H}_p \breve{X}_{p,i-1} = \mathbf{H}_k X_k + \tilde{\mathbf{V}}_{k,i}, \tag{5.25}$$

where $\mathbf{H}_p$ is the $p$th column of $\mathbf{H}$, $\breve{X}_{p,i-1}$ is the BMMSE estimate of the $p$th interference signal at iteration $i-1$, and $\tilde{\mathbf{V}}_{k,i}$ is the composite residual interference and white Gaussian noise, i.e.,

$$\tilde{\mathbf{V}}_{k,i} = \sum_{\substack{p=1 \\ p\neq k}}^{n_T} \mathbf{H}_p \left( X_p - \breve{X}_{p,i-1} \right) + \mathbf{V}.$$

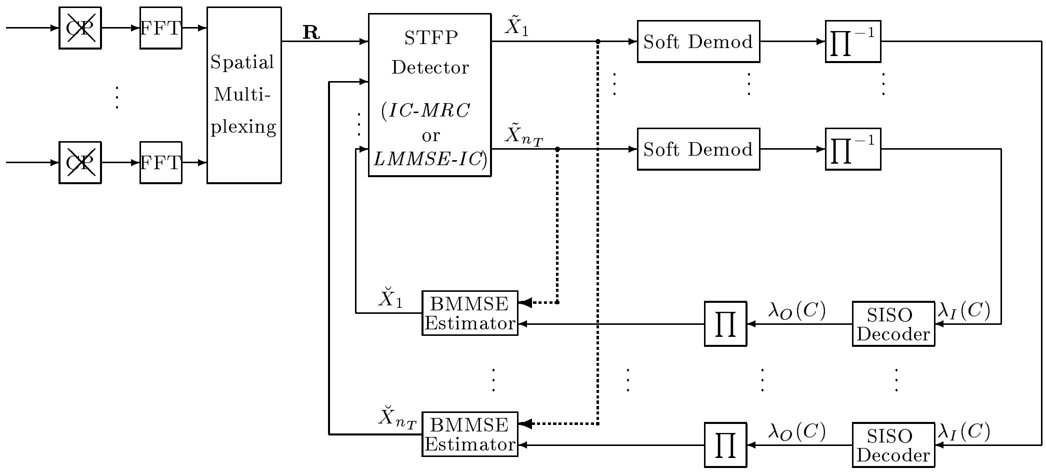


**Figure 5.13: The Bayesian turbo receiver for BICM STFP MIMO-OFDM.**

Making use of $\tilde{\mathbf{R}}_{k,i}$, the decision statistic for the $k$th signal at $i$th iteration, $\tilde{X}_{k,i}$, is thus

$$\tilde{X}_{k,i} = \mathbf{F}_{k,i}^{H}\tilde{\mathbf{R}}_{k,i}, \tag{5.26}$$

where $\mathbf{F}_{k,i}$ denotes the linear filter.

Same as the conventional turbo receiver, the two popular linear filtering schemes can be used in the proposed Bayesian turbo receiver, namely, the MF (or MRC) filter with

$$\mathbf{F}_{k,i} = \mathbf{H}_k \tag{5.27}$$

and the LMMSE filter which will be discussed in detail in Section 5.4.

In the Bayesian turbo receiver, we take a different approach in the interference estimation. We start from the optimal BMMSE estimate of the transmitted signals, and show that when expectation maximization (EM) algorithm [101] is used to reduce computation complexity, we obtain the same IC-MRC linear filtering detector, as given in (5.26) and (5.27). Different from the conventional turbo receiver which uses the statistical interference mean given in (5.7), the BMMSE estimation of the interference signals $\check{X}_{p,i-1}$ is used in IC. This leads to better accuracy, and hence better receiver performance, as shown in the later part of this chapter.

### 5.3.3 Optimal BMMSE Estimate

In Bayesian parameter estimation we assume that the parameter to be estimated is a realization of the random variable $\theta$ with an assigned prior pdf $p(\theta)$. After the data $\mathbf{y}$ are observed, the state of knowledge about $\theta$ is given by the posterior pdf $p(\theta|\mathbf{y})$. The optimal BMMSE estimator that minimizes the MSE averaged over all realizations of $\theta$ and $\mathbf{y}$, i.e., the Bayesian MSE [95], is defined as the mean of the posterior pdf,

$$\breve{\theta} = E(\theta|\mathbf{y}) = \int \theta p(\theta|\mathbf{y}) d\theta.$$

The BMMSE estimator depends on the prior knowledge as well as the observation data. If the prior knowledge is weak relative to that of data, then the estimator will ignore the prior knowledge. Otherwise, the estimator will be "biased" towards the prior mean.

If we assume perfect CSI at the receiver, the BMMSE estimate of the transmitted signals is

$$\breve{\mathbf{X}}_{BMMSE} = \mathcal{E}\{\mathbf{X}|\mathbf{R}\} = \sum_{\mathbf{X}_j \in \Omega^{n_T}} \mathbf{X}_j \; p(\mathbf{X}_j|\mathbf{R}), \tag{5.28}$$

where we assume $n_T \leq (2n_R)$ for GSTBC, and $n_T \leq n_R$ for VBLAST systems, hence the transmitted signal dimension is $n_T \times 1$ in (3.14).

Using Bayes' rule, we obtain the posterior pdf, i.e., the conditional pdf of $\mathbf{X}_j$ given $\mathbf{R}$ as

$$p(\mathbf{X}_j|\mathbf{R}) = \frac{p(\mathbf{X}_j, \mathbf{R})}{p(\mathbf{R})} = \frac{p(\mathbf{R}|\mathbf{X}_j)p(\mathbf{X}_j)}{p(\mathbf{R})} = \frac{p(\mathbf{R}|\mathbf{X}_j)p(\mathbf{X}_j)}{\sum_{\mathbf{X}_i \in \Omega^{n_T}} p(\mathbf{R}|\mathbf{X}_i)p(\mathbf{X}_i)}, \tag{5.29}$$

where

$$p(\mathbf{R}|\mathbf{X}_j) = \alpha \exp\left[-\frac{(\mathbf{R}-\mathbf{H}\mathbf{X}_j)^H(\mathbf{R}-\mathbf{H}\mathbf{X}_j)}{2\sigma^2}\right] \tag{5.30}$$

from (3.14). Here $\alpha$ is a system-dependent constant, and $\alpha = \frac{1}{(2\pi\sigma^2)^{n_R}}$ for VBLAST, and $\alpha = \frac{1}{(2\pi\sigma^2)^{2n_R}}$ for GSTBC.

We can then compute the BMMSE estimate of $\mathbf{X}$, as

$$\begin{aligned}\breve{\mathbf{X}}_{BMMSE} &= \sum_{\mathbf{X}_j \in \Omega^{n_T}} \mathbf{X}_j \; p(\mathbf{X}_j|\mathbf{R}) = \sum_{\mathbf{X}_j \in \Omega^{n_T}} \mathbf{X}_j \frac{p(\mathbf{R}|\mathbf{X}_j) \; p(\mathbf{X}_j)}{\sum_{\mathbf{X}_i \in \Omega^{n_T}} p(\mathbf{R}|\mathbf{X}_i) \; p(\mathbf{X}_i)} \\ &= \frac{\sum_{\mathbf{X}_j \in \Omega^{n_T}} \mathbf{X}_j p(\mathbf{R}|\mathbf{X}_j) \; p(\mathbf{X}_j)}{\sum_{\mathbf{X}_i \in \Omega^{n_T}} p(\mathbf{R}|\mathbf{X}_i) \; p(\mathbf{X}_i)}.\end{aligned} \tag{5.31}$$

Computation of both the numerator and the denominator of (5.31) incurs a complexity of $M^{n_T}$. Simplification is thus desired. Here we adopt the expectation step of the EM algorithm [101] to reduce the computational complexity of BMMSE estimation in (5.31).

### 5.3.4 Bayesian EM MMSE Estimate

The Bayesian EM MMSE estimate of the $k$th transmitted signal at iteration $i$, $\breve{X}_{k,i}$, is derived based on the received signal $\mathbf{R}$, and the estimate of the interference signals at iteration $i-1$, $\breve{X}_{p,i-1}, \quad p \neq k$, as

$$\breve{X}_{k,i} \stackrel{\text{def}}{=} \mathcal{E}\left\{X_k | \mathbf{R}, \overline{\breve{\mathbf{X}}}_{k,i-1}\right\} \tag{5.32}$$

where $\overline{\breve{\mathbf{X}}}_{k,i-1} = \left[\breve{X}_{1,i-1},\ \breve{X}_{2,i-1},\ \cdots, \breve{X}_{k-1,i-1},\ \breve{X}_{k+1,i-1},\ \cdots,\ \breve{X}_{n_T,i-1}\right]^T$. The bar on the vector $\breve{\mathbf{X}}_{k,i-1}$ means the exclusion of the $k$th element $\breve{X}_{k,i-1}$ from it.

Using Bayes' rule, we can further write (5.32) as

$$\begin{aligned}
\breve{X}_{k,i} &= \sum_{S_m \in \Omega} S_m\, p\left(X_k = S_m | \mathbf{R}, \overline{\breve{\mathbf{X}}}_{k,i-1}\right) \\
&= \sum_{S_m \in \Omega} S_m \frac{p\left(X_k = S_m, \mathbf{R}, \overline{\breve{\mathbf{X}}}_{k,i-1}\right)}{p\left(\mathbf{R}, \overline{\breve{\mathbf{X}}}_{k,i-1}\right)} \\
&= \sum_{S_m \in \Omega} S_m \frac{p\left(\mathbf{R} | X_k = S_m, \overline{\breve{\mathbf{X}}}_{k,i-1}\right)\, p\left(X_k = S_m, \overline{\breve{\mathbf{X}}}_{k,i-1}\right)}{\sum_{S_n \in \Omega} p\left(\mathbf{R}, X_k = S_n, \overline{\breve{\mathbf{X}}}_{k,i-1}\right)} \\
&= \frac{\sum_{S_m \in \Omega} S_m\, p\left(\mathbf{R} | X_k = S_m,\ \overline{\breve{\mathbf{X}}}_{k,i-1}\right)\, p\left(X_k = S_m\right)}{\sum_{S_n \in \Omega} p\left(\mathbf{R} | X_k = S_n,\ \overline{\breve{\mathbf{X}}}_{k,i-1}\right)\, p\left(X_k = S_n\right)},
\end{aligned} \tag{5.33}$$

under the assumption that $p\left\{X_k,\ \overline{\breve{\mathbf{X}}}_{k,i-1}\right\} = p\left\{X_k\right\} p\left\{\overline{\breve{\mathbf{X}}}_{k,i-1}\right\}$.

**Remark 1** The computational complexity in the enumerator and denominator of (5.33) is linear with the modulation size.

**Remark 2** In conventional EM receivers as [101] [44], it was assumed that all the transmitted signals are equally probable, hence (5.33) is simplified to

$$\breve{X}_{k,i} = \frac{\sum_{S_m \in \Omega} S_m \; p\left\{\mathbf{R}|X_k = S_m, \; \overline{\breve{\mathbf{X}}}_{k,i-1}\right\}}{\sum_{S_n \in \Omega} p\left\{\mathbf{R}|X_k = S_n, \; \overline{\breve{\mathbf{X}}}_{k,i-1}\right\}}. \tag{5.34}$$

In a coded system, when turbo receiver is implemented, the *a priori* probability of the transmitted signals is available from the previous iteration, as given in (5.6). We thus can use it in (5.33) to obtain the Bayesian EM MMSE estimate. This distinguishes the Bayesian EM estimate from the conventional EM estimate in [101] [44].

Before proceeding to the derivation of $\breve{X}_{k,i}$, we first prove that the BMMSE EM estimate is unbiased.

**Theorem 7.** *The Bayesian EM MMSE estimate $\breve{X}_{k,i}$ is unbiased.*

*Proof:* From (5.32), we have

$$\begin{aligned}
\mathrm{E}\left\{\breve{X}_{k,i}\right\} &= \mathrm{E}\left\{\mathrm{E}\left\{X_k|\mathbf{R}, \overline{\breve{\mathbf{X}}}_{k,i-1}\right\}\right\} \\
&= \mathrm{E}\left\{\sum_{X_k \in \Omega} X_k \; p\left(X_k|\mathbf{R}, \overline{\breve{\mathbf{X}}}_{k,i-1}\right)\right\} \\
&= \int p\left(\mathbf{R}|\overline{\breve{\mathbf{X}}}_{k,i-1}\right) \, d\mathbf{R} \sum_{X_k \in \Omega} X_k \; p\left(X_k|\mathbf{R}, \overline{\breve{\mathbf{X}}}_{k,i-1}\right) \\
&= \sum_{X_k \in \Omega} X_k \int p\left(\mathbf{R}|\overline{\breve{\mathbf{X}}}_{k,i-1}\right) \; p\left(X_k|\mathbf{R}, \overline{\breve{\mathbf{X}}}_{k,i-1}\right) \, d\mathbf{R} \\
&= \sum_{X_k \in \Omega} X_k \int p\left(X_k, \mathbf{R}|\overline{\breve{\mathbf{X}}}_{k,i-1}\right) \, d\mathbf{R} \\
&= \sum_{X_k \in \Omega} X_k p\left(X_k|\overline{\breve{\mathbf{X}}}_{k,i-1}\right) \\
&= \sum_{X_k \in \Omega} X_k p\left(X_k\right) \\
&= \mathrm{E}\left\{X_k\right\}.
\end{aligned}$$

For the modulation schemes considered here, we further have

$$\mathcal{E}\left\{\breve{X}_{k,i}\right\} = 0. \tag{5.35}$$

■

Now we proceed to derive the BMMSE estimate $\breve{X}_{k,i}$. We first assume perfect interference estimation and cancelation. In this case, we have the pdf $p\left\{\mathbf{R}|X_k,\ \overline{\breve{\mathbf{X}}}_{k,i-1}\right\}$ given in (5.30). With the result from (5.56) in Appendix B, we have the Bayesian EM MMSE estimate as

$$\breve{X}_{k,i} = \frac{\sum\limits_{S_m\in\Omega} S_m \exp\left[\dfrac{-\left\|\mathbf{H}_k S_m\right\|^2 + 2\mathrm{R}e\left(S_m^*\tilde{X}_{k,i}\right)}{2\sigma^2}\right] p\left(X_k = S_m\right)}{\sum\limits_{S_n\in\Omega} \exp\left[\dfrac{-\left\|\mathbf{H}_k S_n\right\|^2 + 2\mathrm{R}e\left(S_n^*\tilde{X}_{k,i}\right)}{2\sigma^2}\right] p\left(X_k = S_n\right)}. \tag{5.36}$$

where $\tilde{X}_{k,i}$ is the IC-MRC decision statistic, computed from (5.26) and (5.27).

Equation (5.36) implies that to obtain the Bayesian EM MMSE estimate $\breve{X}_{k,i}$, we need to implement the IC-MRC detector so as to have $\tilde{X}_{k,i}$. We also need to know the *a priori* probability $p\left\{X_k\right\}$, which are obtained from the SISO decoders, as illustrated in Fig. 5.13.

The assumption of perfect IC or zero-residual interference power, however, is unrealistic. An accurate knowledge of $p\left\{\mathbf{R}|X_k,\ \overline{\breve{\mathbf{X}}}_{k,i-1}\right\}$ is important to guarantee the estimation accuracy. To obtain that, we proceed to analyze the statistical properties of the IC-MRC signal.

### Statistics of The IC-MRC Signal

From (5.26) and (5.27), we have the IC-MRC decision statistic as

$$\tilde{X}_{k,i} = g_k X_k + \tilde{v}_{k,i}, \tag{5.37}$$

where

$$\begin{aligned} g_k &= \mathbf{H}_k^H \mathbf{H}_k, \\ \tilde{v}_{k,i} &= \mathbf{H}_k^H \tilde{\mathbf{V}}_{k,i} = \sum_{\substack{p=1 \\ p\neq k}}^{n_T} \mathbf{H}_k^H \mathbf{H}_p \left(X_p - \breve{X}_{p,i-1}\right) + \mathbf{H}_k^H \mathbf{V}. \end{aligned}$$

$\tilde{v}_{k,i}$ is Gaussian distributed from the central limit theorem. As proven in Appendix C, it has mean zero and variance

$$\varsigma^2 = \frac{1}{2}\sum_{\substack{p=1 \\ p\neq k}}^{n_T} \frac{n_R}{n_T^2}\left(1 - \left|\breve{X}_{p,i-1}\right|^2\right) + g_k\sigma^2$$

for spatially uncorrelated WSSUS UPDF and EPDF multipath channels.

We therefore have the following pdf

$$p\left\{\tilde{X}_{k,i}|X_k\right\} = \frac{1}{2\pi\varsigma^2}\ \exp\left(-\frac{\left|\tilde{X}_{k,i} - g_k X_k\right|^2}{2\varsigma^2}\right). \tag{5.38}$$

### IC-MRC Bayesian EM MMSE Estimate

With (5.38), we are ready to derive the IC-MRC Bayesian EM MMSE estimate, as

$$\begin{aligned}\breve{X}_{k,i,\mathrm{IC}-MRC} &= \mathcal{E}\left\{X_k|\tilde{X}_{k,i}\right\} \\ &= \sum_{S_m\in\Omega} S_m \frac{p\left\{\tilde{X}_{k,i}|X_k = S_m\right\} p\left\{X_k = S_m\right\}}{\sum\limits_{S_n\in\Omega} p\left\{\tilde{X}_{k,i}|X_k = S_n\right\} p\left\{X_k = S_n\right\}}.\end{aligned} \tag{5.39}$$

For BPSK modulated signals with Gray mapping rule given in Table 5.1, we have

$$\breve{X}_{k,i,\mathrm{B}PSK,IC-MRC} = \tanh\left(\frac{\lambda_{\mathbf{a}}^{i-1}(k)}{2} + \frac{g_k\tilde{X}_{k,i}}{\varsigma^2}\right). \tag{5.40}$$

For QPSK signals with the Gray mapping rule in Table 5.2, we use the *a priori* probability in (5.6) and have

$$\begin{aligned}\breve{X}_{k,i,\mathrm{QPSK,IC-MRC}} &= \frac{1}{\sqrt{2}}\left[\tanh\left(\frac{\lambda_{\mathbf{a}}^{i-1}(k,1)}{2} + \frac{\mathrm{Re}\left(g_k\tilde{X}_{k,i}\right)}{\sqrt{2}\varsigma^2}\right)\right. \\ &\left. +j\tanh\left(\frac{\lambda_{\mathbf{a}}^{i-1}(k,2)}{2} + \frac{\mathrm{Im}\left(g_k\tilde{X}_{k,i}\right)}{\sqrt{2}\varsigma^2}\right)\right].\end{aligned} \tag{5.41}$$

For 8PSK signals with Gray mapping rule of Table 5.3, its Bayesian EM MMSE estimate is calculated as

$$\begin{aligned}\breve{X}_{k,i,\mathrm{8PSK,IC-MRC}} &= \frac{a\cosh\left(\acute{x}_1\right)\sinh\left(\acute{x}_3\right) + b\cosh\left(\acute{x}_2\right)\sinh\left(\acute{x}_4\right)e^{\lambda_{\mathbf{a}}^{i-1}(k,3)}}{\cosh\left(\acute{x}_1\right)\cosh\left(\acute{x}_3\right) + \cosh\left(\acute{x}_2\right)\cosh\left(\acute{x}_4\right)e^{\lambda_{\mathbf{a}}^{i-1}(k,3)}} \\ &+j\frac{a\cosh\left(\acute{x}_4\right)\sinh\left(\acute{x}_2\right)e^{\lambda_{\mathbf{a}}^{i-1}(k,3)} + b\cosh\left(\acute{x}_3\right)\sinh\left(\acute{x}_1\right)}{\cosh\left(\acute{x}_1\right)\cosh\left(\acute{x}_3\right) + \cosh\left(\acute{x}_2\right)\cosh\left(\acute{x}_4\right)e^{\lambda_{\mathbf{a}}^{i-1}(k,3)}},\end{aligned} \tag{5.42}$$

where

$$\begin{aligned}\acute{x}_1 &= \frac{bg_k\tilde{X}_{k,i,Im}}{\varsigma^2} - \frac{\lambda_{\mathbf{a}}^{i-1}(k,1)}{2}, \\ \acute{x}_2 &= \frac{ag_k\tilde{X}_{k,i,Im}}{\varsigma^2} - \frac{\lambda_{\mathbf{a}}^{i-1}(k,1)}{2}, \\ \acute{x}_3 &= \frac{ag_k\tilde{X}_{k,i,Re}}{\varsigma^2} - \frac{\lambda_{\mathbf{a}}^{i-1}(k,2)}{2}, \\ \acute{x}_4 &= \frac{bg_k\tilde{X}_{k,i,Re}}{\varsigma^2} - \frac{\lambda_{\mathbf{a}}^{i-1}(k,2)}{2}.\end{aligned}$$

For 16QAM signals following the mapping rule in Table 5.4, we have the Bayesian EM MMSE estimate as

$$\begin{aligned}\breve{X}_{k,i,\mathrm{16QAM,IC-MRC}} &= \frac{1}{\sqrt{10}}\Bigg(f_{x,\mathrm{16QAM}}\big(\frac{\mathrm{Re}(g_k\tilde{X}_{k,i})}{\sqrt{10}\varsigma^2},\frac{g_k}{\varsigma^2},\lambda_c^{4l-3},\lambda_c^{4l-2}\big)\\ &\qquad +jf_{x,\mathrm{16QAM}}\big(\frac{\mathrm{Im}(g_k\tilde{X}_{k,i})}{\sqrt{10}\varsigma^2},\frac{\rho_l}{\varsigma^2},\lambda_c^{4l-1},\lambda_c^{4l}\big)\Bigg), \qquad (5.43)\end{aligned}$$

where

$$f_{x,\mathrm{16QAM}}(x,y,\lambda,\gamma)=\frac{3e^{-0.4y-\gamma}\sinh(3x+\frac{\lambda}{2})+\sinh(x+\frac{\lambda}{2})}{e^{-0.4y-\gamma}\cosh(3x+\frac{\lambda}{2})+\cosh(x+\frac{\lambda}{2})},$$

and for 64QAM signals following the mapping rule in Table 5.5, we have the Bayesian EM MMSE estimate as

$$\begin{aligned}\breve{X}_{k,i,\mathrm{64QAM,IC-MRC}} &= \frac{1}{\sqrt{42}}\Bigg(f_{x,\mathrm{64QAM}}\big(\frac{\mathrm{Re}(g_k\tilde{X}_{k,i})}{\sqrt{42}\varsigma^2},\frac{\rho_l}{42\varsigma^2},\lambda_c^{6l-5},\lambda_c^{6l-4},\lambda_c^{6l-3}\big)\\ &\qquad +jf_{x,\mathrm{64QAM}}\big(\frac{\mathrm{Im}(g_k\tilde{X}_{k,i})}{\sqrt{42}\varsigma^2},\frac{\rho_l}{\varsigma^2},\lambda_c^{6l-2},\lambda_c^{6l-1},\lambda_c^{6l}\big)\Bigg), \qquad (5.44)\end{aligned}$$

where

$$f_x(x,y,\lambda,\gamma,\eta)=\frac{7k_7\sinh(7x+\frac{\lambda}{2})+5k_5\sinh(5x+\frac{\lambda}{2})+3k_3\sinh(3x+\frac{\lambda}{2})+\sinh(x+\frac{\lambda}{2})}{k_7\cosh(7x+\frac{\lambda}{2})+k_5\cosh(5x+\frac{\lambda}{2})+k_3\cosh(3x+\frac{\lambda}{2})+\cosh(x+\frac{\lambda}{2})}$$

with

$$k_7=e^{-24y-\gamma},\quad k_5=e^{-12y-\gamma-\eta},\quad \text{and } k_3=e^{-4y-\eta}\ .$$

**Discussions and Remarks**

Comparing the Bayesian EM MMSE estimate in (5.40) and (5.41) with that in [44] and [101] for uncoded systems, we improve the original EM estimate by applying the estimated *a priori* probability from the SISO decoders, rather than simply assuming an equal *a priori* probability. As pointed out in [95], use of prior information will always improve the estimation accuracy.

In turbo receivers, the prior information is estimated from the SISO decoders, hence has limited accuracy especially in the low to medium SNR region. The accuracy is further degraded when punctured code is used. Therefore, when only this estimated prior information is used to calculate the statistical mean of the interference, as in conventional turbo receivers, inaccurate IC ensues. The detrimental effect can lead to performance divergence under some circumstances, e.g., when punctured code is used, as shown by one of our simulation results presented in Fig. 5.18 of Section 5.6.

In BMMSE estimation we make use of not only the estimated *a priori* information from the decoder, but also the decision statistic information from the detector. While the prior information is estimated by exploiting the correlation introduced through BICM, the IC-MRC decision statistic is obtained using the spatial domain information through STFP. The spatial domain information can effectively compensate for the estimation errors due to the *a priori* information inaccuracy, as observed from the MSE comparison depicted in Fig. 5.14 and Fig. 5.15 for a QPSK-modulated $8 \times 8$ VBLAST system. ZFIS initialization is used for the simulation of Fig. 5.14 and LMMSE IS is used for Fig. 5.15. The rate $R_c = \frac{1}{2}$ constraint length $K = 3$ CC is used. It can be seen that for ZFIS initialization, BMMSE estimation leads to MSE reduction of 14dB at iteration 3 and $E_b/N_o =$ 4dB, and 22dB at $E_b/N_o =$ 6dB. With LMMSE IS initialization, about 12dB MSE reduction is obtained at iteration 3 for both $E_b/N_o =$ 4dB and $E_b/N_o =$6dB. This improved estimation accuracy leads to significant BER and FER performance improvement, as will be shown in Section 5.6.

### 5.3.5 The Soft Demodulator

The soft demodulator calculates the updated log-likelihood ratio (LLR) of coded bits, i.e., the extrinsic metric values, for the SISO decoder using the detector output $\tilde{X}_{k,i}$, as described in (5.4). Applying results

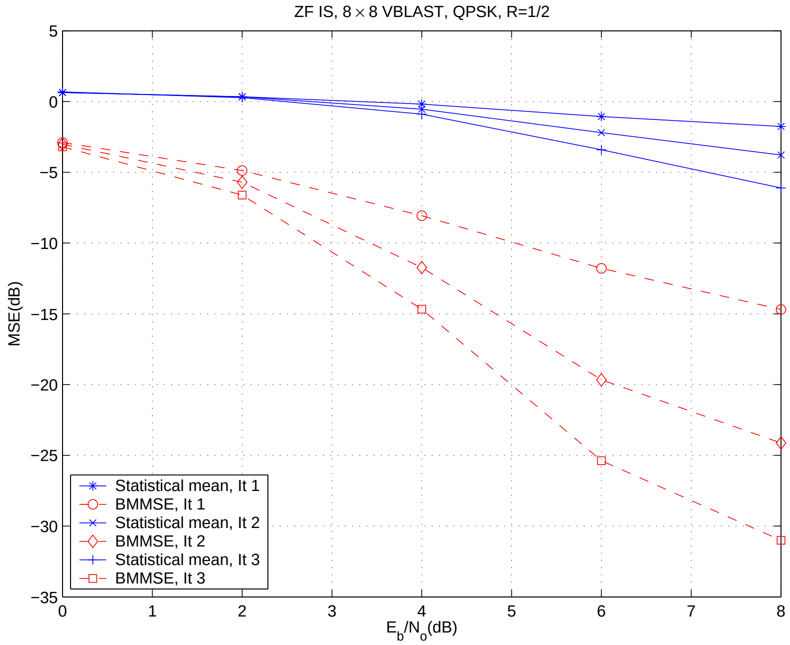


**Figure 5.14: MSE comparison between BMMSE and statistical mean interference estimation for IC-MRC turbo receiver with ZFIS initialization.** $8 \times 8$ **VBLAST, QPSK modulation,** $R_c = \frac{1}{2}$ $K = 3$ **CC.**

.

in (5.5) and (5.38), we can further decompose the extrinsic LLR $\lambda_e^i(k,l)$ as

$$
\begin{aligned}
\lambda_{\mathbf{e}}^i(k,l) &= \log \frac{\sum_{S_k \in \Omega^+} p(\tilde{X}_{k,i}|S_k) \prod_{\substack{j=1\\ j\neq l}}^{\log_2 M} \exp\left[(2c_k^j - 1)\frac{\lambda_{\mathbf{a}}^{i-1}(k,j)}{2}\right]}{\sum_{S_k \in \Omega^-} p(\tilde{X}_{k,i}|S_k) \prod_{\substack{j=1\\ j\neq l}}^{\log_2 M} \exp\left[(2c_k^j - 1)\frac{\lambda_{\mathbf{a}}^{i-1}(k,j)}{2}\right]} \\
&= \log \frac{\sum_{S_k \in \Omega^+} \exp\left(-\frac{\left|\tilde{X}_{k,i} - g_k S_k\right|^2}{2\varsigma^2}\right) \prod_{\substack{j=1\\ j\neq l}}^{\log_2 M} \exp\left[(2c_k^j - 1)\frac{\lambda_{\mathbf{a}}^{i-1}(k,j)}{2}\right]}{\sum_{S_k \in \Omega^-} \exp\left(-\frac{\left|\tilde{X}_{k,i} - g_k S_k\right|^2}{2\varsigma^2}\right) \prod_{\substack{j=1\\ j\neq l}}^{\log_2 M} \exp\left[(2c_k^j - 1)\frac{\lambda_{\mathbf{a}}^{i-1}(k,j)}{2}\right]} \\
&= \log \frac{\sum_{S_k \in \Omega^+} \exp\left[-\frac{\left|\tilde{X}_{k,i} - g_k S_k\right|^2}{2\varsigma^2} + \sum_{\substack{j=1\\ j\neq l}}^{\log_2 M} (2c_k^j - 1)\frac{\lambda_{\mathbf{a}}^{i-1}(k,j)}{2}\right]}{\sum_{S_k \in \Omega^-} \exp\left[-\frac{\left|\tilde{X}_{k,i} - g_k S_k\right|^2}{2\varsigma^2} + \sum_{\substack{j=1\\ j\neq l}}^{\log_2 M} (2c_k^j - 1)\frac{\lambda_{\mathbf{a}}^{i-1}(k,j)}{2}\right]},
\end{aligned}
$$

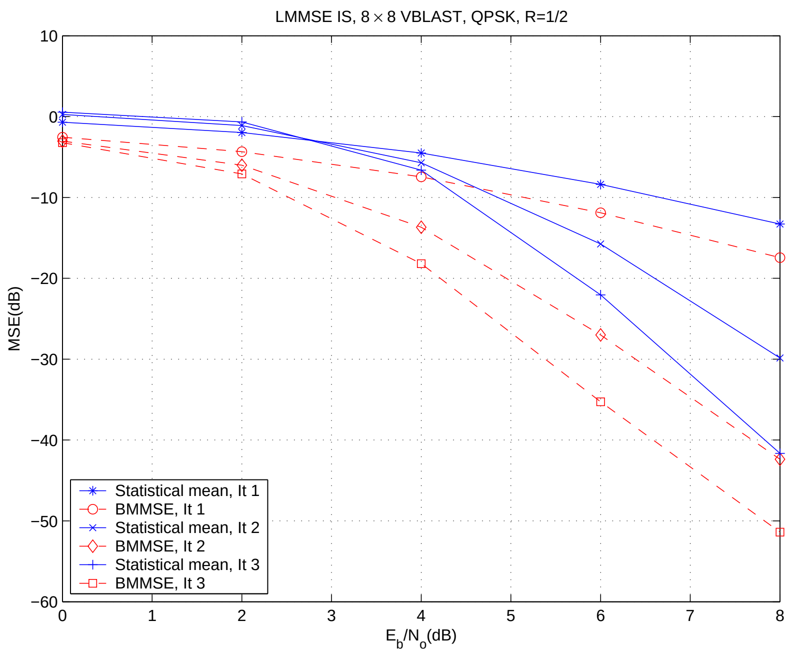


**Figure 5.15: MSE comparison between BMMSE and statistical mean interference estimation for IC-MRC turbo receiver with LMMSEIS initialization.** $8 \times 8$ **VBLAST, QPSK modulation,** $R_c = \frac{1}{2}$ $K = 3$ **CC.**

.

we hence have

$$\begin{aligned}\lambda_{\mathbf{e}}^{i}(k,l) \;&=\; \max{}^{*}_{S_k \in \Omega^{+}} \left[ -\frac{\left|\tilde{X}_{k,i} - g_k S_k\right|^2}{2\varsigma^2} + \sum_{\substack{j=1\\ j\neq l}}^{\log_2 M} (2c_k^j - 1)\frac{\lambda_{\mathbf{a}}^{i-1}(k,j)}{2} \right] \\ &\quad - \max{}^{*}_{S_k \in \Omega^{-}} \left[ -\frac{\left|\tilde{X}_{k,i} - g_k S_k\right|^2}{2\varsigma^2} + \sum_{\substack{j=1\\ j\neq l}}^{\log_2 M} (2c_k^j - 1)\frac{\lambda_{\mathbf{a}}^{i-1}(k,j)}{2} \right], \quad (5.45)\end{aligned}$$

where the $\max^*(\cdot)$ function is defined as

$$a = \max{}^{*} \equiv \log \left[ \sum_{i=1}^{L} \exp\left(a_i\right) \right],$$

and

$$\max{}^{*}(a_1, a_2) = \log\left(e^{a_1} + e^{a_2}\right) = \max(a_1, a_2) + \log\left(1 + e^{-|a_1 - a_2|}\right).$$

## 5.4 The Bayesian LMMSE-IC Turbo Receiver

Similar to conventional LMMSE-IC turbo receivers in [85] and [89], LMMSE filtering can be applied in (5.26) to further suppress the residual interference. Verdú and Poor have proved in [98] that the LMMSE filter can be modeled by an equivalent AWGN channel. We can therefore obtain the BMMSE estimate of the signal for the LMMSE-IC turbo receiver in a very straight-forward manner. The LMMSE filter $\mathbf{F}_{k,i}$ minimizes the MSE between the transmitted signal $X_k$ and the filter output $\tilde{X}_{k,i} = \mathbf{F}_{k,i}^H \tilde{\mathbf{R}}_{k,i}$, and is obtained from the Wiener-Hopf equation as

$$\mathbf{F}_{k,i} = \mathcal{E}\left\{\tilde{\mathbf{R}}_{k,i}\tilde{\mathbf{R}}_{k,i}^H\right\}^{-1} \mathrm{E}\left\{X_k^* \tilde{\mathbf{R}}_{k,i}\right\} = \left(\mathbf{H}\boldsymbol{\Gamma}_{k,i}\mathbf{H}^H + 2\sigma^2\mathbf{I}\right)^{-1}\mathbf{H}_k, \tag{5.46}$$

where $\boldsymbol{\Gamma}_{k,i}$ is a diagonal matrix with elements

$$\gamma_{p,i} = \begin{cases} 1, & p = k \\ \sum_{X_p \in \Omega} p\{X_p\}|X_p|^2 + \left|\breve{X}_{p,i-1}\right|^2 - \breve{X}_{p,i-1}^* \hat{X}_{p,i} - \breve{X}_{p,i-1}\hat{X}_{p,i}^*, & p \neq k, \end{cases} \tag{5.47}$$

and $\hat{X}_{p,i}$ and $\breve{X}_{p,i}$ being the statistical mean and BMMSE estimate of $X_p$, respectively. Please take note that elements $\gamma_{p\neq k,i}$ are computed differently from the conventional LMMSE-IC turbo receivers, which is due to the use of Bayesian estimate in the IC process.

The LMMSE-IC decision statistic is thus

$$\tilde{X}_{k,i} = \mathbf{F}_{k,i}^H \tilde{\mathbf{R}}_{k,i} = \mu_{k,i} X_k + \eta_{k,i}, \tag{5.48}$$

which is an equivalent AWGN model with [98]

$$\mu_{k,i} = \mathcal{E}\left\{\tilde{X}_{k,i} X_k^*\right\} = \left\{\mathbf{H}^H\left[\mathbf{H}\boldsymbol{\Gamma}_{k,i}\mathbf{H}^H + 2\sigma^2\mathbf{I}\right]^{-1}\mathbf{H}\right\}_{kk}, \tag{5.49}$$

$$\nu_{k,i}^2 = \mathrm{V}ar\left\{\tilde{X}_{k,i}\right\} = \mu_{k,i} - \mu_{k,i}^2, \tag{5.50}$$

and $\{\cdot\}_{ij}$ denotes the element at $i$th row and $j$th column.

With the Gaussian model developed in (5.48) - (5.50), we can easily obtain the BMMSE estimate for BPSK signals as

$$\breve{X}_{k,i,\mathrm{B}PSK,LMMSE-IC} = \tanh\left(\frac{\lambda_{\mathbf{a}}^{i-1}(k)}{2} + \frac{\mu_{k,i}\tilde{X}_{k,i}}{\nu_{k,i}^2}\right), \tag{5.51}$$

and for QPSK signals as

$$\breve{X}_{k,i,QPSK,LMMSE-IC} = \frac{1}{\sqrt{2}}\left[\tanh\left(\frac{\lambda_{\mathbf{a}}^{i-1}(k,1)}{2}+\frac{\mathrm{Re}\left(\mu_{k,i}\tilde{X}_{k,i}\right)}{\sqrt{2}\nu_{k,i}^{2}}\right)\right.$$
$$\left.+j\tanh\left(\frac{\lambda_{\mathbf{a}}^{i-1}(k,2)}{2}+\frac{\mathrm{Im}\left(\mu_{k,i}\tilde{X}_{k,i}\right)}{\sqrt{2}\nu_{k,i}^{2}}\right)\right]. \tag{5.52}$$

For 8PSK signals with Gray mapping rule of Table 5.3, its Bayesian EM MMSE estimate is calculated as

$$\breve{X}_{k,i,\mathrm{8PSK,LMMSE-IC}} = \frac{a\cosh(\acute{x}_1)\sinh(\acute{x}_3)+b\cosh(\acute{x}_2)\sinh(\acute{x}_4)\,e^{\lambda_{\mathbf{a}}^{i-1}(k,3)}}{\cosh(\acute{x}_1)\cosh(\acute{x}_3)+\cosh(\acute{x}_2)\cosh(\acute{x}_4)\,e^{\lambda_{\mathbf{a}}^{i-1}(k,3)}}$$
$$+j\frac{a\cosh(\acute{x}_4)\sinh(\acute{x}_2)\,e^{\lambda_{\mathbf{a}}^{i-1}(k,3)}+b\cosh(\acute{x}_3)\sinh(\acute{x}_1)}{\cosh(\acute{x}_1)\cosh(\acute{x}_3)+\cosh(\acute{x}_2)\cosh(\acute{x}_4)\,e^{\lambda_{\mathbf{a}}^{i-1}(k,3)}}, \tag{5.53}$$

where

$$\begin{aligned}
\acute{x}_1 &= \frac{b\mu_{k,i}\tilde{X}_{k,i,Im}}{\nu_{k,i}^2}-\frac{\lambda_{\mathbf{a}}^{i-1}(k,1)}{2},\\
\acute{x}_2 &= \frac{a\mu_{k,i}\tilde{X}_{k,i,Im}}{\nu_{k,i}^2}-\frac{\lambda_{\mathbf{a}}^{i-1}(k,1)}{2},\\
\acute{x}_3 &= \frac{a\mu_{k,i}\tilde{X}_{k,i,Re}}{\nu_{k,i}^2}-\frac{\lambda_{\mathbf{a}}^{i-1}(k,2)}{2},\\
\acute{x}_4 &= \frac{b\mu_{k,i}\tilde{X}_{k,i,Re}}{\nu_{k,i}^2}-\frac{\lambda_{\mathbf{a}}^{i-1}(k,2)}{2}.
\end{aligned}$$

For 16QAM signals following the mapping rule in Table 5.4, we have the Bayesian EM MMSE estimate as

$$\breve{X}_{k,i,\mathrm{16QAM,LMMSE-IC}} = \frac{1}{\sqrt{10}}\left(f_{x,\mathrm{16QAM}}\left(\frac{\mathrm{Re}(\mu_{k,i}\tilde{X}_{k,i})}{\sqrt{10}\nu_{k,i}^2},\frac{\mu_{k,i}}{\nu_{k,i}^2},\lambda_c^{4l-3},\lambda_c^{4l-2}\right)\right.$$
$$\left.+jf_{x,\mathrm{16QAM}}\left(\frac{\mathrm{Im}(\mu_{k,i}\tilde{X}_{k,i})}{\sqrt{10}\nu_{k,i}^2},\frac{\rho_l}{\nu_{k,i}^2},\lambda_c^{4l-1},\lambda_c^{4l}\right)\right), \tag{5.54}$$

where

$$f_{x,\mathrm{16QAM}}(x,y,\lambda,\gamma)=\frac{3e^{-0.4y-\gamma}\sinh(3x+\frac{\lambda}{2})+\sinh(x+\frac{\lambda}{2})}{e^{-0.4y-\gamma}\cosh(3x+\frac{\lambda}{2})+\cosh(x+\frac{\lambda}{2})},$$

and for 64QAM signals following the mapping rule in Table 5.5, we have the Bayesian EM MMSE estimate as

$$\breve{X}_{k,i,\text{64QAM,LMMSE-IC}} = \frac{1}{\sqrt{42}}\left(f_{x,\text{64QAM}}\left(\frac{\text{Re}(\mu_{k,i}\tilde{X}_{k,i})}{\sqrt{42}\nu_{k,i}^2}, \frac{\rho_l}{42\nu_{k,i}^2}, \lambda_c^{6l-5}, \lambda_c^{6l-4}, \lambda_c^{6l-3}\right)\right.$$
$$\left.+jf_{x,\text{64QAM}}\left(\frac{\text{Im}(\mu_{k,i}\tilde{X}_{k,i})}{\sqrt{42}\nu_{k,i}^2}, \frac{\rho_l}{\nu_{k,i}^2}, \lambda_c^{6l-2}, \lambda_c^{6l-1}, \lambda_c^{6l}\right)\right), \quad (5.55)$$

where

$$f_x(x,y,\lambda,\gamma,\eta) = \frac{7k_7\sinh(7x+\frac{\lambda}{2}) + 5k_5\sinh(5x+\frac{\lambda}{2}) + 3k_3\sinh(3x+\frac{\lambda}{2}) + \sinh(x+\frac{\lambda}{2})}{k_7\cosh(7x+\frac{\lambda}{2}) + k_5\cosh(5x+\frac{\lambda}{2}) + k_3\cosh(3x+\frac{\lambda}{2}) + \cosh(x+\frac{\lambda}{2})}$$

with

$$k_7 = e^{-24y-\gamma}, \quad k_5 = e^{-12y-\gamma-\eta}, \quad k_3 = e^{-4y-\eta} \ .$$

## 5.5 SDF Simplification in Bayesian EM Estimate

Incorporation of both the SISO decoder EXT and the soft output detector output in the interference estimation will improve the estimation accuracy, as shown in the MSE performance comparison in Fig. 5.15 and Fig. 5.14, and the BER and FER performance comparison presented in Section 5.6. However, as more variables are included in the SDF, the computational complexity is higher than the conventional turbo receiver. In this section, we will discuss the possible simplification of SDF's in Bayesian EM estimate.

For BPSK and QPSK signals, the SDF's are still a hyperbolic function. Therefore, the clip function given in (5.14) can be applied. For 8PSK, 16QAM and 64QAM signals, the corresponding simplified SDF's need to be re-derived. We propose this as one possible future work for the Bayesian turbo receiver study.

## 5.6 BER and FER Performance

In this section, we present the BER and FER performance of the Bayesian turbo receivers. Again the rate $R_c = \frac{1}{2}$ constraint length $K = 3$ CC is used as the mother code, and puncturing is applied to generate the desired coding rate according to the puncturing pattern given in [99] and [100]. Uniform power delay profiles with sixteen i.i.d. complex Gaussian taps are used for the spatial channels corresponding to each transmit and receive antenna pair, which are assumed uncorrelated in spatial domain. The FFT size is set to 64, and 48 are assigned as data subcarriers. Coding frame length of 96 is used for $R_c = \frac{1}{2}$ QPSK, and 216 is used for $R_c = \frac{3}{4}$ 8PSK, in each parallel streams.

In Fig. 5.16 and Fig. 5.17, we depict the BER and FER performance of IC-MRC Bayesian MMSE receiver for an $8 \times 4$ GSTBC system with $R_c = \frac{1}{2}$ and QPSK modulation. ZFIS is used for initialization. For comparison, the conventional IC-MRC turbo receiver performance using EXT and the lower bound are also depicted in the figure. Several observations can be made from the two figures. First, superior performance is obtained by the Bayesian turbo receiver. Its *second iteration* performance is better than the *fifth iteration* of the conventional receiver at low to medium SNR values, and the same as the conventional one at high SNR. This is because of the improved accuracy in BMMSE interference estimation, as shown in Section 5.3.

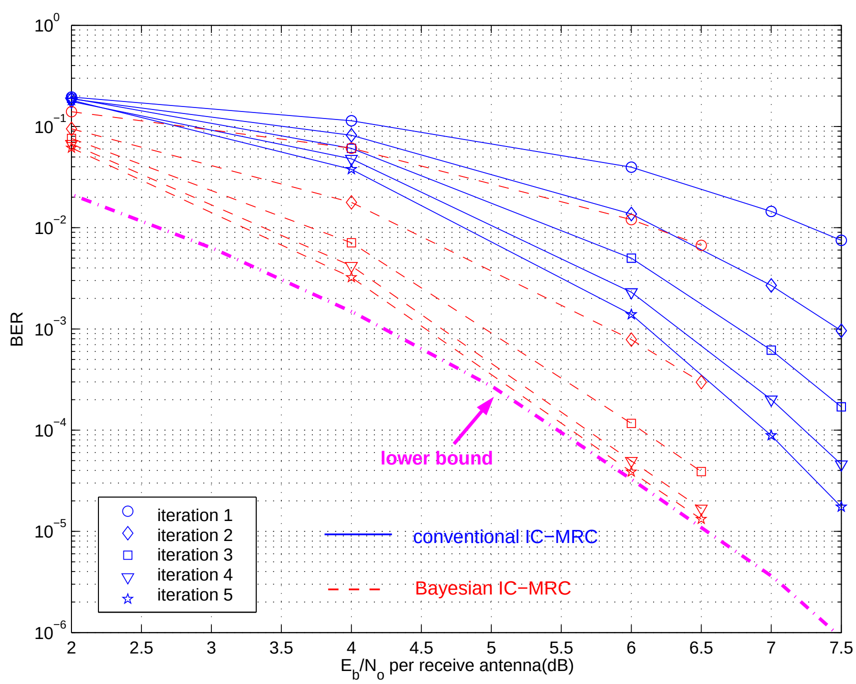


**Figure 5.16: BER performance of Bayesian IC-MRC receiver, $8 \times 4$ GSTBC, QPSK, $R_c = \frac{1}{2}$ $K = 3$ CC.**

With the same number of iterations implemented for both the Bayesian and the conventional turbo receivers, an SNR gain of 1.2 dB can be achieved from the Bayesian receiver at iteration five, at BER =

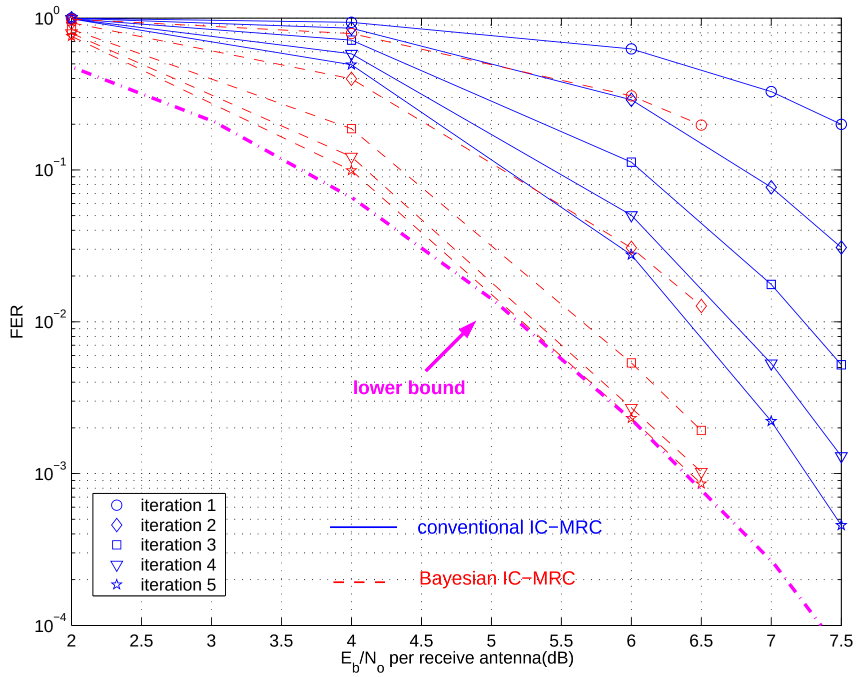


**Figure 5.17: FER performance of Bayesian IC-MRC receiver, $8 \times 4$ GSTBC, QPSK, $R_c = \frac{1}{2}$ $K = 3$ CC.**

$10^{-5}$ and FER = $10^{-2}$. Convergence of the Bayesian receiver to the lower bound appears at SNR = 4.5 dB, corresponding to BER = $6 \times 10^{-4}$ and FER = $3 \times 10^{-2}$, while the conventional receiver does not show obvious convergence in the range of our simulation setups.

Performance advantage of the Bayesian receiver is more obvious when punctured code is used in the system, as illustrated by the BER simulation results in Fig. 5.18 for a $R_c = \frac{3}{4}$ QPSK $8 \times 4$ GSTBC system. In this figure, we present four simulation results, namely, conventional IC-MRC receiver with both ZFIS and LMMSE IS initialization, and Bayesian IC-MRC receiver with ZFIS and LMMSE IS initialization. For each of them, we show the BER at iterations 1, 3, and 5. From the figure, we can see clearly the divergence behaviors of the conventional receivers. LMMSE IS initialization improves the performance, but it can not solve the divergence problem. We believe this is due to puncturing in the CC that degrades the accuracy in the EXT. The Bayesian receivers, with the compensation of the detector's decision statistic, however, can very well avoid the performance divergence. Furthermore, they achieves 5dB gain over the conventional ones with both ZFIS and LMMSE IS initialization, and more importantly,

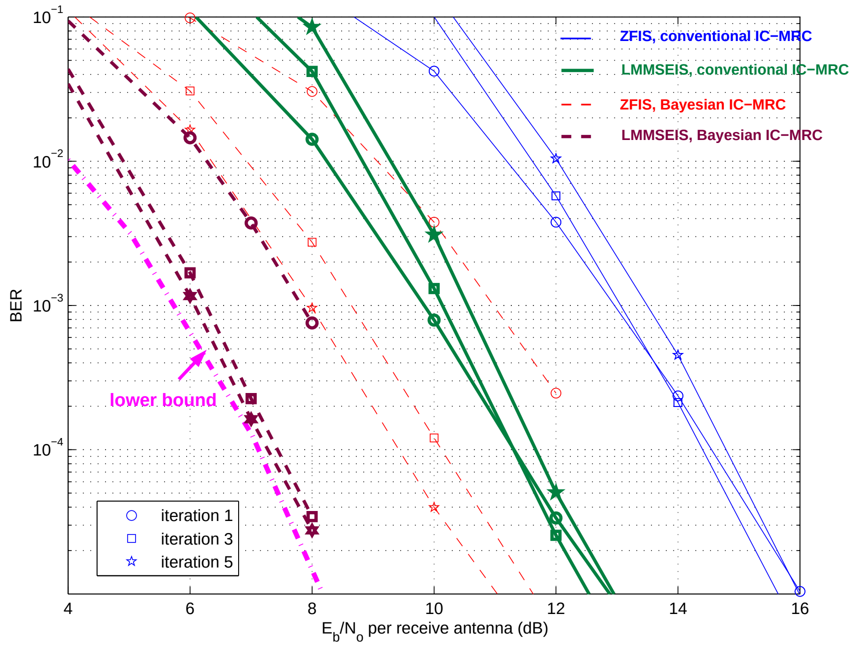


**Figure 5.18: BER performance comparison of Bayesian IC-MRC and conventional IC-MRC receivers, ZFIS and LMMSE IS,** $8 \times 4$ **GSTBC, QPSK,** $R_c = \frac{3}{4}$ **K=3 CC.**

the Bayesian receiver with LMMSE IS initialization converges to the lower bound at BER = $10^{-3}$ and SNR = 6 dB. The additional complexity to achieve these significant gains is only the summation of decision statistic and the *a priori* information in the hyperbolic tangent function, as shown in (5.40) and (5.41).

The BER and FER performances of the Bayesian LMMSE-IC receiver are depicted in Fig. 5.19 and Fig. 5.20, respectively for an $8 \times 8$ VBLAST system with $R_c = \frac{3}{4}$ 8PSK. Similar to the IC-MRC Bayesian receivers, its second iteration performance is better than the fifth iteration conventional receiver at low to medium SNR values due to the improved accuracy in the interference estimation, and approaches the conventional receiver performance at high SNR values due to the dominance of the a priori information in the Bayesian estimate. For all the five iterations presented in the figures, the Bayesian receiver achieves at least 1dB gain over the conventional one.

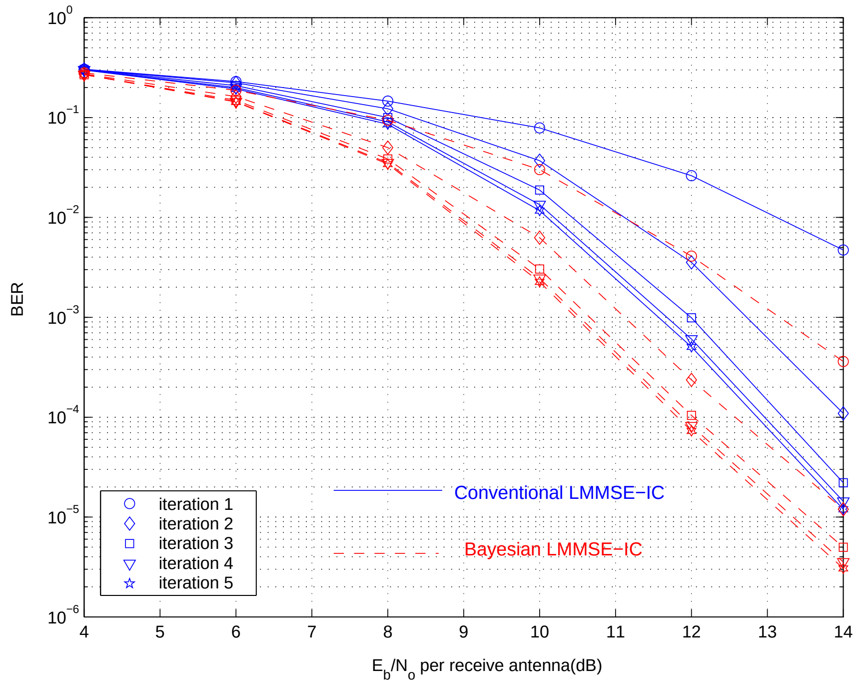


**Figure 5.19: BER performance of Bayesian LMMSE-IC receiver, $8 \times 8$ VBLAST, 8PSK, $R_c = \frac{3}{4}$ K=3 CC.**

## 5.7 Conclusions

We presented our study on turbo receivers for coded MIMO-OFDM systems. In order to reduce the complexity of SDF in conventional turbo receivers, Maclaurin series were used to derive the simplified linear SDF's for the popular MPSK and M-QAM signals. Simulation results show negligible performance degradation from the decision function simplification.

We also proposed a new class of Bayesian MMSE turbo receivers for coded MIMO-OFDM systems. Using EM algorithm, we derive the Bayesian MMSE estimate of the transmitted signals and show that it is a function of both the linear detector decision statistic and the extrinsic information from the soft-input soft-output decoder. The EXT and decision statistic represent information from two different domains, one from coding domain and the other from the interference domain. They are hence not repetitive but complementary in interference estimation. The Bayesian MMSE estimate effectively compensates for the inaccuracy experienced by the statistical mean interference estimation using *only* the extrinsic in-

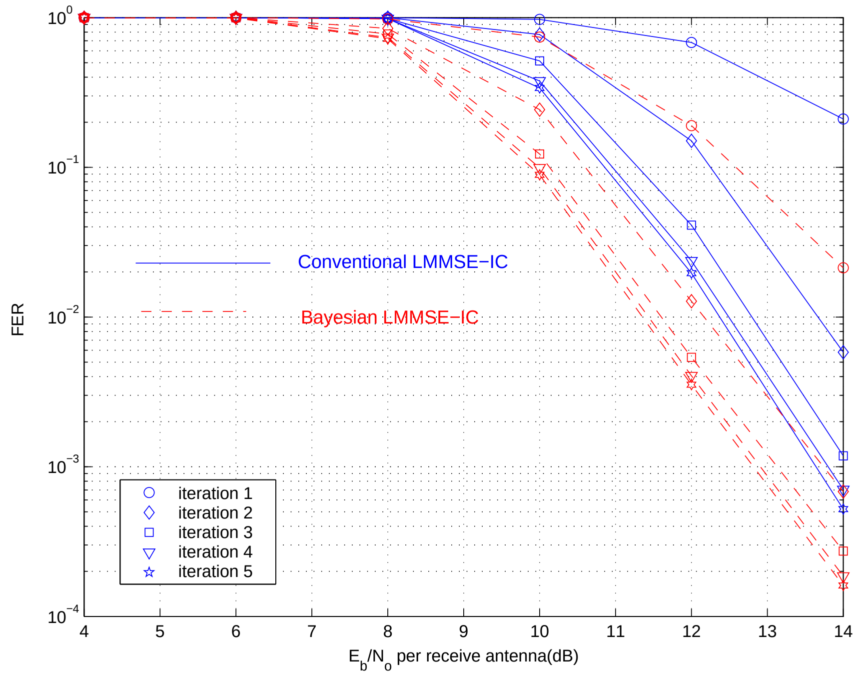


**Figure 5.20: FER performance of Bayesian LMMSE-IC receiver, $8 \times 8$ VBLAST, 8PSK, $R_c = \frac{3}{4}$ K=3 CC.**

formation in conventional turbo receivers. This contributes to the fewer number of iterations needed to achieve convergence, and the SNR gains at same BER and FER performances.

The incorporation of more variables in the interference estimate of Bayesian turbo receiver, however, does not introduce much additional complexity for the BPSK and QPSK modulation signals. The simplified linear SDF used in conventional turbo receivers can also be applied in the Bayesian turbo receivers in a straightforward manner for these two modulation schemes. For other modulation signals, e.g., 8PSK, 16QAm and 64QAM, more complexities will be incurred in getting the Bayesian EM estimate. Simplification of the decision functions are therefore desired.

In the next chapter, we will present the EXIT chart analysis of the Bayesian turbo receivers.

# Appendix A: Detailed Derivation of (5.10), (5.11) and (5.12)

From (5.7), we can calculate $\hat{X}_{m,i}$ following the 8PSK mapping table given in Table 5.3 as

$$\begin{aligned}\hat{X}_{m,i} \;=\; & (a+jb)P(S_{m,I}=a+jb)+(b+ja)P(S_{m,I}=b+ja)\\ & +(-b+ja)P(S_{m,I}=-b+ja)+(-a+jb)P(S_{m,I}=-a+jb)\\ & +(-a-jb)P(S_{m,I}=-a-jb)+(-b-ja)P(S_{m,I}=-b-ja)\\ & +(b-ja)P(S_{m,I}=b-ja)+(a-jb)P(S_{m,I}=a-jb).\end{aligned}$$

Hence we have the real part of $\hat{X}_{m,i}$ calculated as

$$\begin{aligned}&\hat{X}_{m,i,\ Re} =\\ &\quad a\frac{e^{\frac{-\lambda_{\mathbf{a}}(m,1)-\lambda_{\mathbf{a}}(m,2)-\lambda_{\mathbf{a}}(m,3)}{2}}-e^{\frac{-\lambda_{\mathbf{a}}(m,1)+\lambda_{\mathbf{a}}(m,2)-\lambda_{\mathbf{a}}(m,3)}{2}}-e^{\frac{\lambda_{\mathbf{a}}(m,1)+\lambda_{\mathbf{a}}(m,2)-\lambda_{\mathbf{a}}(m,3)}{2}}+e^{\frac{\lambda_{\mathbf{a}}(m,1)-\lambda_{\mathbf{a}}(m,2)-\lambda_{\mathbf{a}}(m,3)}{2}}}{\left(e^{\frac{\lambda_{\mathbf{a}}(m,1)}{2}}+e^{-\frac{\lambda_{\mathbf{a}}(m,1)}{2}}\right)\left(e^{\frac{\lambda_{\mathbf{a}}(m,2)}{2}}+e^{-\frac{\lambda_{\mathbf{a}}(m,2)}{2}}\right)\left(e^{\frac{\lambda_{\mathbf{a}}(m,3)}{2}}+e^{-\frac{\lambda_{\mathbf{a}}(m,3)}{2}}\right)}\\ &\quad +b\frac{e^{\frac{-\lambda_{\mathbf{a}}(m,1)-\lambda_{\mathbf{a}}(m,2)+\lambda_{\mathbf{a}}(m,3)}{2}}-e^{\frac{-\lambda_{\mathbf{a}}(m,1)+\lambda_{\mathbf{a}}(m,2)+\lambda_{\mathbf{a}}(m,3)}{2}}-e^{\frac{\lambda_{\mathbf{a}}(m,1)+\lambda_{\mathbf{a}}(m,2)+\lambda_{\mathbf{a}}(m,3)}{2}}+e^{\frac{\lambda_{\mathbf{a}}(m,1)-\lambda_{\mathbf{a}}(m,2)+\lambda_{\mathbf{a}}(m,3)}{2}}}{\left(e^{\frac{\lambda_{\mathbf{a}}(m,1)}{2}}+e^{-\frac{\lambda_{\mathbf{a}}(m,1)}{2}}\right)\left(e^{\frac{\lambda_{\mathbf{a}}(m,2)}{2}}+e^{-\frac{\lambda_{\mathbf{a}}(m,2)}{2}}\right)\left(e^{\frac{\lambda_{\mathbf{a}}(m,3)}{2}}+e^{-\frac{\lambda_{\mathbf{a}}(m,3)}{2}}\right)}\\ &= \frac{ae^{\frac{-\lambda_{\mathbf{a}}(m,3)}{2}}}{e^{\frac{\lambda_{\mathbf{a}}(m,3)}{2}}+e^{-\frac{\lambda_{\mathbf{a}}(m,3)}{2}}}\tanh\left(-\frac{\lambda_{\mathbf{a}}(m,2)}{2}\right)+\frac{be^{\frac{\lambda_{\mathbf{a}}(m,3)}{2}}}{e^{\frac{\lambda_{\mathbf{a}}(m,3)}{2}}+e^{-\frac{\lambda_{\mathbf{a}}(m,3)}{2}}}\tanh\left(-\frac{\lambda_{\mathbf{a}}(m,2)}{2}\right)\\ &= \frac{ae^{\frac{-\lambda_{\mathbf{a}}(m,3)}{2}}+be^{\frac{\lambda_{\mathbf{a}}(m,3)}{2}}}{e^{\frac{\lambda_{\mathbf{a}}(m,3)}{2}}+e^{-\frac{\lambda_{\mathbf{a}}(m,3)}{2}}}\tanh\left(-\frac{\lambda_{\mathbf{a}}(m,2)}{2}\right),\end{aligned}$$

and the imaginary part of $\hat{X}_{m,i}$ calculated as

$$\begin{aligned}&\hat{X}_{m,i,\ Im} =\\ &\quad b\frac{e^{\frac{-\lambda_{\mathbf{a}}(m,1)-\lambda_{\mathbf{a}}(m,2)-\lambda_{\mathbf{a}}(m,3)}{2}}+e^{\frac{-\lambda_{\mathbf{a}}(m,1)+\lambda_{\mathbf{a}}(m,2)-\lambda_{\mathbf{a}}(m,3)}{2}}-e^{\frac{\lambda_{\mathbf{a}}(m,1)+\lambda_{\mathbf{a}}(m,2)-\lambda_{\mathbf{a}}(m,3)}{2}}-e^{\frac{\lambda_{\mathbf{a}}(m,1)-\lambda_{\mathbf{a}}(m,2)-\lambda_{\mathbf{a}}(m,3)}{2}}}{\left(e^{\frac{\lambda_{\mathbf{a}}(m,1)}{2}}+e^{-\frac{\lambda_{\mathbf{a}}(m,1)}{2}}\right)\left(e^{\frac{\lambda_{\mathbf{a}}(m,2)}{2}}+e^{-\frac{\lambda_{\mathbf{a}}(m,2)}{2}}\right)\left(e^{\frac{\lambda_{\mathbf{a}}(m,3)}{2}}+e^{-\frac{\lambda_{\mathbf{a}}(m,3)}{2}}\right)}\\ &\quad +a\frac{e^{\frac{-\lambda_{\mathbf{a}}(m,1)-\lambda_{\mathbf{a}}(m,2)+\lambda_{\mathbf{a}}(m,3)}{2}}+e^{\frac{-\lambda_{\mathbf{a}}(m,1)+\lambda_{\mathbf{a}}(m,2)+\lambda_{\mathbf{a}}(m,3)}{2}}-e^{\frac{\lambda_{\mathbf{a}}(m,1)+\lambda_{\mathbf{a}}(m,2)+\lambda_{\mathbf{a}}(m,3)}{2}}-e^{\frac{\lambda_{\mathbf{a}}(m,1)-\lambda_{\mathbf{a}}(m,2)+\lambda_{\mathbf{a}}(m,3)}{2}}}{\left(e^{\frac{\lambda_{\mathbf{a}}(m,1)}{2}}+e^{-\frac{\lambda_{\mathbf{a}}(m,1)}{2}}\right)\left(e^{\frac{\lambda_{\mathbf{a}}(m,2)}{2}}+e^{-\frac{\lambda_{\mathbf{a}}(m,2)}{2}}\right)\left(e^{\frac{\lambda_{\mathbf{a}}(m,3)}{2}}+e^{-\frac{\lambda_{\mathbf{a}}(m,3)}{2}}\right)}\\ &= \frac{be^{\frac{-\lambda_{\mathbf{a}}(m,3)}{2}}}{e^{\frac{\lambda_{\mathbf{a}}(m,3)}{2}}+e^{-\frac{\lambda_{\mathbf{a}}(m,3)}{2}}}\tanh\left(-\frac{\lambda_{\mathbf{a}}(m,1)}{2}\right)+\frac{ae^{\frac{\lambda_{\mathbf{a}}(m,3)}{2}}}{e^{\frac{\lambda_{\mathbf{a}}(m,3)}{2}}+e^{-\frac{\lambda_{\mathbf{a}}(m,3)}{2}}}\tanh\left(-\frac{\lambda_{\mathbf{a}}(m,1)}{2}\right)\\ &= \frac{ae^{\frac{\lambda_{\mathbf{a}}(m,3)}{2}}+be^{\frac{-\lambda_{\mathbf{a}}(m,3)}{2}}}{e^{\frac{\lambda_{\mathbf{a}}(m,3)}{2}}+e^{-\frac{\lambda_{\mathbf{a}}(m,3)}{2}}}\tanh\left(-\frac{\lambda_{\mathbf{a}}(m,1)}{2}\right).\end{aligned}$$

Similarly, we can calculate the real part of $\hat{X}_{m,i}$ following the 16QAM mapping table given in

Table 5.4 as

$$
\begin{aligned}
\hat{X}_{m,i,\,Re} &= -3P(S_{m,I}=-3) - P(S_{m,I}=-1) + P(S_{m,I}=1) + 3P(S_{m,I}=3) \\
&= -3P\left(c_m^1=0\right)P\left(c_m^2=0\right) - P\left(c_m^1=0\right)P\left(c_m^2=1\right) \\
&\qquad + P\left(c_m^1=1\right)P\left(c_m^2=1\right) + 3P\left(c_m^1=1\right)P\left(c_m^2=0\right) \\
&= \frac{-3 - e^{\lambda_{\mathbf{a}}(m,2)} + e^{\lambda_{\mathbf{a}}(m,1)+\lambda_{\mathbf{a}}(m,2)} + 3e^{\lambda_{\mathbf{a}}(m,1)}}{\left(1+e^{\lambda_{\mathbf{a}}(m,1)}\right)\left(1+e^{\lambda_{\mathbf{a}}(m,2)}\right)} \\
&= \frac{\left(e^{\lambda_{\mathbf{a}}(m,1)-1}\right)\left(3+e^{\lambda_{\mathbf{a}}(m,2)}\right)}{\left(1+e^{\lambda_{\mathbf{a}}(m,1)}\right)\left(1+e^{\lambda_{\mathbf{a}}(m,2)}\right)} \\
&= \frac{\left(e^{\lambda_{\mathbf{a}}(m,1)/2} - e^{-\lambda_{\mathbf{a}}(m,1)/2}\right)\left(3+e^{\lambda_{\mathbf{a}}(m,2)}\right)}{\left(e^{\lambda_{\mathbf{a}}(m,1)/2} + e^{-\lambda_{\mathbf{a}}(m,1)/2}\right)\left(1+e^{\lambda_{\mathbf{a}}(m,2)}\right)} \\
&= \tanh\left(\frac{\lambda_{\mathbf{a}}(m,1)}{2}\right)\frac{3+e^{\lambda_{\mathbf{a}}(m,2)}}{1+e^{\lambda_{\mathbf{a}}(m,2)}}.
\end{aligned}
$$

Following the same procedure, we can derive the estimate of the imaginary part of $\hat{X}_{m,i}$ for 16QAM, and the estimate of $\hat{X}_{m,i}$ for 64QAM.

## Appendix B. Expansion of $(\mathbf{R}-\mathbf{H}\mathbf{X})^H(\mathbf{R}-\mathbf{H}\mathbf{X})$

$$
\begin{aligned}
&(\mathbf{R}-\mathbf{H}\mathbf{X})^H(\mathbf{R}-\mathbf{H}\mathbf{X})\Big|_{\breve{\mathbf{X}}_{k,i-1}} = \mathbf{R}^H\mathbf{R} - \mathbf{R}^H\mathbf{H}\mathbf{X} - \mathbf{X}^H\mathbf{H}^H\mathbf{R} + \mathbf{X}^H\mathbf{H}^H\mathbf{H}\mathbf{X}\Big|_{\breve{\mathbf{X}}_{k,i-1}} \\
=\;& \mathbf{R}^H\mathbf{R} - \sum_{p=1}^{n_T}\mathbf{R}^H X_p\mathbf{H}_p - \sum_{p=1}^{n_T} X_p^*\mathbf{H}_p^H\mathbf{R} + \sum_{p=1}^{n_T}\sum_{q=1}^{n_T}\mathbf{H}_p^H\mathbf{H}_q X_p^* X_q\Big|_{\breve{\mathbf{X}}_{k,i-1}} \\
=\;& \underbrace{\mathbf{R}^H\mathbf{R} - \sum_{\substack{p=1\\p\neq k}}^{n_T}\mathbf{R}^H\breve{X}_{p,i-1}\mathbf{H}_p - \sum_{\substack{p=1\\p\neq k}}^{n_T}\breve{X}_{p,i-1}^*\mathbf{H}_p^H\mathbf{R} + \sum_{\substack{p=1\\p\neq k}}^{n_T}\sum_{\substack{q=1\\q\neq k}}^{n_T}\mathbf{H}_p^H\mathbf{H}_q\breve{X}_{p,i-1}^* X_q}_{\mathcal{C}} \\
&\qquad + \|\mathbf{H}_k X_k\|^2 - \mathbf{R}^H X_k\mathbf{H}_k - X_k^*\mathbf{H}_k^H\mathbf{R} + \sum_{\substack{p=1\\p\neq k}}^{n_T}\left(\mathbf{H}_p^H\mathbf{H}_k\breve{X}_{p,i-1}^* X_k + H_k^H\mathbf{H}_p X_k^*\breve{X}_{p,i-1}\right) \\
=\;& \mathcal{C} + \|\mathbf{H}_k X_k\|^2 - \mathbf{R}^H X_k\mathbf{H}_k - X_k^*\mathbf{H}_k^H\mathbf{R} + \sum_{\substack{p=1\\p\neq k}}^{n_T}\left(\mathbf{H}_p^H\mathbf{H}_k\breve{X}_{p,i-1}^* X_k + H_k^H\mathbf{H}_p X_k^*\breve{X}_{p,i-1}\right) \\
=\;& \mathcal{C} + \|\mathbf{H}_k X_k\|^2 - \mathbf{H}_k^H X_k^*\left(\mathbf{R} - \sum_{\substack{p=1\\p\neq k}}^{n_T}\mathbf{H}_p\breve{X}_{p,i-1}\right) - \left(\mathbf{R} - \sum_{\substack{p=1\\p\neq k}}^{n_T}\mathbf{H}_p\breve{X}_{p,i-1}\right)^H\mathbf{H}_k X_k \\
=\;& \mathcal{C} + \|\mathbf{H}_k X_k\|^2 - 2\mathrm{R}e\left[\mathbf{H}_k^H X_k^*\left(\mathbf{R} - \sum_{\substack{p=1\\p\neq k}}^{n_T}\mathbf{H}_p\breve{X}_{p,i-1}\right)\right] \\
=\;& \mathcal{C} + \|\mathbf{H}_k X_k\|^2 - 2\mathrm{R}e\left[X_k^*\tilde{X}_{k,i}\right], \qquad (5.56)
\end{aligned}
$$

where $\mathcal{C}$ is a constant term for all the hypotheses of $X_k$, "$\mathrm{Re}$" denotes the real part of the signal, $\mathbf{H}_k$ denotes the $k$th column of channel matrix $\mathbf{H}$, and

$$\tilde{X}_{k,i} = \mathbf{H}_k^H \left( \mathbf{R} - \sum_{\substack{p=1 \\ p \neq k}}^{n_T} \mathbf{H}_p \breve{X}_{p,i-1} \right)$$

is the IC-MRC decision statistic, as shown in (5.37).

# Appendix C. Mean and Variance of the Interference at IC-MRC Output

Defining $\rho_{kp} = \mathbf{H}_k^H \mathbf{H}_p$, we have

$$\begin{aligned}
\mathcal{E}\left\{\tilde{v}_k\right\} &= \mathcal{E}\left\{ \sum_{\substack{p=1 \\ p \neq k}}^{n_T} \rho_{kp} \left( X_p - \breve{X}_{p,i-1} \right) + \mathbf{H}_k^H \mathbf{V} \right\} \\
&= \sum_{\substack{p=1 \\ p \neq k}}^{n_T} \rho_{kp} \mathcal{E}\left\{ X_p - \breve{X}_{p,i-1} \right\} + \mathbf{H}_k^H \mathcal{E}\left\{\mathbf{V}\right\} \\
&= 0,
\end{aligned}$$

and

$$\begin{aligned}
\mathcal{E}\left\{|\tilde{v}_k|^2\right\} &= \mathcal{E}\left\{ \left| \sum_{\substack{p=1 \\ p \neq k}}^{n_T} \rho_{kp} \left( X_p - \breve{X}_{p,i-1} \right) + \mathbf{H}_k^H \mathbf{V} \right|^2 \right\} \\
&= \mathcal{E}\left\{ \left| \sum_{\substack{p=1 \\ p \neq k}}^{n_T} \rho_{kp} \left( X_p - \breve{X}_{p,i-1} \right) \right|^2 \right\} + \mathcal{E}\left\{ \left| \mathbf{H}_k^H \mathbf{V} \right|^2 \right\} \\
&= \mathcal{E}\left\{ \sum_{\substack{p=1 \\ p \neq k}}^{n_T} \sum_{\substack{q=1, \\ q \neq k}}^{n_T} \rho_{kp} \rho_{kq}^* \left( X_p - \breve{X}_{p,i-1} \right) \left( X_q^* - \breve{X}_{q,i-1}^* \right) \right\} + 2\sigma^2 g_k \\
&= \sum_{\substack{p=1 \\ p \neq k}}^{n_T} \sum_{\substack{q=1, \\ q \neq k}}^{n_T} \mathcal{E}\left\{ \rho_{kp} \rho_{kq}^* \right\} \mathcal{E}\left\{ \left( X_p - \breve{X}_{p,i-1} \right) \left( X_q^* - \breve{X}_{q,i-1}^* \right) \right\} + 2\sigma^2 g_k \\
&= \sum_{\substack{p=1 \\ p \neq k}}^{n_T} \mathcal{E}\left\{ |\rho_{kp}|^2 \right\} \left( 1 - \left| \breve{X}_{p,i-1} \right|^2 \right) + 2\sigma^2 g_k
\end{aligned}$$

under the assumption that $\rho_{kp}$ is independent of $X_p$ and $\breve{X}_{p,i-1}$.

For OFDM systems, we have the frequency domain channel coefficients

$$H_{n,k}(m) = \sum_{l=0}^{L-1} h_{n,k}(l) e^{-\frac{2\pi}{N} ml},$$

where $n$ and $k$ are the receive antenna and transmit antenna indices, $n = 1,\ 2,\ \cdots,\ n_R$, $k = 1,\ 2,\ \cdots,\ n_T$, $m$ is the subcarrier index, $m = 0,\ 1,\ 2,\ \cdots,\ N-1$, $N$ is the FFT size, $l$ is the multipath index, $l = 0,\ 1,\ \cdots,\ L-1$ with $L$ being the number of multipaths in the channel corresponding to transmit antenna $n$ and receiver antenna $k$, and $h_{n,k}(l)$ is the time-domain multipath coefficients.

Assuming wide sense stationary uncorrelated scattering (WSSUS) multipath channel for each $(n, k)$ MIMO channel, and spatially uncorrelated for different $(n, k)$ pairs, we will look into two channels: uniform power delay (UPD) profile and exponentially decaying power delay (EPD) profile.

## C.1 Uniform Power Delay Profile

For WSSUS UPD profile, each multipath is i.i.d. complex Gaussian with zero mean and variance $\frac{1}{Ln_T}$. We hence have for $k \neq p$

$$
\begin{aligned}
&\mathcal{E}\left\{|\rho_{kp}|^2\right\} = \mathcal{E}\left\{\mathbf{H}_k^H \mathbf{H}_p \mathbf{H}_p^H \mathbf{H}_k\right\} \\
= \ & \mathcal{E}\left\{\sum_{g=1}^{n_R} H_{g,k}^*(m) H_{g,p}(m) \sum_{n=1}^{n_R} H_{n,k}(m) H_{n,p}^*(m)\right\} \\
= \ & \mathcal{E}\left\{\sum_{g=1}^{n_R}\sum_{n=1}^{n_R} H_{g,k}^*(m) H_{n,k}(m) H_{g,p}(m) H_{n,p}^*(m)\right\} \\
= \ & \mathcal{E}\left\{\sum_{g=1}^{n_R}\sum_{n=1}^{n_R}\left[\sum_{l=0}^{L-1} h_{g,k}^*(l) e^{j\frac{2\pi}{N}ml} \sum_{i=0}^{L-1} h_{n,k}(i) e^{-j\frac{2\pi}{N}mi} \sum_{t=0}^{L-1} h_{g,p}(t) e^{-j\frac{2\pi}{N}mt} \sum_{s=0}^{L-1} h_{n,p}^*(s) e^{j\frac{2\pi}{N}mg}\right]\right\} \\
= \ & \mathcal{E}\left\{\sum_{g=1}^{n_R}\sum_{n=1}^{n_R}\sum_{l=0}^{L-1}\sum_{i=0}^{L-1}\sum_{t=0}^{L-1}\sum_{s=0}^{L-1} h_{g,k}^*(l) h_{n,k}(i) h_{g,p}(t) h_{n,p}^*(s) e^{j\frac{2\pi}{N}m(l-i-t+s)}\right\} \\
= \ & \sum_{g=1}^{n_R}\sum_{n=1}^{n_R}\sum_{l=0}^{L-1}\sum_{i=0}^{L-1}\sum_{t=0}^{L-1}\sum_{s=0}^{L-1} \mathcal{E}\left\{h_{g,k}^*(l) h_{n,k}(i) h_{g,p}(t) h_{n,p}^*(s)\right\} e^{j\frac{2\pi}{N}m(l-i-t+s)}
\end{aligned}
$$

From [102], when $Z_1$, $Z_2$, $Z_3$ and $Z_4$ are zero-mean, stationary complex Gaussian, we have

$$
\mathcal{E}\left\{Z_1^* Z_2^* Z_3 Z_4\right\} = \mathcal{E}\left\{Z_1^* Z_3\right\}\mathcal{E}\left\{Z_2^* Z_4\right\} + \mathcal{E}\left\{Z_1^* Z_4\right\}\mathcal{E}\left\{Z_2^* Z_3\right\},
$$

therefore,

$$\begin{aligned}
\mathcal{E}\left\{|\rho_{kp}|^2\right\} &= \sum_{g=1}^{n_R}\sum_{n=1}^{n_R}\sum_{l=0}^{L-1}\sum_{i=0}^{L-1}\sum_{t=0}^{L-1}\sum_{s=0}^{L-1}\left[\mathcal{E}\left\{h_{g,k}^*(l)h_{n,k}(i)\right\}\mathcal{E}\left\{h_{g,p}(t)h_{n,p}^*(s)\right\}\right. \\
&\qquad \left. +\mathcal{E}\left\{h_{g,k}^*(l)h_{g,p}(t)\right\}\mathcal{E}\left\{h_{n,k}(i)h_{n,p}^*(s)\right\}\right] e^{j\frac{2\pi}{N}m(l-i-t+s)} \\
&= \frac{1}{n_T^2L^2}\sum_{g=1}^{n_R}\sum_{n=1}^{n_R}\sum_{l=0}^{L-1}\sum_{i=0}^{L-1}\sum_{t=0}^{L-1}\sum_{s=0}^{L-1}\left[\delta(g-n)\delta(l-i)\delta(t-s)\right. \\
&\qquad \left. +\delta(k-p)\delta(l-t)\delta(k-p)\delta(i-s)\right] e^{j\frac{2\pi}{N}m(l-i-t+s)} \\
&= \sum_{g=1}^{n_R}\sum_{n=1}^{n_R}\sum_{l=0}^{L-1}\sum_{i=0}^{L-1}\sum_{t=0}^{L-1}\sum_{s=0}^{L-1}\left(\frac{1}{Ln_T}\right)^2\delta(g-n)\delta(l-i)\delta(t-s)e^{j\frac{2\pi}{N}m(l-i-t+s)} \\
&= \frac{n_R}{n_T^2},
\end{aligned}$$

where $\delta(\cdot)$ is the Dirac delta function.

We then have the composite interference and AWGN noise power for the UPD profile as

$$\varsigma^2 = \frac{1}{2}\mathcal{E}\left\{|\tilde{\mu}_{k,i}|^2\right\} = \frac{1}{2}\sum_{\substack{p=1\\p\neq k}}^{n_T}\frac{n_R}{n_T^2}\left(1-\left|\breve{X}_{p,i-1}\right|^2\right) + g_k\sigma^2. \tag{5.57}$$

## C.2 Exponential Power Delay Profile

For WSSUS EPD profile, we have the channel impulse response expressed as

$$h(n) = \sum_{l=0}^{L-1} h_l\delta(n-l) = \sum_{l=0}^{L-1} ae^{-l\beta}\delta(n-l),$$

where $\beta$ is the power loss law exponent and $a$ is the normalization factor such that the MIMO multipath channel does not change the average SNR at its input and output, and $a$ is

$$a = \sqrt{\frac{1-e^{-2\beta}}{n_T\sum_{l=0}^{L-1}e^{-2l\beta}}}\ .$$

Same as the UPD channel, each multipath $h_l$ is complex Gaussian with zero mean and variance $\sigma_l^2 = a^2e^{-2l\beta}$.

Making use of the derivation result for UPD channels, we hence have

$$
\begin{aligned}
&\mathcal{E}\left\{|\rho_{kp}|^2\right\} = \mathcal{E}\left\{\mathbf{H}_k^H\mathbf{H}_p\mathbf{H}_p^H\mathbf{H}_k\right\} \\
= \;& \sum_{g=1}^{n_R}\sum_{n=1}^{n_R}\sum_{l=0}^{L-1}\sum_{i=0}^{L-1}\sum_{t=0}^{L-1}\sum_{s=0}^{L-1}\left[\mathcal{E}\left\{h_{g,k}^*(l)h_{n,k}(i)\right\}\mathcal{E}\left\{h_{g,p}(t)h_{n,p}^*(s)\right\}\right. \\
&\qquad\qquad \left.+\mathcal{E}\left\{h_{g,k}^*(l)h_{g,p}(t)\right\}\mathcal{E}\left\{h_{n,k}(i)h_{n,p}^*(s)\right\}\right] e^{j\frac{2\pi}{N}m(l-i-t+s)} \\
= \;& \sum_{g=1}^{n_R}\sum_{n=1}^{n_R}\sum_{l=0}^{L-1}\sum_{i=0}^{L-1}\sum_{t=0}^{L-1}\sum_{s=0}^{L-1}\left[\sigma_l\sigma_i\sigma_t\sigma_s\delta(g-n)\delta(l-i)\delta(t-s)\right. \\
&\qquad\qquad \left.+\sigma_l\sigma_i\sigma_t\sigma_s\delta(k-p)\delta(l-t)\delta(k-p)\delta(i-s)\right] e^{j\frac{2\pi}{N}m(l-i-t+s)} \\
= \;& \sum_{g=1}^{n_R}\sum_{n=1}^{n_R}\sum_{l=0}^{L-1}\sum_{i=0}^{L-1}\sum_{t=0}^{L-1}\sum_{s=0}^{L-1}\sigma_l\sigma_i\sigma_t\sigma_s\delta(g-n)\delta(l-i)\delta(t-s)e^{j\frac{2\pi}{N}m(l-i-t+s)} \\
= \;& \sum_{g=1}^{n_R}\sum_{l=0}^{L-1}\sum_{t=0}^{L-1}\sigma_l^2\sigma_t^2 \\
= \;& \sum_{g=1}^{n_R}\left[a^2\frac{\sum_{l=0}^{L-1}e^{-2l\beta}}{1-e^{-2\beta}}\right]\left[a^2\frac{\sum_{l=0}^{L-1}e^{-2l\beta}}{1-e^{-2\beta}}\right] \\
= \;& \frac{n_R}{n_T^2},
\end{aligned}
$$

which is the same as the Uniform Power Delay Profile multipath channel.

The composite interference and AWGN noise power for the Exponential Power Delay profile is therefore also the same as Uniform Power Delay Profile multipath channel, as given in (5.57).

# Chapter 6

# EXIT Chart Analysis

In this chapter we present the EXIT chart analysis of the proposed Bayesian MMSE turbo receivers. EXIT chart was first proposed by S. ten Brink to trace the convergence behavior of iterative decoding of turbo codes [103] by observing the mutual information trajectory of $I^{\mathrm{a}}$ and $I^{\mathrm{e}}$, the mutual information between the input *a priori* values $L_{\mathrm{a}}$ to the SISO decoder and the coded bits ($I^{\mathrm{a}}$), and the mutual information between the output extrinsic values $L_e$ from the SISO decoder and the coded bits ($I^{\mathrm{e}}$). The EXIT chart analysis was later applied to trace convergence of other turbo processing algorithms, e.g., for turbo-equalization in [104], for turbo receivers in MIMO systems in [105], etc.. In [106], a good tutorial is given on the EXIT chart analysis in iterative processing.

It is worth mentioning that for turbo and turbo-like codes, e.g., LDPC codes [107][108], serially concatenated convolutional codes (SCCC) [109], etc., another method based on density evolution has also been proposed to analyze the convergence behaviour of iterative decoding, see works by Divsalar *et. al.* [110], Richardson and Urbanke [111][112]. The density evolution analysis tracks the pdf of the EXT as the density evolves from iteration to iteration. As the pdf of the EXT can be approximated by a Gaussian density function [113], the density evolution analysis and the EXIT chart analysis share a lot of commonalities. For other turbo processing algorithms, e.g., turbo equalization and turbo receiver, the EXT output from the soft-output equalizer (turbo equalization) or detector (turbo receiver) can not be approximated by a Gaussian distribution [114]. In this case, Hagenauer and Tüchler proposed to use the time average to replace the statistical expectation in mutual information calculation [114], making use of the ergodicity theorem. This is a very straightforward replacement. The density evolution analysis, however, will become a lot more complex.

The rest of the chapter is organized as follows. In Section 6.1, we give a general overview of the mutual information of EXT and in Section 6.2, we present the EXIT chart derivation of the BMMSE detectors. The numerical results of the BMMSE detector are presented in Section 6.3. Finally the conclusions are drawn in Section 6.4.

## 6.1 Mutual Information of Extrinsic Information

We have briefly discussed the mutual information definition in Section 4.3. That discussion is more focussed on the mutual information between the MIMO channel input $\mathbf{X}$ with fixed modulation and the unconstrained MIMO channel output $\mathbf{R}$. Here we focus on real binary input unconstrained output "channels" with the coded bits as the input and the real-valued LLR information as the output. The LLR can be either the *a priori* LLR information at the SISO detectors/decoders input, or the *extrinsic* LLR information at the SISO modules output. We are interested in determining the mutual information between the coded bits and the input *a priori* LLR, as well as the mutual information between the coded bits and the output *extrinsic* LLR.

Let $X$ and $Y$ be two real valued r.v. with pdf $f(x)$ and $f(y)$ and joint density function $f(x, y)$, then the mutual information between $X$ and $Y$ is defined as [6]

$$\begin{aligned} I(X;Y) &= \mathcal{E}\left\{\log \frac{f(x,y)}{f(x)\cdot f(y)}\right\} \\ &= \int\int f(x,y)\log \frac{f(x,y)}{f(x)\cdot f(y)}dxdy \\ &= \int\int f(y|x)f(x)\log \frac{f(y|x)}{f(y)}dxdy. \end{aligned}$$

For binary input unconstrained output AWGN channel $Y = X + Z$ where $x \in \{+1, -1\}$, and $z$ is statistically independent AWGN with zero mean and variance $\sigma^2$, we have

$$I(X;Y) = \sum_{x=\pm 1}\int_{-\infty}^{+\infty} f(y|x)p(x)\log \frac{f(y|x)}{f(y)}dy,$$

with $p(x)$ being the probability mass function (pmf). The maximum mutual information is achieved for equally likely input $x$ [6], which is

$$I(X;Y) = \frac{1}{2}\sum_{x=\pm 1}\int_{-\infty}^{+\infty} f(y|x)\log\frac{f(y|x)}{f(y)}dy,$$

and

$$\begin{aligned} f(y) &= \frac{1}{2}f(y|x=+1) + \frac{1}{2}f(y|x=-1) \\ &= \frac{1}{2}\frac{1}{\sqrt{2\pi}\sigma}\left(e^{-\frac{(y-1)^2}{2\sigma^2}} + e^{-\frac{(y+1)^2}{2\sigma^2}}\right). \end{aligned}$$

As extensively discussed in [113], [103], [104], [110], and [115], the *a priori* and extrinsic information, i.e., the LLR values $(\lambda|c)$ at the input and output of the SISO decoder can be well approximated by a conditionally independent and identically distributed Gaussian random variable satisfying the "consistency" condition

$$|\mathcal{E}(\lambda|c)| = \frac{\sigma_{\mathrm{a}}^2}{2},$$

where $\sigma_{\mathrm{a}}^2$ is its variance. The conditional pdf is therefore

$$f(\lambda|c) = \frac{1}{\sqrt{2\pi}\sigma_{\mathrm{a}}}e^{-\frac{(\lambda - c\frac{\sigma_{\mathrm{a}}^2}{2})^2}{2\sigma_{\mathrm{a}}^2}} = \frac{1}{\sqrt{2\pi}\sigma_{\mathrm{a}}}e^{-\frac{(-\lambda - c\frac{\sigma_{\mathrm{a}}^2}{2})^2}{2\sigma_{\mathrm{a}}^2}}e^{c\lambda} = f(-\lambda|c)e^{c\lambda}. \tag{6.1}$$

The mutual information between $\Lambda$ and the coded bits $C$ is therefore only dependent on $\sigma_{\mathrm{a}}$, and written as

$$
\begin{aligned}
I(\Lambda;C) &= \frac{1}{2}\sum_{x=\pm1}\int_{-\infty}^{+\infty} f(\lambda|c)\log\frac{f(\lambda|c)}{f(\lambda)}d\lambda \qquad (6.2)\\
&= \frac{1}{2}\left[\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\sigma_{\mathrm{a}}}e^{-\frac{\left(\lambda-\frac{\sigma_{\mathrm{a}}^2}{2}\right)^2}{2\sigma_{\mathrm{a}}^2}}\log_2\frac{\frac{1}{\sqrt{2\pi}\sigma_{\mathrm{a}}}e^{-\frac{\left(\lambda-\frac{\sigma_{\mathrm{a}}^2}{2}\right)^2}{2\sigma_{\mathrm{a}}^2}}}{\frac{1}{2}\left[\frac{1}{\sqrt{2\pi}\sigma_{\mathrm{a}}}e^{-\frac{\left(\lambda-\frac{\sigma_{\mathrm{a}}^2}{2}\right)^2}{2\sigma_{\mathrm{a}}^2}}+\frac{1}{\sqrt{2\pi}\sigma_{\mathrm{a}}}e^{-\frac{\left(\lambda+\frac{\sigma_{\mathrm{a}}^2}{2}\right)^2}{2\sigma_{\mathrm{a}}^2}}\right]}d\lambda\right.\\
&\qquad\left.+\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\sigma_{\mathrm{a}}}e^{-\frac{\left(\lambda+\frac{\sigma_{\mathrm{a}}^2}{2}\right)^2}{2\sigma_{\mathrm{a}}^2}}\log_2\frac{\frac{1}{\sqrt{2\pi}\sigma_{\mathrm{a}}}e^{-\frac{\left(\lambda+\frac{\sigma_{\mathrm{a}}^2}{2}\right)^2}{2\sigma_{\mathrm{a}}^2}}}{\frac{1}{2}\left[\frac{1}{\sqrt{2\pi}\sigma_{\mathrm{a}}}e^{-\frac{\left(\lambda-\frac{\sigma_{\mathrm{a}}^2}{2}\right)^2}{2\sigma_{\mathrm{a}}^2}}+\frac{1}{\sqrt{2\pi}\sigma_{\mathrm{a}}}e^{-\frac{\left(\lambda+\frac{\sigma_{\mathrm{a}}^2}{2}\right)^2}{2\sigma_{\mathrm{a}}^2}}\right]}d\lambda\right]\\
&= \frac{1}{2}\left[\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\sigma_{\mathrm{a}}}e^{-\frac{\left(\lambda-\frac{\sigma_{\mathrm{a}}^2}{2}\right)^2}{2\sigma_{\mathrm{a}}^2}}\log_2\frac{2}{1+e^{-\lambda}}d\lambda+\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\sigma_{\mathrm{a}}}e^{-\frac{\left(\lambda+\frac{\sigma_{\mathrm{a}}^2}{2}\right)^2}{2\sigma_{\mathrm{a}}^2}}\log_2\frac{2}{1+e^{\lambda}}d\lambda\right]\\
&= \int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\sigma_{a}}e^{-\frac{\left(\lambda-\frac{\sigma_{a}^2}{2}\right)^2}{2\sigma_{a}^2}}\log_2\frac{2}{1+e^{-\lambda}}d\lambda\\
&= 1-\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\sigma_{\mathrm{a}}}e^{-\frac{\left(\lambda-\frac{\sigma_{\mathrm{a}}^2}{2}\right)^2}{2\sigma_{\mathrm{a}}^2}}\log_2\left(1+e^{-\lambda}\right)d\lambda. \qquad (6.3)
\end{aligned}
$$

The actual value of $I(\Lambda;C)$ can be obtained through numerical evaluation of (6.3).

For notational convenience, (6.3) is defined as a mapping function between the mutual information $I(\Lambda;C)=I^{\mathrm{a}}$ and $\sigma_{\mathrm{a}}^2$ [103], i.e.,

$$
I^{\mathrm{a}}=J(\sigma_{\mathrm{a}}^2)=1-\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\sigma_{\mathrm{a}}}e^{-\frac{\left(\lambda-\frac{\sigma_{\mathrm{a}}^2}{2}\right)^2}{2\sigma_{\mathrm{a}}^2}}\log_2\left(1+e^{-\lambda}\right)d\lambda. \qquad (6.4)
$$

For some SISO processors, e.g., SISO equalizer in ISI channels [104], SISO detector for CDMA channels [85] and MIMO channels [105], it is difficult to analytically define the pdf of the extrinsic LLR values. In this case, the ergodicity theorem can be used to replace the statistical expectation by time average, and the mutual information can be computed for a large number of samples as [114] [106]

$$I(\Lambda; C) = 1 - \mathcal{E}\left\{1 + e^{-\lambda}\right\} \approx 1 - \frac{1}{N}\sum_{n=1}^{N} \log_2\left(1 + e^{-c_n \lambda_n}\right). \tag{6.5}$$

For accurate approximation by (6.5), the coded block length $N$ needs to be large enough.

## 6.2 Derivation of EXIT Chart of SISO Bayesian Detectors

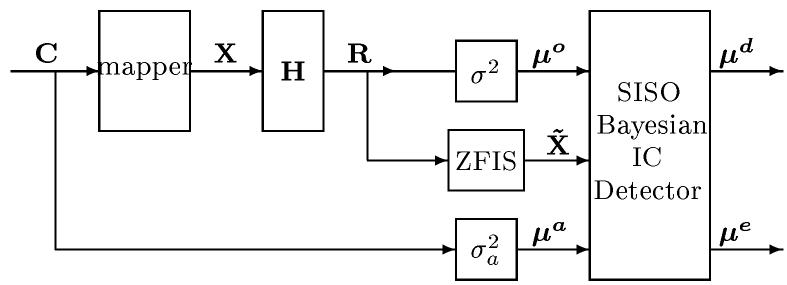


**Figure 6.1: Block diagram for the EXIT chart derivation of the SISO Bayesian MMSE detecor.**

In this section, we present the derivation of EXIT chart analysis for the soft-input soft-output Bayesian MMSE detectors according to the block diagram depicted in Fig. 6.1. The coded and interleaved bits are denoted by the binary $(\log_2 M n_T)$-vector $\mathbf{c} = \left[\mathbf{c}_1^T, \ \mathbf{c}_2^T, \ \cdots, \ \mathbf{c}_{n_T}^T\right]^T$, where $\mathbf{c}_i = \left[c_{i,1}, \ c_{i,2}, \ \cdots, \ c_{i,\log_2 M}\right]^T$, $c_{i,j} \in \{\pm 1\}$, and M denotes the constellation size of the modulation and $\log_2 M$ denotes the number of bits per symbol. The binary vector $\mathbf{c}$ is mapped to the symbol vector $\mathbf{x}$ with length $n_T$ according to the mapping rule $\mathbf{x} = \mathbf{m}(\mathbf{c})$. In this chapter we consider only Gray mapping.

If direct mapping is used in the MIMO precoding, the length-$n_T$ vector $\mathbf{m}(\mathbf{c})$ will be transmitted through the $n_T \times n_R$ MIMO channel represented by $\mathbf{H}$, with its $(i, j)$-th element denoting the channel corresponding to the transmit-receive antenna pair $(j, i)$. The extrinsic message input to the BMMSE detector $\mu^o$ is generated according to the conditional pdf

$$p(\mathbf{r}|\mathbf{c}) = \frac{1}{(2\pi\sigma^2)^{n_R}} \exp\left(-\frac{\|\mathbf{r} - \mathbf{H}\mathbf{m}(\mathbf{c})\|^2}{2\sigma^2}\right), \tag{6.6}$$

sampled at all possible values of $\mathbf{c} \in \{\pm 1\}^{n_T \log_2 M}$. The *a priori* information input to the Bayesian IC detector is generated according to

$$\lambda_{ij}^{\mathrm{a}}|c_{ij} \sim \mathcal{N}\left(c_{ij}\frac{\sigma_{\mathrm{a}}^2}{2}, \sigma_{\mathrm{a}}^2\right) \tag{6.7}$$

under the assumption that the *a priori* LLR is the interleaved version of EXT from the SISO decoder and it is therefore Gaussian distributed [103] with mean $c_{ij}\frac{\sigma_{\mathrm{a}}^2}{2}$ and variance $\sigma_{\mathrm{a}}^2$, as given in (6.1). The mutual information between the input *a priori* information and the coded bits can therefore be computed following (6.4).

The BMMSE estimate $\breve{x}_{k,i}$ in the IC-MRC receiver is computed using the input *a priori* LLR, the received signal $\mathbf{r}$, and the ZFIS estimate, i.e.,

$$\breve{\mathbf{x}}_{i-1} = \mathbf{H}^{\dagger}\mathbf{r},$$

where $\mathbf{H}^{\dagger}$ denotes the pseudo inverse of $\mathbf{H}$.

The BMMSE estimate in the LMMSE-IC receiver is computed using the input *a priori* LLR, the received signal $\mathbf{r}$, and the LMMSE-IS estimate, i.e.,

$$\breve{\mathbf{x}}_{i-1} = \mathbf{H}^{H}\left(\mathbf{H}\mathbf{H}^{H} + 2\sigma^2\mathbf{I}\right)^{-1}\mathbf{r},$$

where $\mathbf{H}^{H}\left(\mathbf{H}\mathbf{H}^{H} + 2\sigma^2\mathbf{I}\right)^{-1}$ is the LMMSE filter.

After $\breve{x}_{k,i}$, $k = 1, \cdots, n_T$ is obtained, either IC-MRC or LMMSE-IC is performed, and the extrinsic LLR of each coded bits $c(k,l)$ is computed, following (5.4). As given in Chapter 5, depending on whether an IC-MRC or LMMSE-IC filtering scheme is applied, the statistical properties of $p(\tilde{X}_{k,i}|S_k)$ in (5.4) need to be changed accordingly.

After the extrinsic LLR $p\left(\lambda_{\mathrm{e}}(k,l)|c(k,l)\right)$ is obtained, we will use (6.5) to compute the output mutual information of the extrinsic LLR as [114]

$$I^{\mathrm{e}}(\lambda_{\mathrm{e}}; C) = 1 - \frac{1}{N}\sum_{n=1}^{N}\log_2\left[1 + \exp\left(-c(n)\,\lambda_{\mathrm{e}}(n)\right)\right].$$

## 6.3 Numerical Results of SISO Bayesian MMSE Detectors

In this section, we present the numerical results of the EXIT chart analysis for the SISO Bayesian MMSE detectors. We use both the $4 \times 4$ static channel matrix given in [105] and random CSCG channels to evaluate the EXIT performance. For easy reference, the static channel matrix from [105] is given in the Appendix of this Chapter.

### 6.3.1 EXIT Chart with the Static $4 \times 4$ Channel

We first depict in Fig. 6.2 the EXIT chart of the conventional and Bayesian detectors for a $4 \times 4$ MIMO system using the static channel. We follow the approach in [96] to derive the EXIT chart for conventional IC-MRC and LMMSE-IC receivers. QPSK modulated signals are used and the noise power is set to $\sigma^2 = 0.1990$.

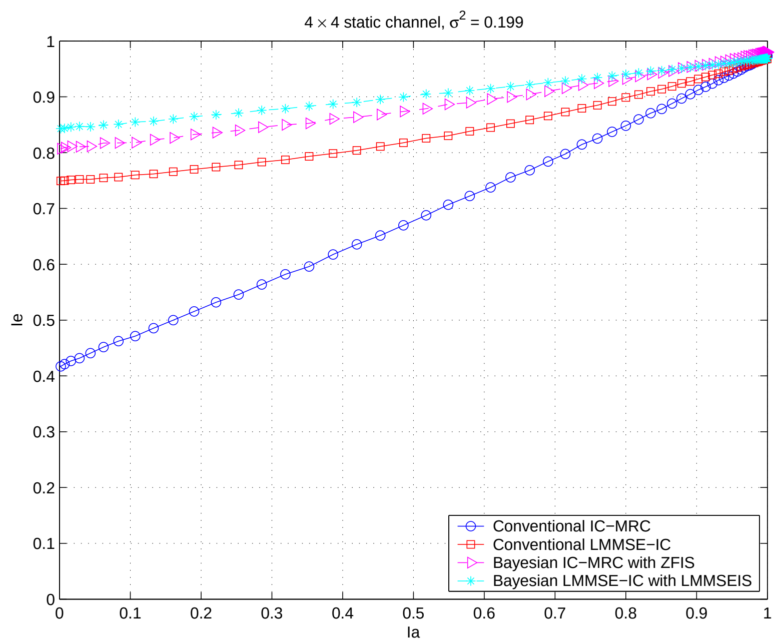


**Figure 6.2: Mutual information transfer function comparison of the conventional and Bayesian MMSE detectors. Static channel, QPSK modulation.** $\sigma^2 = 0.1990$

It is obvious from the figure that the Bayesian IC-MRC receiver outperforms the conventional IC-MRC receiver as it achieves a higher value of $I^{\mathrm{e}}$ at any given $I^{\mathrm{a}}$. More importantly, the Bayesian IC-MRC receiver also outperforms the conventional LMMSE-IC receiver. This is of great practical importance as the soft decision function for QPSK signals remain as hyperbolic-tangent function in the Bayesian receiver, but no matrix inversion is needed in the iterations. The complexity is therefore significantly reduced. As expected, the Bayesian LMMSE-IC receiver improves the output mutual information over the conventional LMMSE-IC receiver.

One more observation from the figure is that the smaller the input $I^{\mathrm{a}}$ is, the larger the difference between the Bayesian and the conventional detectors' output $I^{\mathrm{e}}$ values becomes. This is expected, as at small input $I^{\mathrm{a}}$'s, the BMMSE interference estimation accuracy is greatly improved by using the additional detector decision statistic, and when $I^{\mathrm{a}}$ increases, the *a priori* probability becomes more and more dominant in the BMMSE estimate, making it closer and closer to the *a priori* estimate, i.e., the statistical mean. The performance gap between the Bayesian and the conventional receivers will become smaller correspondingly.

The EXIT chart performance is further verified for QPSK signals with the noise power set to $\sigma^2 = 0.1256$. The results are depicted in Fig. 6.3. Same observations can be made. That is, the Bayesian IC-MRC receiver outperforms both the conventional IC-MRC and the conventional LMMSE-IC receivers, and by using Bayesian MMSE estimation at the LMMSE-IC output, further improvement can be gained over the Bayesian IC-MRC receivers.

The EXIT chart analysis results for 8PSK signals are presented in Fig. 6.4 and Fig. 6.5 with the noise power values of $\sigma^2 = 0.1990$, and $\sigma^2 = 0.1256$, respectively. We only depict the mutual information curves for LMMSE-IC receivers. As expected, the Bayesian LMMSE-IC receiver has higher output mutual information than the conventional receiver. Compared with QPSK modulation results in Fig. 6.2 and Fig. 6.3, however, we can notice that at same input mutual information value, higher output mutual information can be obtained for QPSK signals than the 8PSK signals. At $\sigma^2 = 0.1256$, when the input mutual information reaches the maximum value of 1, the Bayesian LMMSE-IC detector has the output mutual information value approaching 1 as well when QPSK signals are used, but it can only get to about 0.875 for 8PSK signals.

### 6.3.2 EXIT Chart with Random CSCG $4 \times 4$ Channel

Now we present the EXIT chart analysis in $4 \times 4$ random CSCG channels. Shown in Fig. 6.6 is the EXIT chart of the conventional and Bayesian IC-MRC detectors for QPSK modulation, which is obtained

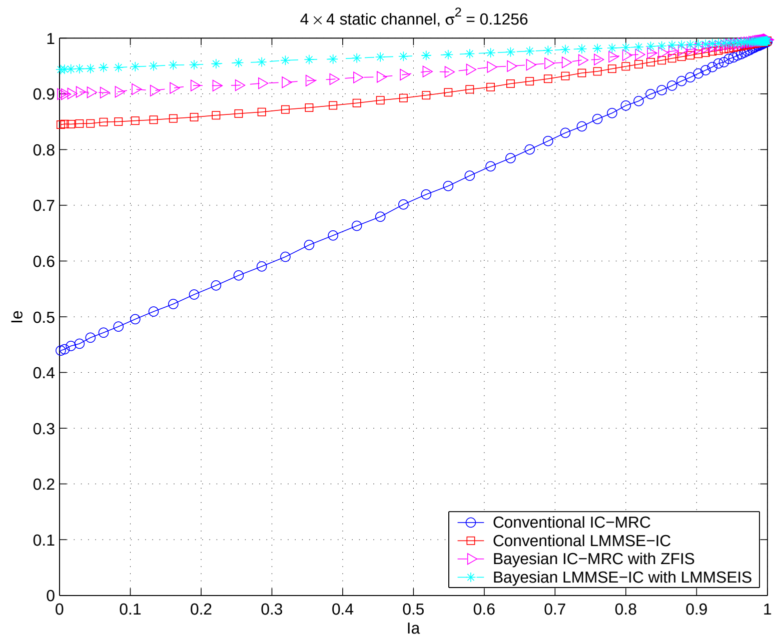


**Figure 6.3: Mutual information transfer function comparison of the conventional and Bayesian MMSE detectors. Static channel, QPSK modulation.** $\sigma^2 = 0.1256$

through averaging over 50000 realizations of the random channel. The average SNR per receive antenna is set to 6dB. The results further prove the superior performance of Bayesian IC-MRC receiver over the conventional one, especially at low to medium values of the input mutual information. When the input mutual information approaches 1, the difference between the conventional and the Bayesian IC-MRC detectors' output mutual information will diminish. The EXIT charts comparison between the conventional and Bayesian IC-MRC receivers for SNR = 8dB is depicted in Fig. 6.7 and same observation can be made from this figure.

The LMMSE-IC turbo receiver EXIT chart analysis results for QPSK signals are depicted in Fig. 6.8 with SNR = 6dB. Similar to the static channel analysis, Bayesian estimation based on the LMMSE-IC output can further improve the performance, especially for low to medium input mutual information values. Comparing the results with that in Fig. 6.6, we can also see that the LMMSE-IC filtering scheme is superior to IC-MRC, exhibited by the higher output mutual information values especially when the input mutual information values are not very high. The Bayesian IC-MRC receiver, however, is superior to the

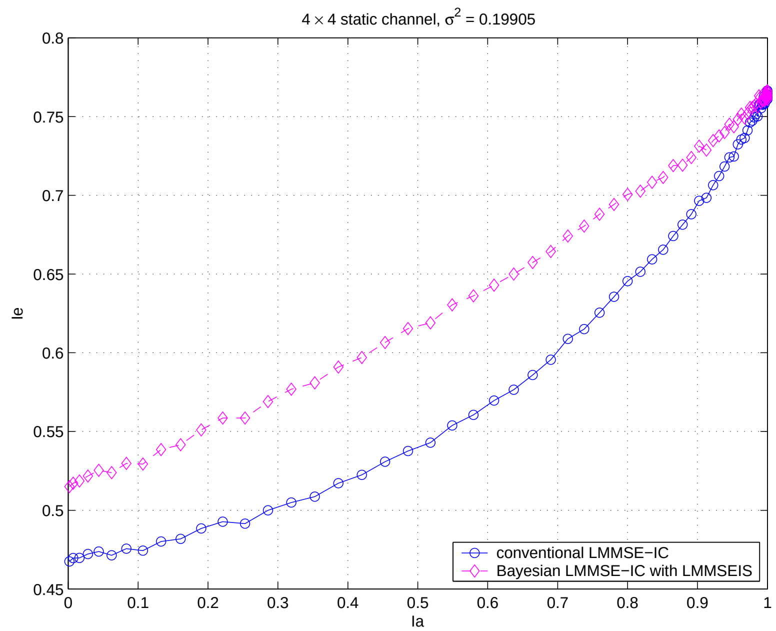


**Figure 6.4: Mutual information transfer function comparison of the conventional and Bayesian MMSE detectors. Static channel, 8PSK modulation.** $\sigma^2 = 0.1990$**.**

conventional LMMSE-IC receiver. This is consistent with the results for static channels, as well as the BER and FER performances we presented in Chapter 5.

Same as the static channel case, we study the EXIT chart of LMMSE-IC receivers for 8PSK signals. The results are presented in Fig. 6.9 for SNR = 8dB, and in Fig. 6.10 for SNR = 6dB. Both results demonstrates the better performance of Bayesian detector than the conventional detector.

### 6.3.3 Convergence Analysis with the Static $4 \times 4$ Channel

With the EXIT chart, we are also able to study the convergence behavior of the turbo receivers. The $4 \times 4$ static MIMO channel given in [105] is used and QPSK modulation is considered. For error control code, we again use the rate-half constraint length $K = 3$ convolutional code with generation function $(5, 7)_{\text{octal}}$.

Presented in Fig. 6.11 is the results of IC-MRC receivers with noise power of $\sigma^2 = 0.199$. The trajectories of the conventional and Bayesian receivers clearly show that the Bayesian receiver requires much fewer iterations to achieve the same performance. Furthermore, performance improvement in the

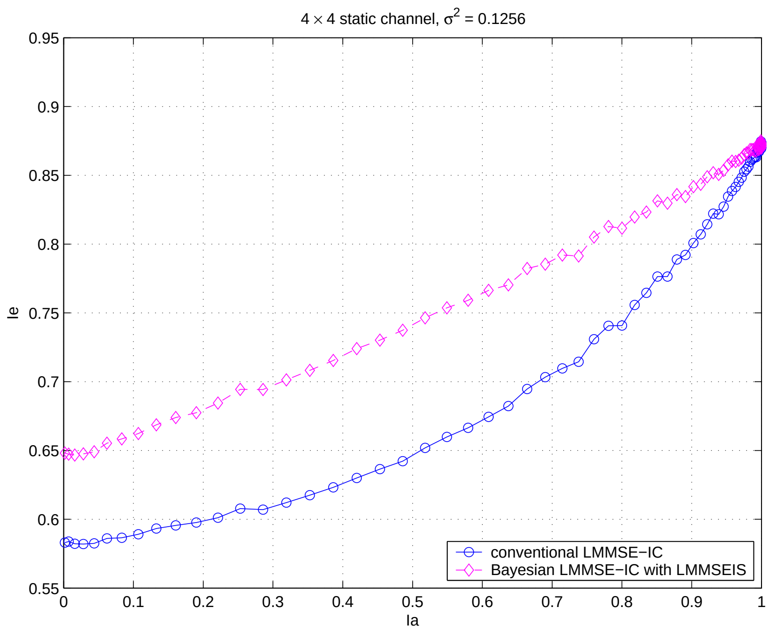


**Figure 6.5: Mutual information transfer function comparison of the conventional and Bayesian MMSE detectors. Static channel, 8PSK modulation. $\sigma^2 = 0.1256$.**

initial iterations of the Bayesian receiver is substantial.

We next present the results of LMMSE-IC receivers in Fig. 6.12 with the noise power set to $\sigma^2 = 0.285$. Several observations can be made from the figure. First, for the initial iteration, the Bayesian receiver will lead to higher output mutual information of the decoder, implying that the BER and FER performance is better than the conventional receiver. Second, in terms of number of iterations to achieve convergence, the improvement of the Bayesian receiver is not as much as the IC-MRC receiver case. This is because of the interference suppression capability of the LMMSE filter at the interference cancelation output. Even when there is relatively high residual interference in the conventional LMMSE-IC receiver, the LMMSE filter can suppress it effectively and provide a not-so-bad bit metric value to the decoder. This can also be verified by the small difference of the output mutual information values when the input mutual information is set to $I^{\mathrm{a}}(\mathrm{rec}) = 0$. Therefore, for implementation complexity consideration, the Bayesian LMMSE-IC receiver is more applicable for the punctured code and higher modulation schemes, as we have already indicated in Chapter 5, and further confirmed with the EXIT chart analysis result for

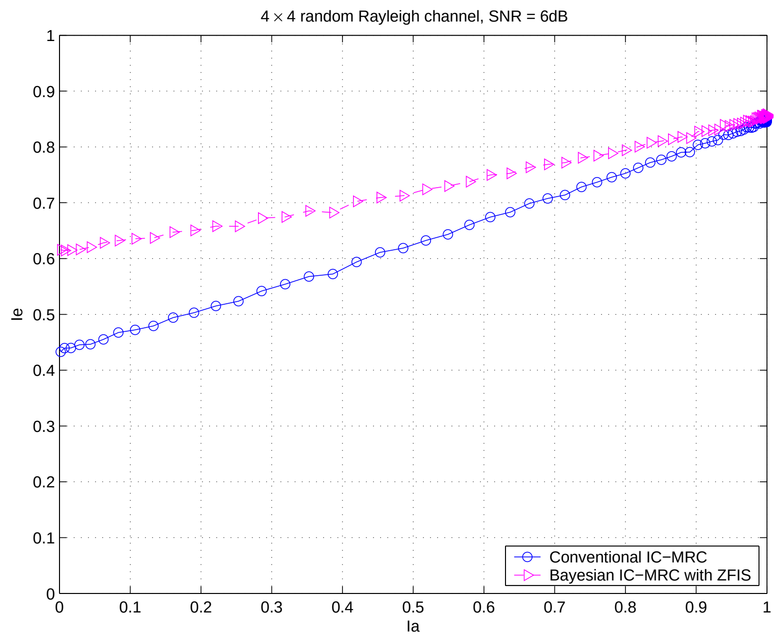


**Figure 6.6: Mutual information transfer function comparison of the conventional and Bayesian IC-MRC detectors. Random Rayleigh fading channel, QPSK modulation. Receive SNR = 6 dB.**

8PSK signals in Fig. 6.13.

## 6.4 Conclusions

The EXIT chart analysis of the Bayesian MMSE IC-MRC turbo receiver is presented and compared with the conventional turbo receivers. Our extensive results show that the Bayesian MMSE IC-MRC turbo receiver has much higher output mutual information than the conventional turbo receivers, thus verifying its superior BER and FER performance shown in Chapter 5. Furthermore, the detector and decoder trajectories have shown that much fewer number of iterations is required by the Bayesian turbo receiver to achieve convergence.

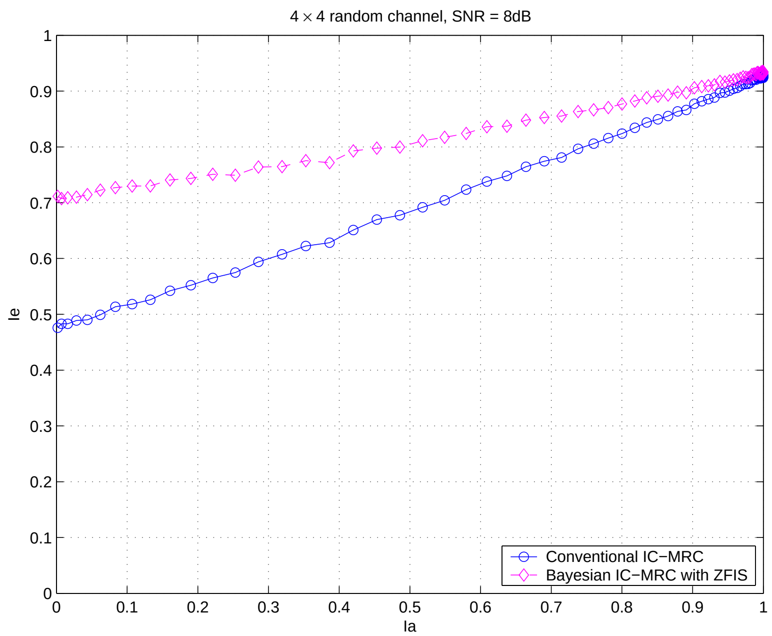


**Figure 6.7: Mutual information transfer function comparison of the conventional and Bayesian IC-MRC detectors. Random Rayleigh fading channel, QPSK modulation. Receive SNR = 8 dB.**

## Appendix - The $4 \times 4$ Static Channel Used in some EXIT Charts Generation

The $4 \times 4$ static channel used in some EXIT charts generation in this chapter was taken from [105], and is reproduced as below:

$$\begin{bmatrix} 0.926 - 1.187i & -0.24 - 0.535i & 1.305 + 0.184i & 0.483 - 0.852i \\ -0.432 - 0.235i & 0.448 + 0.122i & -0.32 - 0.007i & 0.507 + 0.417i \\ -0.211 - 0.877i & 0.649 + 0.294i & 0.316 - 0.209i & -0.969 - 0.312i \\ -0.198 - 0.688i & -1.054 + 0.14i & 0.44 + 0.371i & 0.948 - 0.304i \end{bmatrix}$$

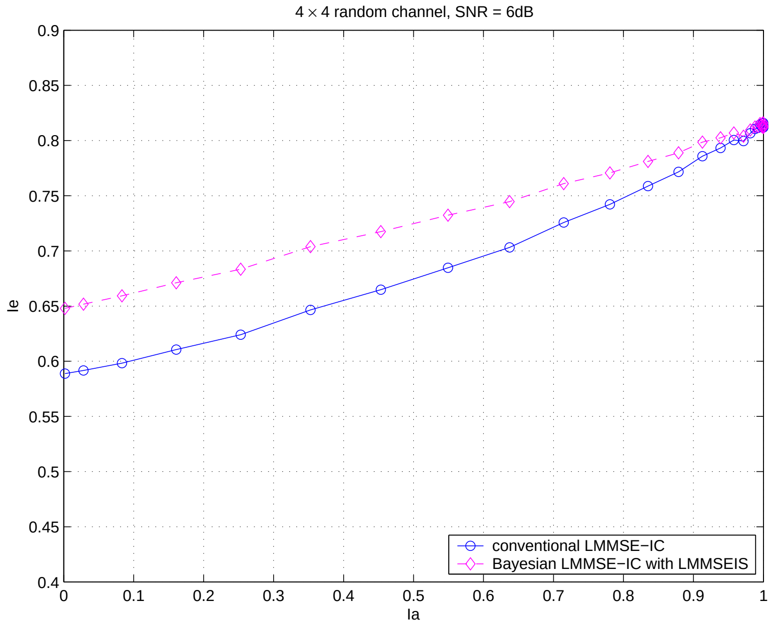


**Figure 6.8: Mutual information transfer function comparison of the conventional and Bayesian LMMSE-IC detectors. Random Rayleigh fading channel, QPSK modulation. Receive SNR = 6 dB.**

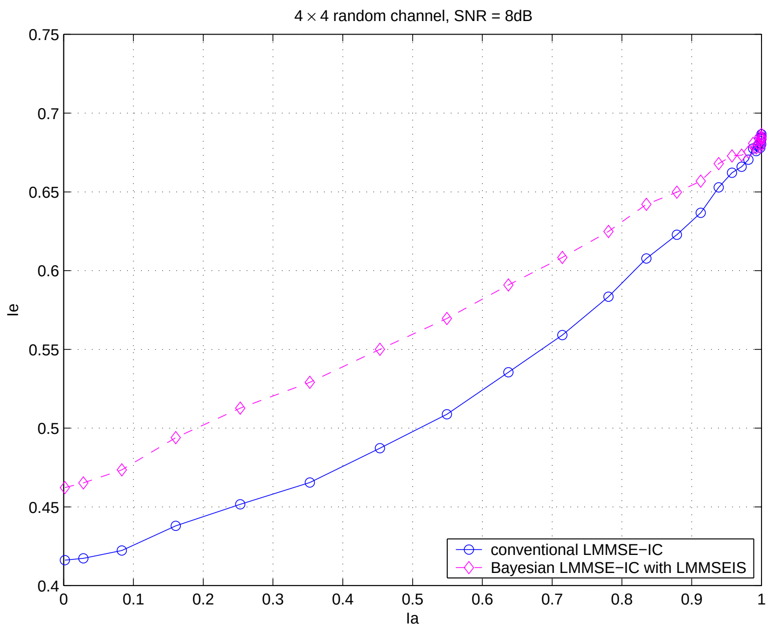


**Figure 6.9: Mutual information transfer function comparison of the conventional and Bayesian LMMSE-IC detectors. Random Rayleigh fading channel, 8PSK modulation. Receive SNR = 8 dB.**

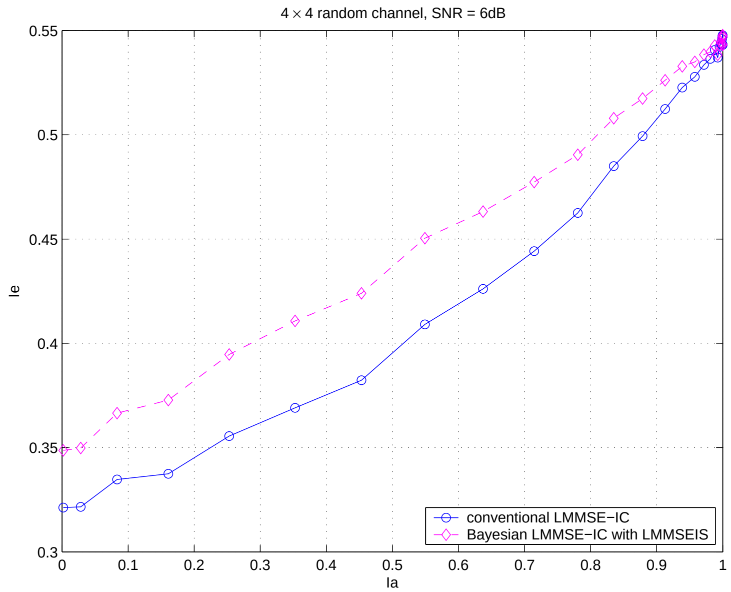


**Figure 6.10: Mutual information transfer function comparison of the conventional and Bayesian LMMSE-IC detectors. Random Rayleigh fading channel, 8PSK, receive SNR = 6 dB.**

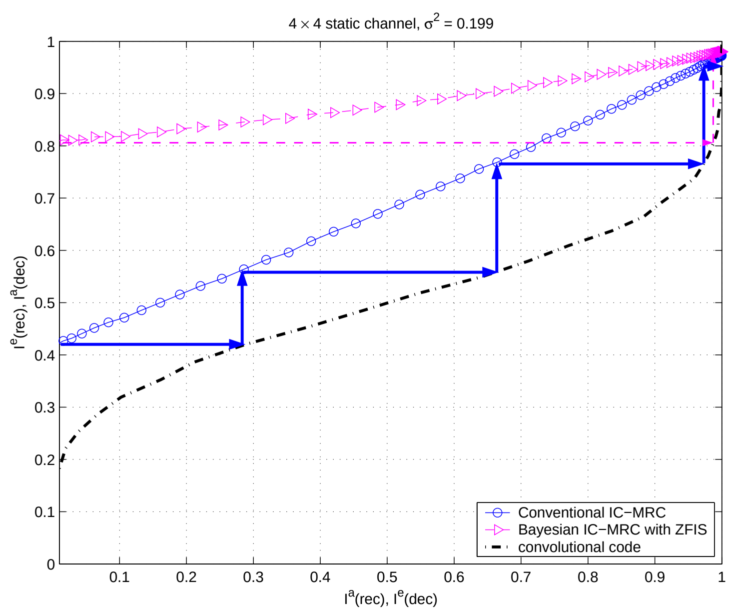


**Figure 6.11: Mutual information transfer function comparison of the conventional and Bayesian IC-MRC turbo receivers, and decoding path for the turbo receivers with $K = 3$ CC. Static channel, QPSK, $\sigma^2 = 0.199$.**

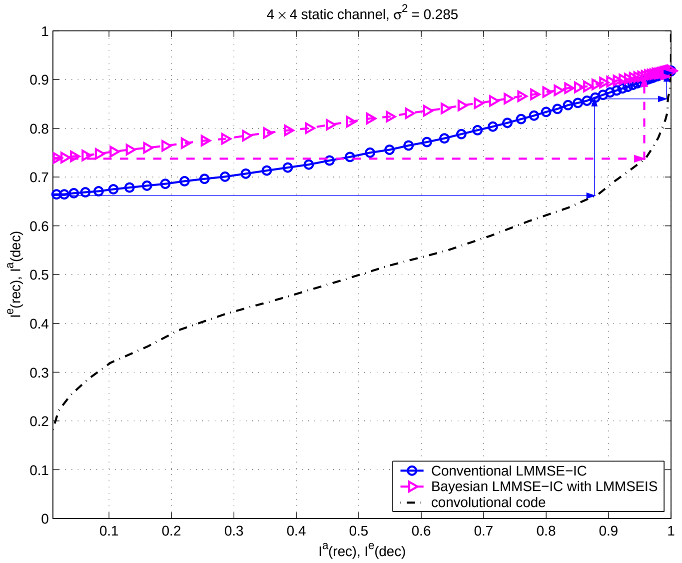


**Figure 6.12: Mutual information transfer function comparison of the conventional and Bayesian LMMSE-IC turbo receivers, and decoding path for the turbo receivers with $R_c = \frac{1}{2}$ $K = 3$ CC. Static channel, QPSK, $\sigma^2 = 0.285$.**

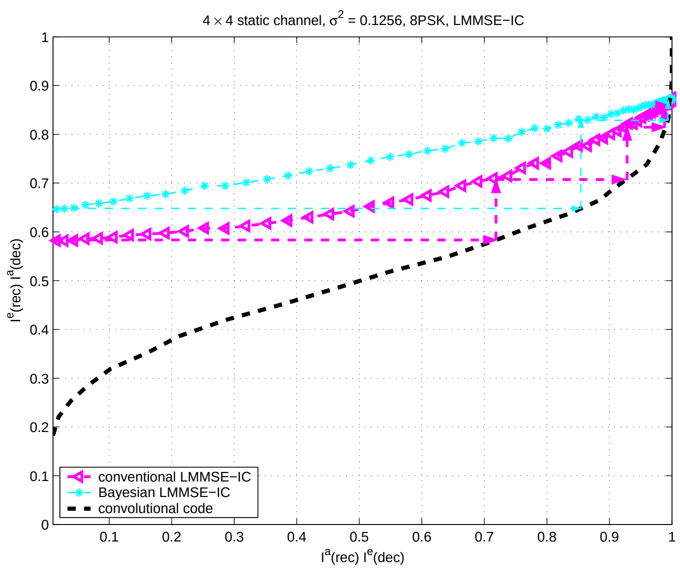


**Figure 6.13: Mutual information transfer function comparison of the conventional and Bayesian LMMSE-IC turbo receivers, and decoding path for the turbo receivers with $R_c = \frac{1}{2}$ $K = 3$ CC. Static channel, 8PSK, $\sigma^2 = 0.1256$.**

# Chapter 7

# Training Signal Design and Channel Estimation

Receivers based on coherent detection need the channel information, which makes channel estimation essential in the MIMO detector. In this chapter, we will focus on training sequence assisted channel estimation for packet-based MIMO-OFDM for WLAN application. Due to the low mobility in this network, a quasi-static channel can be assumed for each packet. Training signals are thus needed only at the beginning of the packet. Training signals that are transmitted at the beginning of a packet are sometimes called "preambles".

Following the linear matrix algebraic model defined in Chapter 3, we will first study the frequency-domain channel estimation (FDCE). Based on the minimum mean squared error (MMSE) criteria for least squares (LS) channel estimation, we will define the basic orthogonal training signals (OTS) structure and derive the LS and LMMSE channel estimation algorithms in Section 7.2. With this OTS structure, LS and LMMSE channel estimation can be obtained by linear matrix filtering of the received frequency domain signal with fixed parameters. Therefore, it is very attractive for practical implementation.

The preamble length in the OTS scheme, however, should be at least equal to the number of transmit antennas. The transmission efficiency can thus be severely degraded especially when the number of active transmit antennas is large. We hence also propose in Section 7.2 a switched subcarrier preamble scheme (SSPS) in which the transmit antennas are divided into subsets, and OTS are transmitted in alternative subsets of subcarriers in each group. With the SSPS, only subsets of the frequency domain channel estimates can be obtained directly from the preambles. We therefore propose three interpolation algorithms, namely, linear interpolation which assumes high correlation between neighboring subcarriers only, LMMSE interpolation which will make use of a more realistic channel correlation in the different subcarriers, and DFT-based LS interpolation which assumes a fixed number of multipaths in the MIMO radio channel as well as making use of the time- and frequency-domain relationship of the channel parameters. The performance of the different frequency-domain channel estimation algorithms will be presented in Section 7.2.4.

Another way to reduce the overhead length of preambles is to make use of time-domain channel estimation algorithms. Section (TDCE) 7.3 is dedicated to the preamble designs for TDCE. Based on the frame-work derived in [116], we prove that in addition to the CDD-based preamble (CDDP) sequence proposed in [116], the SSPS with transmission of only its first OFDM symbol, also satisfies the criteria of simple channel estimation and minimum MSE. Compared with the CDDP sequence, the SSPS has the advantage of smaller PAPR which is easily achieved with the reduced number of active subcarriers in one OFDM symbol. Finally in Section 7.4, we conclude the chapter.

## 7.1 Contributions of this Chapter

The main contributions of this chapter are:

- Developed the optimal frequency domain training sequence design;
- Analyzed the MSE performance of the LS and LMMSE FDCE, explicitly proved the requirement for the mismatched channel correlation and SNR for robust channel estimation;
- Proposed a simple training signal design for optimal TDCE.

## 7.2 Preamble Design for Frequency-Domain Channel Estimation

### 7.2.1 The LS Channel Estimation

Recall from Chapter 3 that when no precoding is considered, the frequency-domain received signal at each subcarrier is written as

$$\boldsymbol{\mathcal{R}}_k = \boldsymbol{\mathcal{W}}_k \boldsymbol{\mathcal{S}}_k + \boldsymbol{\mathcal{N}}_k, \tag{7.1}$$

where $k$ is the subcarrier index, $\boldsymbol{\mathcal{R}}$ and $\boldsymbol{\mathcal{N}}$ are $n_R \times 1$ vectors representing the received signal and the AWGN at the $n_R$ receive antennas, $\boldsymbol{\mathcal{S}}$ is a $n_T \times 1$ vector denoting the transmitted training signal at the $n_T$ antennas, and $\boldsymbol{\mathcal{W}}$ is the $n_R \times n_T$ channel matrix.

Excluding the AWGN term in (7.1), we observe a linear relation between the channel parameters and the received signals. For training sequence assisted channel estimation, solving this linear equation will lead to the LS channel estimates. In order to do this, pilot signal with length $n_T$ OFDM symbols is needed. The received signal at subcarrier $k$ during the training period of $n_T$ OFDM symbols can then be written as:

$$\underline{\boldsymbol{\mathcal{R}}}_k = \boldsymbol{\mathcal{W}}_k \underline{\boldsymbol{\mathcal{S}}}_k + \underline{\boldsymbol{\mathcal{N}}}_k, \tag{7.2}$$

where $\underline{\boldsymbol{\mathcal{R}}}_k = \begin{bmatrix} \boldsymbol{\mathcal{R}}_{k,1} & \boldsymbol{\mathcal{R}}_{k,2} & \cdots & \boldsymbol{\mathcal{R}}_{k,n_T} \end{bmatrix}$ is an $n_R \times n_T$ matrix representing the received signal at the $n_R$ antennas, subcarrier $k$ during the training period, $\underline{\boldsymbol{\mathcal{S}}}_k = \begin{bmatrix} \boldsymbol{\mathcal{S}}_{k,1} & \boldsymbol{\mathcal{S}}_{k,2} & \cdots & \boldsymbol{\mathcal{S}}_{k,n_T} \end{bmatrix}$ is an $n_T \times n_T$ square matrix representing the training signals at subcarrier $k$ with a length of $n_T$ OFDM symbols, and $\underline{\boldsymbol{\mathcal{N}}}_k = \begin{bmatrix} \boldsymbol{\mathcal{N}}_{k,1} & \boldsymbol{\mathcal{N}}_{k,2} & \cdots & \boldsymbol{\mathcal{N}}_{k,n_T} \end{bmatrix}$ of size $n_R \times n_T$ represents the AWGN.

The LS channel estimates are then obtained by right multiplying $\underline{\boldsymbol{\mathcal{S}}}_k^{-1}$ with $\underline{\boldsymbol{\mathcal{R}}}_k$, as

$$\hat{\boldsymbol{\mathcal{W}}}_{k,LS} = \underline{\boldsymbol{\mathcal{R}}}_k \underline{\boldsymbol{\mathcal{S}}}_k^{-1}. \tag{7.3}$$

As long as $\underline{\boldsymbol{\mathcal{S}}}_k$ is a non-singular matrix, $\underline{\boldsymbol{\mathcal{S}}}_k^{-1}$ exists. Furthermore, computation of $\underline{\boldsymbol{\mathcal{S}}}_k^{-1}$ can be done off-line. Channel estimation is then just a linear combination of the received signals at the different antennas.

In order to derive the optimal $\underline{\boldsymbol{\mathcal{S}}}_k$, we first look at the MSE of the LS estimates, written as

$$\begin{aligned}
\text{MSE} &= \mathcal{E}(||\hat{\boldsymbol{\mathcal{W}}}_{k,LS} - \boldsymbol{\mathcal{W}}_k||^2) \\
&= \mathcal{E}(||\underline{\boldsymbol{\mathcal{N}}}_k \underline{\boldsymbol{\mathcal{S}}}_k^{-1}||^2) \\
&= \mathcal{E}(\sum_{n_r=1}^{n_R} |\underline{\boldsymbol{\mathcal{N}}}_{k,n_r} \underline{\boldsymbol{\mathcal{S}}}_k^{-1}|^2) \\
&= \mathcal{E}\left(\sum_{n_r=1}^{n_R} \underline{\boldsymbol{\mathcal{N}}}_{k,n_r} \underline{\boldsymbol{\mathcal{S}}}_k^{-1} \underline{\boldsymbol{\mathcal{S}}}_k^{-H} \underline{\boldsymbol{\mathcal{N}}}_{k,n_r}^H\right) \\
&= \sum_{n_r=1}^{n_R} \mathcal{E}\left[\text{trace}\left(\underline{\boldsymbol{\mathcal{N}}}_{k,n_r}^H \underline{\boldsymbol{\mathcal{N}}}_{k,n_r} \underline{\boldsymbol{\mathcal{S}}}_k^{-1} \underline{\boldsymbol{\mathcal{S}}}_k^{-H}\right)\right] \\
&= 2\sigma^2 \sum_{n_r=1}^{n_R} \text{trace}\left(\underline{\boldsymbol{\mathcal{S}}}_k^{-1} \underline{\boldsymbol{\mathcal{S}}}_k^{-H}\right) \\
&= 2\sigma^2 n_R \, \text{trace}\left(\underline{\boldsymbol{\mathcal{S}}}_k^{-1} \underline{\boldsymbol{\mathcal{S}}}_k^{-H}\right), \qquad (7.4)
\end{aligned}$$

where $\underline{\boldsymbol{\mathcal{N}}}_{k,n_r}$ denotes the $n_r$th row of $\underline{\boldsymbol{\mathcal{N}}}_k$.

Performing eigen-decomposition on $\underline{\boldsymbol{\mathcal{S}}}_k$ as $\underline{\boldsymbol{\mathcal{S}}}_k = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^H$, where $\mathbf{U}\mathbf{U}^H = \mathbf{U}^H\mathbf{U} = \mathbf{I}$, $\Lambda = \text{diag}\left(\lambda_1,\ \lambda_2,\ \cdots,\ \lambda_{n_T}\right)$ with $\lambda_i$ being the $i$th eigenvalue of $\underline{\boldsymbol{\mathcal{S}}}_k$, we have $\underline{\boldsymbol{\mathcal{S}}}_k^{-1} = \mathbf{U}\boldsymbol{\Lambda}^{-1}\mathbf{U}^H$, $\underline{\boldsymbol{\mathcal{S}}}_k^{-H} = \mathbf{U}\boldsymbol{\Lambda}^{-H}\mathbf{U}^H$, and hence

$$\begin{aligned}
\text{MSE} &= 2\sigma^2 n_R \, \text{trace}\left(\mathbf{U}\Lambda^{-1}\mathbf{U}^H\mathbf{U}\Lambda^{-H}\mathbf{U}^H\right) \\
&= 2\sigma^2 n_R \, \text{trace}\left(\boldsymbol{\Lambda}^{-1}\boldsymbol{\Lambda}^{-H}\right) \\
&= 2\sigma^2 n_R \sum_{n_t=1}^{n_T} \frac{1}{|\lambda_{n_t}|^2}. \qquad (7.5)
\end{aligned}$$

Therefore, minimum MSE is obtained when $\sum_{n_t=1}^{n_T} \frac{1}{|\lambda_{n_t}|^2}$ is minimized. This is achieved when $\frac{1}{|\lambda_{n_t}|^2} = \text{constant}$, $n_t = 1,\ \cdots,\ n_T$.

As

$$\begin{aligned}
\sum_{n_t=1}^{n_T} |\lambda_{n_t}|^2 &= \text{trace}\left(\boldsymbol{\Lambda}^H\boldsymbol{\Lambda}\right) \\
&= \text{trace}\left(\mathbf{U}\boldsymbol{\Lambda}^H\mathbf{U}^H\mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^H\right) \\
&= \text{trace}\left(\underline{\boldsymbol{\mathcal{S}}}_k^H\underline{\boldsymbol{\mathcal{S}}}_k\right) \\
&= n_T^2,
\end{aligned}$$

when we use constant modulus modulated training signals and the total transmission power per subcarrier is $n_T$, the minimum MSE is obtained when $|\lambda_{n_t}|^2 = n_T$, $n_t = 1,\ 2,\ \cdots,\ n_T$. This can be obtained by the following training signals.

**Orthogonal Training Signals (OTS)** The OTS can be obtained by extending a preamble sequence designed for a single-transmit single-receive system, as

$$\underline{\boldsymbol{\mathcal{S}}}_k = T_k \underline{\boldsymbol{\mathcal{M}}}, \tag{7.6}$$

where $T_k$ is the constant modulus training signal at the $k$th subcarrier for single-transmit single-receive OFDM system with $|T_k|^2 = 1$, and $\underline{\boldsymbol{\mathcal{M}}}$ is an $n_T \times n_T$ orthogonal matrix satisfying

$$\underline{\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{M}}}^H = \underline{\boldsymbol{\mathcal{M}}}^H \underline{\boldsymbol{\mathcal{M}}} = n_T \mathbf{I}.$$

**Remark 1** In order to have the same transmission power at each antenna, so as to have the same dynamic range requirement for the power amplifiers, the orthogonal matrix $\underline{\boldsymbol{\mathcal{M}}}$ should have its element taking values of $\exp(j\theta_i)$, with $\{\theta_i\}$ being some discrete phase values taken from $[0, 2\pi)$. For example, when $n_T$ is 1, 2, 4, or a multiple of 4, $\underline{\boldsymbol{\mathcal{M}}}$ can be set to the Walsh Hadamard matrix of size $n_T$. Otherwise, the $n_T \times n_T$ DFT matrix can be used.

**Remark 2** One special case for the OTS is $\underline{\boldsymbol{\mathcal{M}}} = \sqrt{n_T}\mathbf{I}_{n_T}$. This implies that same training signal is transmitted from only one antenna per OFDM symbol. The LS MIMO channel estimation is very simple in this case as it falls back to the single-transmit antenna channel estimation problem. On the other hand, it involves antenna and RF circuits switching on and off in a very short time interval, e.g., 4 microsecond ($4\mu S$) for the IEEE 802.11n systems [4], hence the performance may get degraded if the RF circuits can not get stable within such short interval. This special case of preamble design is sometimes referred as *Switched Antenna Preamble Scheme (SAPS)*.

**Remark 3** We can easily see that in this OTS preamble design, each transmit antenna uses the same preamble with a pre-defined phase rotation. Therefore, the properties of the single antenna

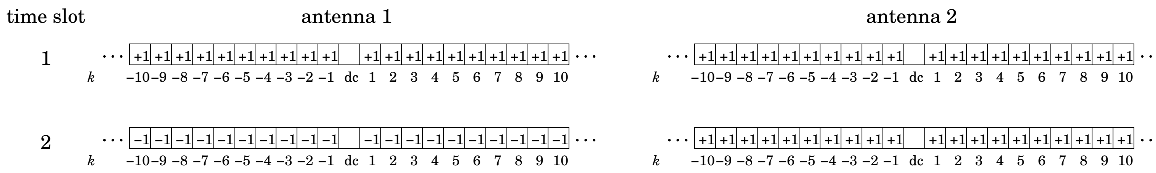


**Figure 7.1: Orthogonal training sequence design for 2 transmit antennas.**

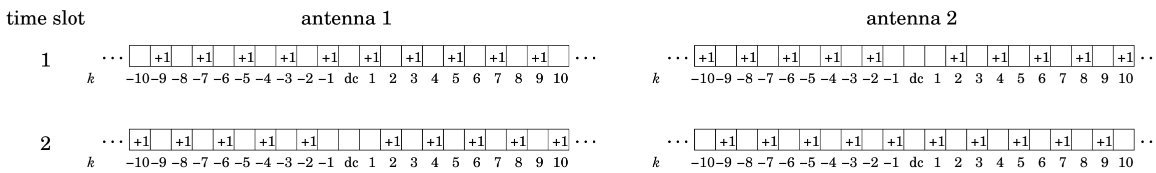


**Figure 7.2: Switched subcarrier preamble scheme for 2 transmit antennas**

OFDM training signals are maintained. Moreover, MIMO channel estimation for each subcarrier is obtained by the same linear matrix filter. The implementation is thus very simple.

In Figure 7.1, the orthogonal training signal design for $n_T = 2$ is illustrated, in which $k$ denotes the subcarrier index, $+1$ means that the original SISO pilot signal $T_k$ is transmitted in this subcarrier, and $-1$ denotes that $-T_k$ is transmitted. In the first time slot (OFDM symbol), both antennas transmit $T_k$. In the second time slot, antenna 1 transmits $-T_k$ and antenna 2 transmits $T_k$ so as to obtain the orthogonality.

**Switched Subcarrier Preamble Scheme (SSPS)** The SSPS can be taken as another special case of OTS. Same as SAPS, training signal is transmitted *only once* per transmit antenna and per subcarrier. While in SAPS, preamble transmission is "switched" from one transmit antenna to the next, in SSPS, training signal is "switched" from one subcarrier subset to the next subcarrier subset, over all transmit antennas. The training signal in the $k$th subcarrier is now

$$\underline{\boldsymbol{\mathcal{S}}}_k = \sqrt{n_T}\mathrm{diag}\left(S_{k,1},\ S_{k,2},\ \cdots,\ S_{k,n_T}\right). \tag{7.7}$$

In Figure 7.2, the SSPS is illustrated for $n_T = 2$ systems.

Comparing the two "optimal" frequency domain training signal schemes, SSPS has the following advantages:

- simpler computation in channel estimation as only scalar operation is needed in order to calculate the LS channel estimates;
- lower PAPR as fewer number of subcarriers are active per OFDM symbol per transmit antenna.

On the other hand, the OTS scheme requires the same and fixed matrix filter for each subcarrier, hence facilitates better circuit reuse in channel estimation.

The minimum MSE per subcarrier is

$$\min(\text{MSE}) = 2\sigma^2 n_R. \tag{7.8}$$

The minimum MSE for all the transmit-receive antenna pairs and all the $N$ subcarriers is thus

$$\min(\text{MSE}_{\text{LS}}) = 2\sigma^2 n_R\, N. \tag{7.9}$$

### 7.2.2 The Frequency Domain LMMSE Channel Estimation

The LS channel estimates are obtained by using only the knowledge of the training signals, which can be further improved by making use of the correlation information in frequency and spatial domains. If we assume that the spatial domain channel response is uncorrelated, and that the power delay profile has the same statistical properties for all the single-transmit single-receive channels corresponding to each transmit-receive antenna pair in the MIMO system, we can then apply a frequency-domain LMMSE filter to the LS channel estimates in (7.3), as [117]:

$$\hat{\mathbf{H}}_{n_r,n_t,\text{LMMSE}} = \mathbf{R}_{HH}\left(\mathbf{R}_{HH} + \frac{\beta}{\text{SNR}}\mathbf{I}\right)^{-1}\hat{\mathbf{H}}_{n_r,n_t,\text{LS}}, \tag{7.10}$$

where subscripts $n_r$ and $n_t$ denote the indices of the receive and transmit antennas, respectively, SNR is the per subcarrier per transmit antenna signal to noise ratio of the training signals, $\hat{\mathbf{H}}_{n_r,n_t,\text{LS}}$ denotes the $N \times 1$ LS channel estimates corresponding to transmit-receive antenna pair $(n_t, n_r)$, $\beta$ is a constant depending on the training signal's constellation. As given in [117], $\beta = 1$ if MPSK training signals are used. $\mathbf{R}_{HH} = \mathcal{E}(\mathbf{H}_{n_r n_t}\mathbf{H}_{n_r n_t}^H)$ is the channel autocorrelation matrix which is independent of $n_r$ and $n_t$ when we assume the same statistical properties for each single-transmit single-receive antenna channel in the MIMO-OFDM system.

As shown in Chapter 3,

$$\mathbf{H} = \mathbf{F}\mathbf{h},$$

where elements of $\mathbf{F}$ is $f_{k,n} = \exp\left(-j\frac{2\pi}{N}kn\right) = W_N^{kn}$, $W_N = \exp\left(-j\frac{2\pi}{N}\right)$, and $\mathbf{h}$ is the time domain channel vector with $\mathbf{h} = [h_0 \ h_1 \ \cdots \ h_{L-1} \ 0 \ \cdots \ 0]^T$. Assuming WSSUS channel, i.e., $\mathcal{E}\left\{h_i h_j^*\right\} = \mathcal{E}\left\{|h_i|^2\right\}\delta(i-j)$, we have

$$\begin{aligned}
\mathbf{R}_{HH} &= \mathcal{E}(\mathbf{H}\mathbf{H}^H) = \mathcal{E}\left(\mathbf{F}\mathbf{h}\mathbf{h}^H\mathbf{F}^H\right) \\
&= \mathbf{F}\mathcal{E}\left(\mathbf{h}\mathbf{h}^H\right)\mathbf{F}^H \\
&= \mathbf{F}\ \text{diag}\underbrace{\left(\mathcal{E}\left(|h_0|^2\right),\ \mathcal{E}\left(|h_1|^2\right),\ \cdots,\ \mathcal{E}\left(|h_{L-1}|^2\right),\ 0,\ \cdots,\ 0\right)}_{\Lambda_{\mathbf{h}}}\mathbf{F}^H. \qquad (7.11)
\end{aligned}$$

Therefore, if the channel PDF information is available, $\mathbf{R}_{HH}$ can be computed and used in LMMSE channel estimation.

The MSE of LMMSE channel estimation, when perfect PDF information is used, can be derived as follows:

$$\begin{aligned}
\text{MSE}_{\text{LMMSE}} &= \sum_{n_r=1}^{n_R}\sum_{n_t=1}^{n_T}\mathcal{E}\left\{\left\|\mathbf{H}_{n_r,n_t} - \hat{\mathbf{H}}_{n_r,n_t,\text{LMMSE}}\right\|^2\right\} \\
&= n_T\ n_R\mathcal{E}\left\{\left\|\mathbf{H} - \hat{\mathbf{H}}_{\text{LMMSE}}\right\|^2\right\} \\
&= n_T\ n_R\mathcal{E}\left\{\left(\mathbf{H} - \hat{\mathbf{H}}_{\text{LMMSE}}\right)^H\left(\mathbf{H} - \hat{\mathbf{H}}_{\text{LMMSE}}\right)\right\} \\
&= n_T\ n_R\mathcal{E}\left\{\left(\mathbf{H}^H - \mathbf{R}^H\mathbf{G}_{\text{LMMSE}}^H\right)\left(\mathbf{H} - \mathbf{G}_{\text{LMMSE}}\mathbf{R}\right)\right\} \\
&= n_T\ n_R\mathcal{E}\left\{\text{trace}\left[\left(\mathbf{H} - \mathbf{G}_{\text{LMMSE}}\mathbf{R}\right)\left(\mathbf{H}^H - \mathbf{R}^H\mathbf{G}_{\text{LMMSE}}^H\right)\right]\right\} \\
&= n_T\ n_R\text{trace}\left[\mathbf{R}_{HH} - \mathbf{G}_{\text{LMMSE}}\mathbf{R}_{RR}\mathbf{G}_{\text{LMMSE}}^H\right] \\
&= n_T\ n_R\text{trace}\left[\mathbf{R}_{HH} - \mathbf{G}_{\text{LMMSE}}\left(\mathbf{X}\mathbf{R}_{HH}\mathbf{X}^H + 2\sigma^2\mathbf{I}\right)\mathbf{G}_{\text{LMMSE}}^H\right],
\end{aligned}$$

where $\mathbf{G}_{\text{LMMSE}}$ is the frequency domain LMMSE channel estimation filter defined in (7.10), $\mathbf{R}_{RR} = \mathcal{E}\left(\mathbf{R}\mathbf{R}^H\right) = \mathcal{E}\left(\mathbf{X}\mathbf{R}_{HH}\mathbf{X}^H + 2\sigma^2\mathbf{I}\right)$ with $\mathbf{X} = \text{diag}\left(x_0,\ x_1,\ \cdots,\ x_{N-1}\right)$ being the frequency domain training signals.

Applying the results in (7.10) and (7.11), we have for MPSK training signals,

$$
\begin{aligned}
& \mathbf{G}_{\text{LMMSE}} \left(\mathbf{X}\mathbf{R}_{HH}\mathbf{X}^{H}\right) \mathbf{G}_{\text{LMMSE}}^{H} \\
= \;& \mathbf{R}_{HH}\left(\mathbf{R}_{HH}+\frac{1}{\text{SNR}}\mathbf{I}\right)^{-1}\mathbf{X}^{H}\mathbf{X}\mathbf{R}_{HH}\mathbf{X}^{H}\mathbf{X}\left(\mathbf{R}_{HH}+\frac{1}{\text{SNR}}\mathbf{I}\right)^{-1}\mathbf{R}_{HH} \\
= \;& \mathbf{R}_{HH}\left(\mathbf{R}_{HH}+\frac{1}{\text{SNR}}\mathbf{I}\right)^{-1}\mathbf{R}_{HH}\left(\mathbf{R}_{HH}+\frac{1}{\text{SNR}}\mathbf{I}\right)^{-1}\mathbf{R}_{HH} \\
= \;& \mathbf{F}\Lambda_{\mathbf{h}}\mathbf{F}^{H}\left(\mathbf{F}\Lambda_{\mathbf{h}}\mathbf{F}^{H}+\frac{1}{\text{SNR}}\mathbf{I}\right)^{-1}\mathbf{F}\Lambda_{\mathbf{h}}\mathbf{F}^{H}\left(\mathbf{F}\Lambda_{\mathbf{h}}\mathbf{F}^{H}+\frac{1}{\text{SNR}}\mathbf{I}\right)^{-1}\mathbf{F}\Lambda_{\mathbf{h}}\mathbf{F}^{H} \\
= \;& \mathbf{F}\Lambda_{\mathbf{h}}\mathbf{F}^{H}\left(\mathbf{F}\Lambda_{\mathbf{h}}\mathbf{F}^{H}+\frac{1}{\text{SNR}}\mathbf{F}\mathbf{F}^{H}\right)^{-1}\mathbf{F}\Lambda_{\mathbf{h}}\mathbf{F}^{H}\left(\mathbf{F}\Lambda_{\mathbf{h}}\mathbf{F}^{H}+\frac{1}{\text{SNR}}\mathbf{F}\mathbf{F}^{H}\right)^{-1}\mathbf{F}\Lambda_{\mathbf{h}}\mathbf{F}^{H} \\
= \;& \mathbf{F}\Lambda_{\mathbf{h}}\left(\Lambda_{\mathbf{h}}+\frac{1}{\text{SNR}}\mathbf{I}\right)^{-1}\Lambda_{\mathbf{h}}\left(\Lambda_{\mathbf{h}}+\frac{1}{\text{SNR}}\mathbf{I}\right)^{-1}\Lambda_{\mathbf{h}}\mathbf{F}^{H},
\end{aligned}
$$

and

$$
\mathbf{G}_{\text{LMMSE}}\mathbf{G}_{\text{LMMSE}}^{H} = \mathbf{F}\Lambda_{\mathbf{h}}\left(\Lambda_{\mathbf{h}}+\frac{1}{\text{SNR}}\mathbf{I}\right)^{-2}\Lambda_{\mathbf{h}}\mathbf{F}^{H}.
$$

Therefore,

$$
\begin{aligned}
& \text{trace}\left[\mathbf{R}_{HH}-\mathbf{G}_{\text{LMMSE}}\left(\mathbf{X}\mathbf{R}_{HH}\mathbf{X}^{H}+2\sigma^{2}\mathbf{I}\right)\mathbf{G}_{\text{LMMSE}}^{H}\right] \\
= \;& \sum_{l=0}^{L-1}\left[|h_l|^2-\frac{|h_l|^6}{\left(|h_l|^2+\frac{1}{\text{SNR}}\right)^2}-\frac{2\sigma^2|h_l|^4}{\left(|h_l|^2+\frac{1}{\text{SNR}}\right)^2}\right] \\
= \;& \sum_{l=0}^{L-1}\frac{\frac{|h_l|^2}{\text{SNR}}}{|h_l|^2+\frac{1}{\text{SNR}}} \\
= \;& \sum_{l=0}^{L-1}\frac{|h_l|^2}{1+\frac{|h_l|^2}{2\sigma^2}}
\end{aligned}
$$

and the MSE of the channel estimates is

$$
\begin{aligned}
\text{MSE}_{\text{LMMSE}} &= n_T\, n_R\sum_{l=0}^{L-1}\frac{|h_l|^2}{1+\frac{|h_l|^2}{2\sigma^2}} && (7.12)\\
&< n_T\, n_R\sum_{l=0}^{L-1}2\sigma^2 && (7.13)\\
&= Ln_T\, n_R\, 2\sigma^2. && (7.14)
\end{aligned}
$$

As long as $Ln_T \leq N$, the LMMSE channel estimates have smaller MSE than the LS estimates.

**MSE with Channel Correlation Mismatch**

Equation 7.12) gives the minimum MSE when the exact correlation matrix $\mathbf{R}_{HH}$ is known to the LMMSE channel estimator. Now we will look at the MSE when there is mismatch between the $\mathbf{R}_{HH}$ used in channel estimation and the exact one.

Denoting the mismatched correlation matrix as $\tilde{\mathbf{R}}_{HH} = \mathbf{F}\tilde{\Lambda}_{\mathbf{h}}\mathbf{F}^H$, we then have the LMMSE filter expressed as

$$\begin{aligned}\tilde{\mathbf{G}}_{\text{LMMSE}} &= \tilde{\mathbf{R}}_{HH}\left(\tilde{\mathbf{R}}_{HH} + 2\sigma^2\mathbf{I}\right)^{-1}\mathbf{X}^H \\ &= \mathbf{F}\tilde{\Lambda}_{\mathbf{h}}\left(\tilde{\Lambda}_{\mathbf{h}} + 2\sigma^2\mathbf{I}\right)^{-1}\mathbf{F}^H\mathbf{X}^H,\end{aligned}$$

and

$$\begin{aligned}&\tilde{\mathbf{G}}_{\text{LMMSE}}\left(\mathbf{X}\mathbf{R}_{HH}\mathbf{X}^H + 2\sigma^2\mathbf{I}\right)\tilde{\mathbf{G}}_{\text{LMMSE}}^H \\ =\ &\mathbf{F}\tilde{\Lambda}_{\mathbf{h}}\left(\tilde{\Lambda}_{\mathbf{h}} + 2\sigma^2\mathbf{I}\right)^{-1}\mathbf{F}^H\mathbf{X}^H\left(\mathbf{X}\mathbf{F}\Lambda_{\mathbf{h}}\mathbf{F}^H X^H + 2\sigma^2\mathbf{F}\mathbf{F}^H\right)\mathbf{X}\mathbf{F}\left(\tilde{\Lambda}_{\mathbf{h}} + 2\sigma^2\mathbf{I}\right)^{-1}\tilde{\Lambda}_{\mathbf{h}}\mathbf{F}^H \\ =\ &\mathbf{F}\tilde{\Lambda}_{\mathbf{h}}\left(\tilde{\Lambda}_{\mathbf{h}} + 2\sigma^2\mathbf{I}\right)^{-1}\left(\Lambda_{\mathbf{h}} + 2\sigma^2\mathbf{I}\right)\left(\tilde{\Lambda}_{\mathbf{h}} + 2\sigma^2\mathbf{I}\right)^{-1}\tilde{\Lambda}_{\mathbf{h}}\mathbf{F}^H\end{aligned}$$

hence the mismatched MSE expressed as

$$\begin{aligned}\widetilde{\text{MSE}}_{\text{LMMSE}} &= n_T\, n_R \text{trace}\left[\mathbf{R}_{HH} - \tilde{\mathbf{G}}_{\text{LMMSE}}\left(\mathbf{X}\mathbf{R}_{HH}\mathbf{X}^H + 2\sigma^2\mathbf{I}\right)\tilde{\mathbf{G}}_{\text{LMMSE}}^H\right] \\ &= n_T\, n_R \text{trace}\left[\Lambda_{\mathbf{h}} - \tilde{\Lambda}_{\mathbf{h}}\left(\tilde{\Lambda}_{\mathbf{h}} + 2\sigma^2\mathbf{I}\right)^{-1}\left(\Lambda_{\mathbf{h}} + 2\sigma^2\mathbf{I}\right)\left(\tilde{\Lambda}_{\mathbf{h}} + 2\sigma^2\mathbf{I}\right)^{-1}\tilde{\Lambda}_{\mathbf{h}}\right] \\ &= n_T\, n_R \sum_{l=0}^{L-1}\left[\lambda_l - \frac{\tilde{\lambda}_l^2\left(\lambda_l^2 + 2\sigma^2\right)}{\left(\tilde{\lambda}_l + 2\sigma^2\right)^2}\right]\end{aligned}$$

We therefore have the difference between the mismatched MSE and the MMSE as

$$
\begin{aligned}
&\widetilde{\mathrm{MSE}}_{\mathrm{LMMSE}} - \mathrm{MSE}_{\mathrm{LMMSE}} \\
= \;& n_T\, n_R \sum_{l=0}^{L-1} \left[ \lambda_l - \frac{\tilde{\lambda}_l^2 \left(\lambda_l^2 + 2\sigma^2\right)}{\left(\tilde{\lambda}_l + 2\sigma^2\right)^2} - \frac{2\sigma^2 \lambda_l}{\lambda_l + 2\sigma^2} \right] \\
= \;& n_T\, n_R \sum_{l=0}^{L-1} \left[ \frac{\lambda_l^2}{\lambda_l + 2\sigma^2} - \frac{\tilde{\lambda}_l^2 \left(\lambda_l^2 + 2\sigma^2\right)}{\left(\tilde{\lambda}_l + 2\sigma^2\right)^2} \right] \\
= \;& n_T\, n_R \sum_{l=0}^{L-1} \left[ \frac{\lambda_l^2 \left(\tilde{\lambda}_l + 2\sigma^2\right)^2 - \tilde{\lambda}_l^2 \left(\lambda_l + 2\sigma^2\right)^2}{\left(\lambda_l + 2\sigma^2\right)\left(\tilde{\lambda}_l + 2\sigma^2\right)^2} \right] \\
= \;& n_T\, n_R \sum_{l=0}^{L-1} \left[ \frac{\left[\lambda_l \left(\tilde{\lambda}_l + 2\sigma^2\right) + \tilde{\lambda}_l \left(\lambda_l + 2\sigma^2\right)\right] \left[\lambda_l \left(\tilde{\lambda}_l + 2\sigma^2\right) - \tilde{\lambda}_l \left(\lambda_l + 2\sigma^2\right)\right]}{\left(\lambda_l + 2\sigma^2\right)\left(\tilde{\lambda}_l + 2\sigma^2\right)^2} \right] \\
= \;& n_T\, n_R \sum_{l=0}^{L-1} 4\sigma^2 \left[ \frac{\left[\lambda_l \tilde{\lambda}_l + \sigma^2 \left(\lambda_l + \tilde{\lambda}_l\right)\right] \left(\lambda_l - \tilde{\lambda}_l\right)}{\left(\lambda_l + 2\sigma^2\right)\left(\tilde{\lambda}_l + 2\sigma^2\right)^2} \right]
\end{aligned}
$$

Therefore, if $\lambda_l > \tilde{\lambda}_l$, the LMMSE channel estimates with mismatched channel correlation matrix will have higher MSE than the perfect case.

From [117], we have $h_l = Ce^{-\frac{l}{\tau_{\mathrm{rms}}}}$ for the exponential power delay profile[1]. We have have $\lambda_l = |h_l|^2 = C^2 e^{-\frac{2l}{\tau_{rms}}}$. If $\tau_{\mathrm{rms}} > \tilde{\tau}_{\mathrm{rms}}$, we have $\lambda_l > \tilde{\lambda}_l$, hence $\widetilde{\mathrm{MSE}}_{\mathrm{LMMSE}} > \mathrm{MSE}_{\mathrm{LMMSE}}$. That is, i.e., if a less correlated channel is assumed, the LMMSE channel estimates will suffer MSE degradation. On the other hand, if $\tau_{\mathrm{rms}} < \tilde{\tau}_{\mathrm{rms}}$, i.e., we use the channel correlation matrix with less correlation, no MSE degradation will be encountered.

[1] In this case, an infinite length multipath channel is assumed, hence when only considering $L$ taps in the LMMSE channel estimation, there will be some energy leakage to the remaining paths. The significant part of the signal power, however, can still be captured in the first $L$ multipaths, especially when $\tau_{rms}$ is small.

**MSE With SNR Mismatch**

If the channel correlation matrix is perfectly known, but the estimated SNR value has some discrepancy from the actual one, we have the LMMSE filter written as

$$\begin{aligned}\tilde{\mathbf{G}}_{\mathrm{LMMSE}} &= \mathbf{R}_{HH}\left(\mathbf{R}_{HH}+2\tilde{\sigma}^2\mathbf{I}\right)^{-1}\mathbf{X}^H\\ &= \mathbf{F}\Lambda_{\mathbf{h}}\left(\Lambda_{\mathbf{h}}+2\tilde{\sigma}^2\mathbf{I}\right)^{-1}\mathbf{F}^H\mathbf{X}^H,\end{aligned}$$

and the SNR-mismatched MSE expressed as

$$\begin{aligned}\widetilde{\mathrm{MSE}}_{\mathrm{LMMSE}} &= n_T\, n_R \mathrm{trace}\left[\mathbf{R}_{HH}-\tilde{\mathbf{G}}_{\mathrm{LMMSE}}\left(\mathbf{X}\mathbf{R}_{HH}\mathbf{X}^H+2\sigma^2\mathbf{I}\right)\tilde{\mathbf{G}}_{\mathrm{LMMSE}}^H\right]\\ &= n_T\, n_R \mathrm{trace}\left[\Lambda_{\mathbf{h}}-\Lambda_{\mathbf{h}}\left(\Lambda_{\mathbf{h}}+2\tilde{\sigma}^2\mathbf{I}\right)^{-1}\left(\Lambda_{\mathbf{h}}+2\sigma^2\mathbf{I}\right)\left(\Lambda_{\mathbf{h}}+2\tilde{\sigma}^2\mathbf{I}\right)^{-1}\Lambda_{\mathbf{h}}\right]\\ &= n_T\, n_R\sum_{l=0}^{L-1}\left[\lambda_l-\frac{\lambda_l^2}{\left(\lambda_l+2\tilde{\sigma}^2\right)^2}\left(\lambda_l+2\sigma^2\right)\right]\\ &= n_T\, n_R\sum_{l=0}^{L-1}\lambda_l\left[1-\frac{\lambda_l\left(\lambda_l+2\sigma^2\right)}{\left(\lambda_l+2\tilde{\sigma}^2\right)^2}\right].\end{aligned}$$

If the estimated SNR is larger than the actual one, i.e., $\tilde{\sigma}^2<\sigma^2$, we have

$$\begin{aligned}1-\frac{\lambda_l\left(\lambda_l+2\sigma^2\right)}{\left(\lambda_l+2\tilde{\sigma}^2\right)^2} &< 1-\frac{\lambda_l\left(\lambda_l+2\sigma^2\right)}{\left(\lambda_l+2\sigma^2\right)^2}\\ &= 1-\frac{\lambda_l}{\left(\lambda_l+2\sigma^2\right)}\\ &= \frac{2\sigma^2}{\left(\lambda_l+2\sigma^2\right)},\end{aligned}$$

and

$$\begin{aligned}\widetilde{\mathrm{MSE}}_{\mathrm{LMMSE}} &< n_T\, n_R\sum_{l=0}^{L-1}\lambda_l\frac{2\sigma^2}{\left(\lambda_l+2\sigma^2\right)}\\ &= n_T\, n_R\sum_{l=0}^{L-1}\frac{|h_l|^2}{1+\frac{|h_l|^2}{2\sigma^2}}\\ &= \mathrm{MSE}_{\mathrm{LMMSE}},\end{aligned}$$

i.e., no degradation is caused to the MSE. On the other hand, if the estimated SNR is lower than the actual value, a degradation will be caused to the MSE.

### 7.2.3 Interpolation-based Channel Estimation

As discussed in Section 7.2.1 and 7.2.2, LS and LMMSE channel estimations can be obtained if training signals with $n_T$ OFDM symbols are sent and each subcarrier's training signal matrix $\underline{\boldsymbol{\mathcal{S}}}_k = T_k\underline{\boldsymbol{\mathcal{M}}}$ is a non-singular matrix. When the number of transmit antennas is large, this preamble scheme could decrease the system throughput severely. We therefore in this section consider a more generalized switched subcarrier preamble scheme in which the transmit antennas and the subcarriers are both divided into $n_G$ groups, and training signals are transmitted in different subset of subcarriers in each transmit antenna group. Therefore, the preamble period needed can be reduced from $n_T$ to $\frac{n_T}{n_G}$. For example, if there are four antennas at the transmitter, we can divide them into two groups and send training signals at even number subcarriers for the first antenna group, and odd number subcarriers for the second group. As there are two antennas in each group, we can set $\underline{\boldsymbol{\mathcal{M}}}$ equal to the Walsh Hadamard matrix of dimension 2, and the training signal period is reduced from four to two. LS channel estimates can be obtained for even and odd number subcarriers for 1st and 2nd transmit antenna groups according to (7.3), respectively. Channel estimates for odd number subcarriers of the 1st antenna group and even number subcarriers of the 2nd transmit antenna group will be obtained by interpolation. The training signal period can be further reduced to one OFDM symbol if the transmit antennas are divided into four groups.

In this thesis, we consider three types of interpolation, namely, linear interpolation, LMMSE interpolation, and DFT-based LS interpolation.

#### Linear Interpolation

Assuming that the transmit antennas are divided into two groups, $\hat{H}_{n_r,n_t,k-1}$ and $\hat{H}_{n_r,n_t,k+1}$ are the LS estimates obtained from (7.3), then channel estimate at the $k$th subcarrier can be obtained through linear interpolation as follows:

$$\tilde{H}_{n_r,n_t,k} = \frac{\hat{H}_{n_r,n_t,k-1} + \hat{H}_{n_r,n_t,k+1}}{2}, \tag{7.15}$$

which is a linear matrix filtering operation as follows:

$$\begin{bmatrix} \tilde{H}_{n_r,n_t,k-1} \\ \tilde{H}_{n_r,n_t,k} \\ \tilde{H}_{n_r,n_t,k+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & \epsilon & \frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{H}_{n_r,n_t,k-1} \\ 0 \\ \hat{H}_{n_r,n_t,k+1} \end{bmatrix},$$

where $\epsilon$ represents any complex number as it does not affect the results.

### LMMSE Interpolation

Linear interpolation expressed in (7.15) assumes a channel correlation matrix $\mathbf{R}_{HH}$ with its elements defined as:

$$R_{i,j} = \begin{cases} \alpha & \text{when } |i-j| = 1, \\ 1 & \text{when } i = j, \\ 0 & \text{otherwise} \end{cases}$$

where $\alpha$ is a real number and $\alpha \in (0,1)$. This suggests that more accurate estimates could be obtained if the real channel correlation information is applied in the interpolation. In this case, not only the neighboring subcarriers, but all the available subcarriers' channel estimates will be used to calculate the missing subcarriers' channel parameters, and the contribution from different subcarriers is determined by their correlation. In our study, we use $\mathbf{W}\mathbf{R}_{HH}\left(\mathbf{R}_{HH} + \frac{1}{\text{SNR}}\mathbf{I}\right)^{-1}$ as the interpolation filter, where $\mathbf{W}$ is a normalization matrix, and $\mathbf{R}_{HH}\left(\mathbf{R}_{HH} + \frac{1}{\text{SNR}}\mathbf{I}\right)^{-1}$ is the LMMSE filter. This is the reason we call this the *LMMSE interpolation*. The simulation results presented in Section 7.2.4 will show that this interpolation scheme has better performance than linear interpolation and it is also robust to the $\mathbf{R}_{HH}$ and SNR mismatches.

### DFT-based LS Interpolation

As defined in Chapter 3, we assume a sample-spaced channel whose excess delay is no greater than the cyclic prefix length, and the time- and frequency-domain channel parameters are related by FFT and IFFT. Taking these into consideration, we propose a DFT-based LS interpolation. The derivation is as follows.

LS channel estimates for the subcarriers with training signals can be obtained according to (7.3), which will be denoted as $\hat{\mathbf{H}}_{n_r,n_t,\text{pilot}}$. Denoting the channel estimates for the other subcarriers as $\hat{\mathbf{H}}_{n_r,n_t,\text{missing}}$, we can express the channel estimates $\hat{\mathbf{H}}_{n_r,n_t}$ as:

$$\hat{\mathbf{H}}_{n_r,n_t} = \mathbf{P} \begin{bmatrix} \hat{\mathbf{H}}_{n_r,n_t,\text{pilot}} \\ \hat{\mathbf{H}}_{n_r,n_t,\text{missing}} \end{bmatrix},$$

where $\mathbf{P}$ represents a permutation matrix of size $N \times N$. As $L$ multipaths are assumed in the time domain channel, we therefore have the following relation:

$$\mathbf{G}^H \hat{\mathbf{H}}_{n_r,n_t} = \mathbf{G}^H \mathbf{P} \begin{bmatrix} \hat{\mathbf{H}}_{n_r,n_t,\text{pilot}} \\ \hat{\mathbf{H}}_{n_r,n_t,\text{missing}} \end{bmatrix} = \mathbf{0}_{N-L}, \tag{7.16}$$

where $\mathbf{G}$ is the last $(N-L)$ columns of the Fourier transform matrix $\mathbf{F}$. Letting $\mathbf{G}^H\mathbf{P} = [\mathbf{G}_T \ \mathbf{G}_M]$ so as to re-write (7.16) as:

$$[\mathbf{G}_T \ \mathbf{G}_M] \begin{bmatrix} \hat{\mathbf{H}}_{n_r,n_t,\text{pilot}} \\ \hat{\mathbf{H}}_{n_r,n_t,\text{missing}} \end{bmatrix} = \mathbf{0}_{N-L}. \tag{7.17}$$

we will have the following relation:

$$\mathbf{G}_T \hat{\mathbf{H}}_{n_r,n_t,\text{pilot}} = -\mathbf{G}_M \hat{\mathbf{H}}_{n_r,n_t,\text{missing}}, \tag{7.18}$$

which leads to:

$$\hat{\mathbf{H}}_{n_r,n_t,\text{missing}} = -(\mathbf{G}_M^H \mathbf{G}_M)^{-1} \mathbf{G}_M^H \mathbf{G}_T \hat{\mathbf{H}}_{n_r,n_t,\text{pilot}}. \tag{7.19}$$

This is a LS estimation of $\hat{\mathbf{H}}_{n_r,n_t,\text{missing}}$ from $\hat{\mathbf{H}}_{n_r,n_t,\text{pilot}}$, which suggests the name of *LS interpolation*.

IFFT can then be applied to the above frequency domain estimates to obtain a $L$-tap time domain channel estimates. The final frequency domain channel estimates will be computed by applying FFT to the $L$-tap time domain channel estimates. This IFFT and FFT operation can filter out some AWGN noise and thus improve the estimation accuracy.

### 7.2.4 Simulation Results

In this section, we will present our simulation results. For each SISO OFDM corresponding to one transmit-receive antenna pair, the system parameters defined in IEEE 802.11a [2] are used. That is, the FFT size is $N = 64$, the number of used subcarriers is $P = 52$, and the number of guard subcarriers is 12. The CP length is $L_{\text{CP}} = 16$. The long preamble given in [2] are used to construct the MIMO

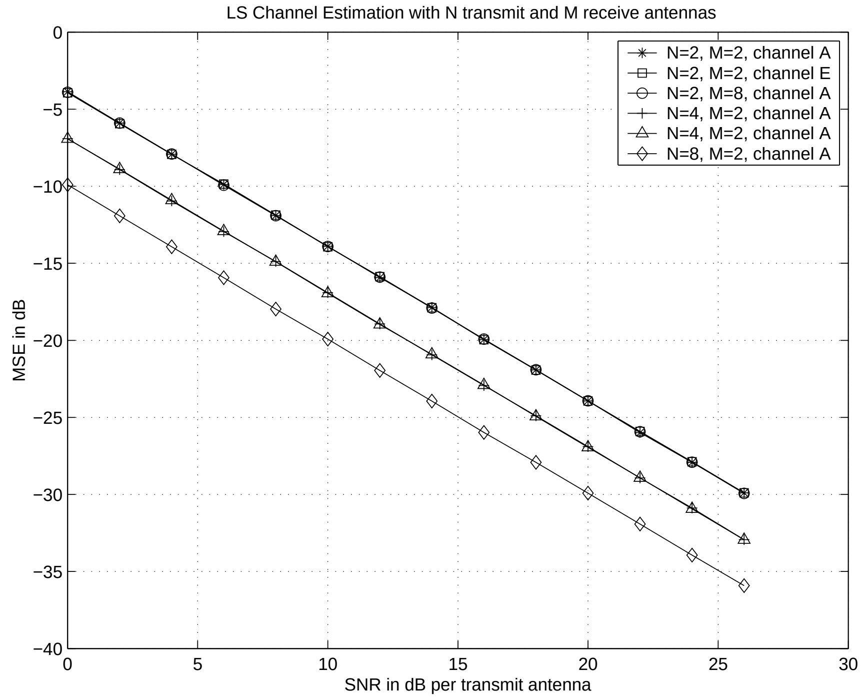


**Figure 7.3: MSE vs. SNR for LS channel estimation with N transmit and M receive antennas.**

preambles. For the 20MHz channel, two channel models are used, namely, Channel A with $\tau_{\mathrm{rms}} = 50$ ns, and Channel E with $\tau_{\mathrm{rms}=}250$ ns. For frequency domain channel, Channel E is thus less correlated than Channel A. Both channels assume an exponentially decaying power delay profile with 16 multipaths which are sample-spaced and independently generated using Jake's model [118]. The mean squared error (MSE) for the frequency domain channel estimates is used for performance comparison.

Depicted in Figure 7.3 are the MSE versus SNR per transmit antenna performances for the LS channel estimation algorithms with different number of transmit and receive antennas. We can observe from the figure that the MSE decreases linearly with the increasing SNR's. We can also observe that when the number of transmit antennas is the same, the MSE is the same for different channel models and different number of receive antennas, which is due to the fact that same power is transmitted per antenna. Therefore, the more the transmit antennas, the more the total power per receive antenna, which results in MSE drop when the transmit antenna number is increased.

We then depict in Figure 7.4 the LMMSE performance for a $2 \times 2$ MIMO-OFDM system. Shown in the same figure is the LS performance for Channel A. A few LMMSE filters are tested, namely, the

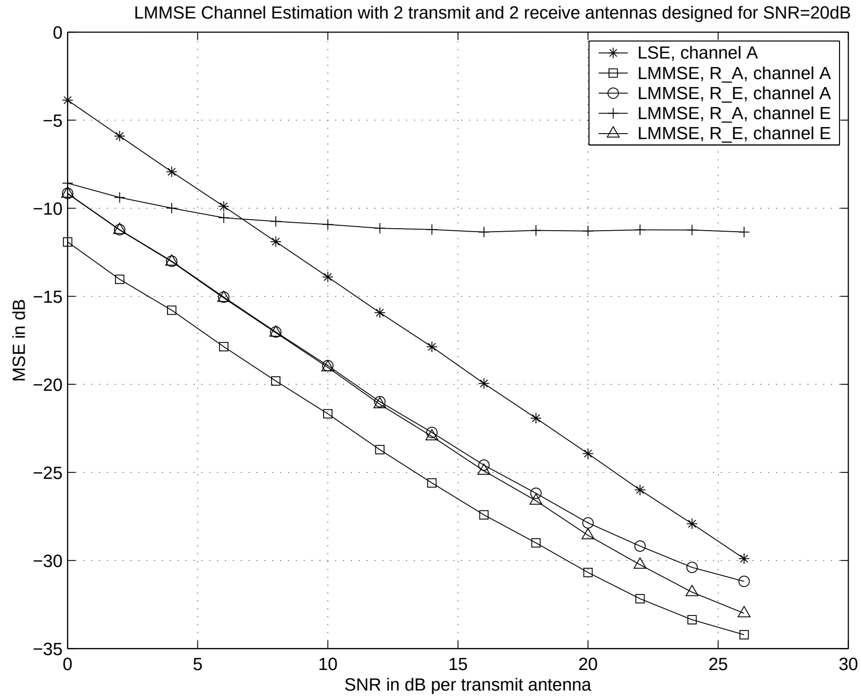


**Figure 7.4: MSE vs. SNR for LMMSE Channel Estimation with 2 transmit and 2 receive antennas.**

LMMSE filter designed for Channel A used for Channel A or Channel E, and LMMSE filter designed for Channel E used for estimation of Channel A or Channel E. In all these LMMSE filters, a fixed SNR of 20dB is used in the LMMSE filter calculation. Studying the curves in Figure 7.4, we can observe that a fixed SNR of 20dB will result in a MSE error floor in higher than 20dB SNR regions. For low SNR values, the performance is very good. Therefore, as long as we fix the SNR value to the highest possible realistic SNR's, the LMMSE estimation performance is very robust to the SNR mismatch. We can also observe from this figure that using a less correlated LMMSE filter (Channel E) to estimate a more correlated channel (Channel A), good MSE performance can still be obtained in the low to medium SNR regions. In high SNR regions, a correlation matrix mismatch of this type will result in some error floor. However, if a more correlated channel matrix (Channel A) is used to estimate a not so correlated channel (Channel E), very poor performance will result, in almost all the SNR regions of interest.

We then present our simulation results based on interpolations for switched subcarrier preamble schemes in Figure 7.5. Comparing the three interpolation schemes for Channel A, we can observe that in the low to medium SNR regions, linear interpolation and DFT-based LS interpolation have the same per-

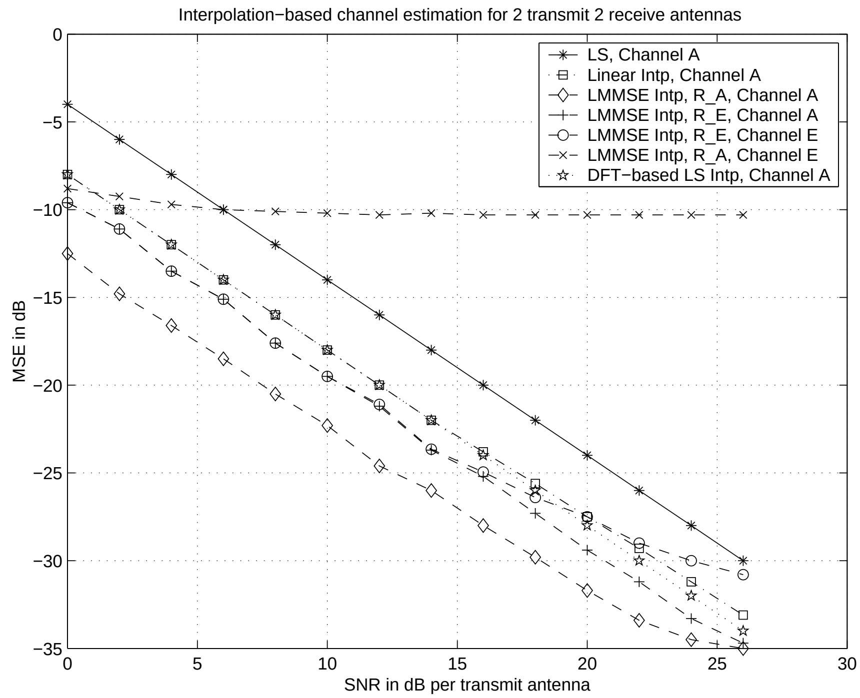


**Figure 7.5: Interpolation-based channel estimation for switched subcarrier scheme.**

formance, and in high SNR regions, the later scheme has slightly better performance. While for LMMSE interpolation, even in the mismatched case (Channel E's correlation matrix used for Channel A, fixed SNR value of 20 dB in the interpolation filter), it demonstrates better performance than the other two schemes in all the SNR regions simulated. Similar to LMMSE channel estimation, LMMSE interpolation is robust to channel model mismatch if a not so correlated channel is used in the correlation matrix computation, and it is robust to SNR mismatch as well if a high SNR value is used in computing the correlation matrix.

## 7.3 Preamble Design for Time-Domain Channel Estimation

In FDCE, we need to compute $n_R n_T N$ parameters, thus need a minimum of $n_T$ OFDM symbols of training signals when no interpolation is relied on in obtaining certain subsets of the channel estimates. The $n_R n_T N$ frequency domain channel coefficients, however, are computed from $n_R n_T L$ time domain channel coefficients. Estimation of the time domain coefficients need only $\lceil \frac{n_T L}{N} \rceil$ symbols of training signals, hence reduces the overhead significantly. In this section, we will consider TDCE for channels with $\lceil \frac{n_T L}{N} \rceil = 1$.

### 7.3.1 The Time-Domain Channel Estimation Algorithm

Again we focus on packet transmission in a block fading channel, and the training signals are transmitted only at the beginning of a packet. The received signal during the training period is expressed as

$$\mathbf{R}(k) = \begin{bmatrix} R_1(k) \\ R_2(k) \\ \vdots \\ R_{n_R}(k) \end{bmatrix} = \begin{bmatrix} H_{1,1}(k) & H_{1,2}(k) & \cdots & H_{1,n_T}(k) \\ H_{2,1}(k) & H_{2,2}(k) & \cdots & H_{2,n_T}(k) \\ & \vdots & \vdots & \\ H_{n_R,1}(k) & H_{n_R,2}(k) & \cdots & H_{n_R,n_T}(k) \end{bmatrix} \begin{bmatrix} S_1(k) \\ S_2(k) \\ \vdots \\ S_{n_R}(k) \end{bmatrix} + \mathbf{V}(k), \quad (7.20)$$

which is the same as in (7.2), except that we consider only one OFDM symbol's training here.

The frequency domain channel coefficients are computed as

$$H_{n_r,n_t}(k) = \sum_{l=0}^{L-1} h_{n_r,n_t}(l) W_N^{lk},$$

where $W_N = \exp\left(-j\frac{2\pi}{N}\right)$.

We therefore have the received signal at antenna $n_r$ and subcarrier $k$ as

$$\begin{aligned} R_{n_r}(k) &= \sum_{n_t=1}^{n_T} H_{n_r,n_t}(k) S_{n_t}(k) + V_{n_r}(k) \\ &= \sum_{n_t=1}^{n_T} S_{n_t}(k) \sum_{l=0}^{L-1} h_{n_r,n_t}(l) W_N^{lk} + V_{n_r}(k) \\ &= \sum_{n_t=1}^{n_T} \sum_{l=0}^{L-1} h_{n_r,n_t}(l) S_{n_t}(k) W_N^{lk} + V_{n_r}(k). \end{aligned} \quad (7.21)$$

Following [116], we define the MSE cost function of the channel estimates $\left\{\tilde{h}_{n_r,n_t}(l)\right\}$ as

$$\begin{aligned} & J\left(\tilde{h}_{n_r,n_t}(l)\right) \\ = & \sum_{n=0}^{N-1} \sum_{n_r=1}^{n_R} \left| R_{n_r}(n) - \sum_{n_t=1}^{n_T} \sum_{l=0}^{L-1} \tilde{h}_{n_r,n_t}(l) S_{n_t}(n) W_N^{ln} \right|^2, \quad (7.22) \\ & n_r = 1,\ 2,\ \cdots, n_R,\ \ n_t = 1,\ 2,\ \cdots, n_T,\ \ l = 0,\ 1,\ \cdots, L-1, \end{aligned}$$

and the solution to the following equation is the LS TDCE:

$$\frac{\partial J}{\partial \tilde{h}_{n_r,n_t}(l)} = \sum_{n=0}^{N-1} \left[ R_{n_r}(n) S_{n_t}^*(n) W_N^{-ln} - \sum_{p=1}^{n_T} \sum_{m=0}^{L-1} \tilde{h}_{n_r,p}(m) S_p(n) S_{n_t}^*(n) W_N^{n(m-l)} \right] = 0, \quad (7.23)$$

which is equivalent to

$$\sum_{n=0}^{N-1} R_{n_r}(n)S_{n_t}^*(n)W_N^{-ln} - \sum_{P=1}^{n_T}\sum_{m=0}^{L-1}\sum_{n=0}^{N-1} S_p(n)S_{n_t}^*(n)W_N^{-nl}\tilde{h}_{n_r,p}(m)W_N^{nm} = 0. \tag{7.24}$$

Similar to [116], we define

$$p_{n_r,n_t}(l) = \sum_{n=0}^{N-1} R_{n_r}(n)S_{n_t}^*(n)W_N^{-ln}, \tag{7.25}$$

$$q_{p,n_t}(l) = \sum_{n=0}^{N-1} S_p(n)S_{n_t}^*(n)W_N^{-nl}, \tag{7.26}$$

where $n_t,\ p = 1,\ 2,\ \cdots,\ n_T$, and $n_r = 1,\ 2,\ \cdots,\ n_R$.

We then have

$$p_{n_r,n_t}(l) = \sum_{p=1}^{n_T}\sum_{m=0}^{L-1} q_{p,n_t}(l)\tilde{h}_{n_r,p}(m)W_N^{nm}, \quad L = 0,\ 1,\ \cdots,\ L-1. \tag{7.27}$$

Further defining

$$\mathbf{p}_{n_r,n_t} = [p_{n_r,n_t}(0)\ \ p_{n_r,n_t}(1)\ \ \cdots\ \ p_{n_r,n_t}(L-1)]^T, \tag{7.28}$$

$$\mathbf{p}_{n_r} = \left[\mathbf{p}_{n_r,1}^T\ \ \mathbf{p}_{n_r,2}^T\ \ \cdots\ \ \mathbf{p}_{n_r,n_T}^T\right]^T, \tag{7.29}$$

$$\mathbf{Q}_{p,n_t} = \begin{bmatrix} q_{p,n_t}(0) & q_{p,n_t}(-1) & \cdots & q_{p,n_t}(-L+1) \\ q_{p,n_t}(1) & q_{p,n_t}(0) & \cdots & q_{p,n_t}(-L+2) \\ & \vdots & \vdots & \\ q_{p,n_t}(L-1) & q_{p,n_t}(L-2) & \cdots & q_{p,n_t}(0) \end{bmatrix}_{L\times L}, \tag{7.30}$$

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_{1,1} & \mathbf{Q}_{1,2} & \cdots & \mathbf{Q}_{1,n_T} \\ \mathbf{Q}_{2,1} & \mathbf{Q}_{2,2} & \cdots & \mathbf{Q}_{2,n_T} \\ \vdots & \vdots & & \\ \mathbf{Q}_{n_T,1} & \mathbf{Q}_{n_T,2} & \cdots & \mathbf{Q}_{n_T,n_T} \end{bmatrix}_{n_TL\times n_TL}, \tag{7.31}$$

$$\tilde{\mathbf{h}}_{n_r,p} = \left[\tilde{h}_{n_r,p}(0)\ \ \tilde{h}_{n_r,p}(1)\ \ \cdots\ \ \tilde{h}_{n_r,p}(L-1)\right]_{L\times 1}^T, \tag{7.32}$$

$$\tilde{\mathbf{h}}_{n_r} = \left[\tilde{\mathbf{h}}_{n_r,1}^T\ \ \ \tilde{\mathbf{h}}_{n_r,2}^T\ \ \ \cdots\ \ \ \tilde{\mathbf{h}}_{n_r,n_T}^T\right]_{n_TL\times 1}^T, \tag{7.33}$$

where

$$n_r = 1,\ 2,\ \cdots,\ n_R, \qquad n_t,\ p = 1,\ 2,\ \cdots,\ n_T,$$

we will have

$$\mathbf{p}_{n_r} = \mathbf{Q}\tilde{\mathbf{h}}_{n_r}. \tag{7.34}$$

This is the same as in [119] except that [119] considered only a $2 \times 2$ MIMO system.

One thing to note is that $\mathbf{Q}$ is the same for different receive antennas. We therefore can form the linear relation of $\mathbf{p}_{n_r}$, $\mathbf{Q}$, and $\tilde{\mathbf{h}}_{n_r}$ as follows:

$$\begin{bmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \vdots \\ \mathbf{p}_{n_R} \end{bmatrix} = \begin{bmatrix} \mathbf{Q} & & & & \\ & \mathbf{Q} & & \underline{\mathbf{0}} & \\ & & \ddots & & \\ & \underline{\mathbf{0}} & & \ddots & \\ & & & & \mathbf{Q} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{h}}_1 \\ \tilde{\mathbf{h}}_2 \\ \vdots \\ \tilde{\mathbf{h}}_{n_R} \end{bmatrix}. \tag{7.35}$$

The time-domain channel estimates are therefore obtained as

$$\begin{bmatrix} \hat{\mathbf{h}}_1 \\ \hat{\mathbf{h}}_2 \\ \vdots \\ \hat{\mathbf{h}}_{n_R} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}^{-1} & & & & \\ & \mathbf{Q}^{-1} & & \underline{\mathbf{0}} & \\ & & \ddots & & \\ & \underline{\mathbf{0}} & & \ddots & \\ & & & & \mathbf{Q}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \vdots \\ \mathbf{p}_{n_R} \end{bmatrix}. \tag{7.36}$$

### Computation of $\mathbf{p}$

From the definition in (7.25), we can further write the calculation of $\mathbf{p}$ as

$$\begin{aligned} p_{n_r,n_t}(l) &= \sum_{n=0}^{N-1} R_{n_r}(n) S_{n_t}^*(n) W_N^{-ln} \\ &= \sum_{n=0}^{N-1} S_{n_t}^*(n) W_N^{-ln} \left( \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} r_{n_r}(k) W_N^{nk} \right) \\ &= \sum_{k=0}^{N-1} r_{n_r}(k) \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} S_{n_t}^*(n) W_N^{-n(l-k)} \\ &= \sum_{k=0}^{N-1} r_{n_r}(k) \left( \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} S_{n_t}(n) W_N^{-n(k-l)} \right)^* \\ &= \sum_{k=0}^{N-1} r_{n_r}(k) s_{n_t}^*(k-l), \end{aligned} \tag{7.37}$$

where $R_{n_r}(n)$ and $r_{n_r}(k)$, $n,\ k = 0,\ 1,\ \cdots,\ N-1$ denote the frequency and time domain signals at receive antenna $n_r$, $s^*_{n_t}(k-l)$ denotes the time-domain training signal at transmit antenna $n_t$ which is cyclicly shifted by $l$ symbols, and $n_r = 1,\ 2,\ \cdots,\ n_R$, $n_t = 1,\ 2,\ \cdots,\ n_T$.

**Computation of $\mathbf{Q}$**

It has been shown in [119] that when $\mathbf{Q} = N\mathbf{I}$, not only the channel estimation is simplified greatly by eliminating the need for matrix inversion, but also the MSE is optimal. In [116], a cyclic shift-based training sequence satisfying $\mathbf{Q} = N\mathbf{I}$ was proposed. Here we propose another simple and "optimal" training signal scheme based on subcarrier switching.

### 7.3.2 Subcarrier Switching Training Sequence

The frequency domain subcarrier switching training signal is expressed as

$$S_p(n) = \sqrt{n_T} S(n)\delta\left((n-p+1) \mod n_T\right), \tag{7.38}$$

where $p = 1,\ 2,\ \cdots,\ n_T$, $n = 0,\ 2,\ \cdots,\ N-1$, $|S(n)|^2 = 1$, $\delta(n)$ is the Dirac delta function, and the factor $\sqrt{n_T}$ is to normalize the average transmission power per subcarrier per transmit antenna to 1.

We therefore have the following relationship:

$$S_p(n)S^*_{n_t}(n) = n_T\delta(p-n_t)\delta\left((n-p+1) \mod n_T\right),$$

and $q_{p,n_t}(l)$ calculated following (7.26) as:

$$\begin{aligned} q_{p,n_t}(l) &= \sum_{n=0}^{N-1} S_p(n)S^*_{n_t}(n)W_N^{-nl} \\ &= \begin{cases} 0, & p \neq n_t, \\ \displaystyle\sum_{n=0}^{N-1} |S_p(n)|^2 W_N^{-nl}, & p = n_t. \end{cases} \end{aligned}$$

As

$$\sum_{n=0}^{N-1} |S_p(n)|^2 W_N^{-nl} = n_T \sum_{\substack{n=p-1 \\ \Delta n = n_T}}^{N-1} W_N^{-nl},$$

when $(N \mod n_T) = 0$, we have

$$\begin{aligned}\sum_{n=0}^{N-1} |S_p(n)|^2 W_N^{-nl} &= n_T \sum_{k=1}^{\frac{N}{n_T}-1} W_N^{-kn_T l} W_N^{-(p-1)l} = n_T W_N^{-(p-1)l} \sum_{k=1}^{\frac{N}{n_T}-1} W_{\frac{N}{n_T}}^{-kl} \\ &= \begin{cases} N W_N^{-(p-1)l}, & \text{when } \left(l \mod \frac{N}{n_T}\right) = 0 \\ 0, & \text{otherwise} \end{cases}\end{aligned}$$

where $W_{\frac{N}{n_T}} = \exp\left(-j\frac{2\pi}{\frac{N}{n_T}}\right) = \exp\left(-j\frac{2\pi n_T}{N}\right)$.

As we want to have

$$\mathbf{Q}_{p,n_t} = N\mathbf{I}_{L\times L}\delta(p - n_t), \tag{7.39}$$

i.e., $q_{p,n_t}(l) = N\delta(l)$, we should have the following relation satisfied

$$L - 1 < \frac{N}{n_T},$$

hence

$$n_T < \frac{N}{L-1},$$

i.e., the maximum number of antennas supported by this training signal scheme is $\lfloor \frac{N}{L-1} \rfloor$.

The time domain channel estimates are obtained as

$$\begin{bmatrix} \hat{\mathbf{h}}_1 \\ \hat{\mathbf{h}}_2 \\ \vdots \\ \hat{\mathbf{h}}_{n_R} \end{bmatrix} = \frac{1}{N} \begin{bmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \vdots \\ \mathbf{p}_{n_R} \end{bmatrix}. \tag{7.40}$$

**Remark** Compared with the CCDP in [116], the switched subcarrier training sequence has lower PAPR, due to the fact that the fewer number of subcarriers are active. The MSE performance of both schemes are exactly the same.

### 7.3.3 Windowing on the Time-Domain Channel Estimates

After the time-domain channel coefficients are obtained, the frequency-domain coefficients can be obtained from FFT. Before applying the FFT, some windowing functions, e.g., Hamming window, Hanning window, or Blackman window, can be applied in order to further minimize the MSE, as proposed in [120]. It was also shown in [120] that the Blackman windowing function provides as good as or even better BER performance than the LMMSE channel estimation scheme.

## 7.4 Conclusions

We have presented several results of our study on MIMO OFDM channel estimation. Based on the linear matrix algebraic model, we have derived a general frequency domain preamble structure which is just a simple extension from the SISO OFDM preamble. Therefore, the good properties, such as low PAPR, of the SISO OFDM preamble can be maintained. We then developed the least squares and linear minimum mean squared error channel estimation algorithms for this proposed preamble scheme. We further proposed a switched subcarrier preamble scheme which needs fewer OFDM symbols in the training sequence and therefore the transmission efficiency is improved. Three interpolation schemes, namely, linear interpolation, LMMSE interpolation and DFT-based LS interpolation are proposed, among which the LMMSE interpolation scheme demonstrates the best performance, even in the mismatch case. As both LMMSE channel estimation and LMMSE interpolation can be implemented with fixed parameter values in the matrix filter, the implementation is very simple and therefore attractive for practical deployment. For time-domain channel estimation, we proved that the switched subcarrier training sequence satisfies the optimal MSE criteria and supports simple channel estimation. Compared with the cyclic-shift-based training sequence proposed in [116], the switched subcarrier training sequence has lower PAPR.

## Appendix A - Definition of First Order Derivative to A Complex Variable

The first order derivative of a function $y = f(x)$ to the complex variable $x = x_r + j\, x_i$ is defined as

$$\frac{dy}{dx} = \frac{1}{2}\left(\frac{\partial y}{\partial x_r} - j\,\frac{\partial y}{\partial x_i}\right). \tag{7.41}$$

Some of the special cases which are used in this chapter are listed below.

1.

$$\begin{aligned} & y = ax = ax_r + jax_i \\ & \tfrac{dy}{dx} = a. \end{aligned} \tag{7.42}$$

2.

$$\begin{aligned} & y = ax^* = ax_r - jax_i \\ & \tfrac{dy}{dx} = 0. \end{aligned} \tag{7.43}$$

3.

$$\begin{aligned} & y = |x|^2 = x_r^2 + x_i^2 \\ & \tfrac{dy}{dx} = 2x^*. \end{aligned} \tag{7.44}$$

4.

$$\begin{aligned} & y = |a - bx|^2 = |a|^2 - a^*bx - ab^*x^* + |b|^2|x|^2 \\ & \tfrac{dy}{dx} = -by^*. \end{aligned} \tag{7.45}$$

## Appendix B - Definition of First Order Derivative of a Scalar to A Complex Matrix

The first order derivative of a scalar-valued function $y = f(\mathbf{x})$ to the complex $M \times N$ matrix $\mathbf{x}$ is defined as

$$\frac{dy}{d\mathbf{x}} = \begin{bmatrix} \frac{\partial y}{\partial x_{11}^*} & \frac{\partial y}{\partial x_{12}^*} & \cdots & \frac{\partial y}{\partial x_{1N}^*} \\ & \ddots & & \\ & & \ddots & \\ \frac{\partial y}{\partial x_{M1}^*} & \frac{\partial y}{\partial x_{M2}^*} & \cdots & \frac{\partial y}{\partial x_{MN}^*} \end{bmatrix}, \tag{7.46}$$

where $x_{ij}$, $i = 1, \cdots, M$, $j = 1, \cdots, N$, is the element of $\mathbf{x}$.

Defining a $M \times 1$ column vector $\mathbf{a}$ and a $N \times 1$ column vector $\mathbf{b}$, we have the following results:

1.

$$\begin{aligned} & y = \mathbf{a}^H\mathbf{x}\mathbf{b} \\ & \tfrac{dy}{d\mathbf{x}} = \mathbf{0}_{M\times N}. \end{aligned} \tag{7.47}$$

2.

$$\begin{aligned} & y = \mathbf{b}^H\mathbf{x}^H\mathbf{a} \\ & \tfrac{dy}{dx} = \mathbf{a}\mathbf{b}^H. \end{aligned} \tag{7.48}$$

3.

$$\begin{aligned} y &= c \\ \frac{dy}{d\mathbf{x}} &= \mathbf{0}_{M\times N}. \end{aligned} \tag{7.49}$$

4.

$$\begin{aligned} y &= \mathbf{b}^H\mathbf{x}^H\mathbf{x}\mathbf{b} \\ \frac{dy}{d\mathbf{x}} &= \mathbf{x}\mathbf{b}\mathbf{b}^H. \end{aligned} \tag{7.50}$$

# Chapter 8

# Conclusions and Recommendations for Future Work

## 8.1 Conclusions

We have addressed several issues associated with transmit and receive techniques for MIMO-OFDM systems. Our main contributions are summarized next.

Driven by the motivation of achieving optimal tradeoff between the multiplexing gain and diversity gain for MIMO-OFDM channels, especially for asymmetric MIMO-OFDM channels, we studied several linear and non-linear precoding schemes which can map fewer spatial streams to more transmit antennas. In order to unify the analysis, we developed a linear signal model and systematically compared their ergodic capacity, outage capacity, and diversity performances. In this process, we developed the closed form equation for the spatial spreading systems using random matrix theory. We also proved that the $4 \times 2$ groupwise space-time block coding and quasi-orthogonal space-time block coding perform exactly the same in ergodic capacity sense. A two-dimensional linear pre-transformed MIMO-OFDM system was proposed which can achieve full diversity and full diversity simultaneously.

Exploitation of the diversity and multiplexing gains in the MIMO-OFDM channel relies on not only an effective precoding scheme at the transmitter, but also on an optimal and efficient receiver. In this thesis, we dedicated our effort to the iterative algorithms using “turbo principle”. We proposed the linear soft decision functions for high-order modulation signals which can significantly reduce the computational complexity in signal estimation but at the same time maintain the BER performance. More importantly, we proposed a novel Bayesian minimum mean squared error turbo receiver. Compared with the conventional turbo receivers in the literature which make use of only the extrinsic information from the decoder for interference estimation and cancelation, the proposed Bayesian turbo receiver uses both the decoder extrinsic information *and* the detector decision statistic for interference estimation. As a result, the estimation accuracy is greatly improved, especially in low to medium SNR regions. This also contributes to the 1.5 dB improvement at BER performance of $10^{-5}$, and the better convergence behavior of the turbo process.

We also developed the extrinsic information transfer chart for the proposed Bayesian turbo receivers. Compared with the conventional turbo receivers, the proposed Bayesian turbo receivers demonstrated a much higher output mutual information, proving its superior performance. When plotted with the extrinsic information transfer chart of the decoder, the trajectories of the Bayesian receivers also exhibit much faster convergence than the conventional receivers.

Our next contribution lies in the systematic study of training signal design for both frequency-domain and time-domain channel estimation in MIMO-OFDM systems. Minimum mean squared error-achieving preamble schemes have been proposed which require very simple filtering calculation to obtain the channel estimates.

## 8.2 Recommendations for Future Work

The following issues can be studied further as continuation of the research in this thesis.

### 8.2.1 Space-Time-Frequency Processing for Spatially Correlated Channels

We have studied the precoding schemes under the assumption of no spatial correlation in the MIMO-OFDM channels. This assumption, however, becomes weaker when the antenna spacing is reduced, especially for the receive antennas at the terminal. Therefore, it is important to look into the precoding schemes in the spatially correlated channels and propose effective solutions.

### 8.2.2 Low-Complexity Near Optimal Receiver Algorithms for 2DLPT MIMO-OFDM

The two dimensional linear pre-transformed MIMO-OFDM system achieves full diversity with maximum likelihood detection receiver. Receiver algorithms are therefore desired which can effectively exploit the diversity gains with affordable complexity.

### 8.2.3 Extension of 2DLPT to Single-Carrier Cyclic-Prefix MIMO Systems

The 2DLPT can simultaneously achieve full capacity and full diversity when the transform is unitary. With the similarity between the MIMO-OFDM channels and the MIMO single-carrier cyclic prefix (SCCP) channels, it is expected that similar transform can be applied to MIMO-SCCP to achieve full capacity and full diversity.

### 8.2.4 Incorporation of Channel Estimation in the Bayesian Turbo Receiver

We have proposed the Bayesian turbo receivers and studied their performance under the assumption of perfect channel estimation. We have also proposed several preamble designs to support optimal channel estimation. As a natural continuation, incorporation of channel estimation into the Bayesian turbo receiver by using the proposed preamble schemes needs to be studied. The corresponding EXIT chart analysis needs to be developed as well.

### 8.2.5 Soft Decision Function Simplification in Bayesian EM Estimate

Incorporation of both the SISO decoder EXT and the soft output detector output in the interference estimation will improve the estimation accuracy, hence better BER and FER performance, as discussed in Section 5.6. However, as more variables are included in the soft decision function, the computational complexity in the signal estimation will become higher than the conventional turbo receiver, especially for high-order modulation schemes such as MQAM. Therefore, possible simplification of SDF's in Bayesian EM estimate is desired.

# Bibliography


[1] “Part 11: wireless LAN medium access control (MAC) and physical layer (PHY) specifications: Higher-speed physical layer extension in the 2.4 GHz band.” IEEE std 802.11b-1999: Supplement to ANSI/IEEE std 802.11, 1999 Edition, Sept. 1999.

[2] “Part 11: wireless LAN medium access control (MAC) and physical layer (PHY) specifications: High-speed physical layer in the 5GHz band.” IEEE std 802.11a-1999: Supplement to IEEE std 802.11-1999, Sept. 1999.

[3] “Part 11: wireless LAN medium access control (MAC) and physical layer (PHY) specifications-Amendment 4: Further Higher Data Rate Extension in the 2.4 GHz band.” IEEE std 802.11g-2003: Amendment to IEEE Std 802.11$^{\mathrm{T}M}$, 1999 Edition (Reaff 2003), June 2003.

[4] IEEE 802.11 TGn, “http://www.ieee802.org/11/.”

[5] S. Cherry, “Edholm’s law of bandwidth,” *IEEE Spectrum*, vol. 41, pp. 50–60, July 2004.

[6] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: John Wiley&Sons, 1986.

[7] G. J. Foschini and M. J. Gans, “On limits of wireless communications in a fading environment when using multiple antennas,” *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, 1998.

[8] I. Emre Telatar, “Capacity of multi-antenna Gaussian channels,” *European Transactions on Telecommunications*, vol. 10, pp. 585–595, Nov/Dec 1999.

[9] G. J. Foschini, "Layered space-time architecture for wireless communications in a fading environment when using multi-element antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, 1996.

[10] G. J. Foschini, G. D. Golden, R. A. Valenzuela, and P. W. Wolnianski, "Simplified processing for high spectral efficiency wireless communication employing multi-element arrays," *IEEE J. Selected Areas Commun.*, vol. 17, pp. 1841–1852, Nov. 1999.

[11] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading Channels: Information-Theoretic and Communications Aspects," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2619–2692, Oct. 1998.

[12] J. G. Proakis, *Digital Communications, Fourth Edition*. McGraw-Hill, 2001.

[13] L. J. Cimini, Jr., "Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing," *IEEE Trans. Commun.*, vol. COM-33, 1985.

[14] W. Y. Zou and Y. Wu, "COFDM: An overview," *IEEE Trans. Broadcast*, vol. 41, pp. 1–8, March 1995.

[15] Richard van Nee and Ramjee Prasad, *OFDM Wireless Multimedia Communications*. Artech House, 2000.

[16] "Broadband radio access networks BRAN: HIPERLAN type 2 technical specification: physical layer." ETSI/OTS/BRAN 0023003, Oct. 1999.

[17] S. Lin and D. J. Constello, Jr., *Error Control Coding: Fundamentals and Applications*. Prentical Hall, 1983.

[18] S. B. Wicker, *Error Control Systems for Digital Communication and Storage*. Prentice Hall International, Inc., 1995.

[19] J. Hagenauer, "Rate-compatible punctured convolutional codes (RCPC codes) and their applications," *IEEE Trans. Commun.*, vol. 36, pp. 389–400, April 1988.

[20] A. J. Viterbi, *CDMA: Principles of spread spectrum communication*. Addison Wesley Pub. Co., 1995.

[21] V. K. Garg, K. Smolik, and J. E. Wilkes, *Applications of CDMA in wireless/personal communications*. Prentice Hall PTR, 1997.

[22] A. Wittneben, "Basestation modulation diversity for digital SIMULCAST," in *IEEE Vehicular Technology Conference*, pp. 848–853, 1991.

[23] A. Wittneben, "A new bandwidth efficient transmit antenna modulation diversity scheme for linear digital modulation," in *IEEE International Conference on Communications*, vol. 3, pp. 1630–1634, 1993.

[24] J. H. Winters, "The diversity gain of transmit diversity in wireless systems with Rayleigh fading," in *IEEE International Conference on Communications*, vol. 2, pp. 1121–1125, 1994.

[25] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communications: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744–765, March 1998.

[26] R. G. Gallager, *Information Theory and Reliable Communications*. New York: John Wiley&Sons, 1968.

[27] A. Edelman, *Eigenvalues and Condition Numbers of Random Matrices*. PhD thesis, Department of Mathematics, MIT, Cambridge, 1989.

[28] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series, and Products*. New York: Academic Press corercted and enlarged ed., 1980.

[29] B. Hochwald, T. L. Marzetta, and V. Tarokh, "Multi-antenna channel-hardening and its implications for rate feedback and scheduling," *IEEE Trans. Inform. Theory*, vol. 50, pp. 1893–1909, Sept. 2004.

[30] A. L. Moustakas, S. H. Simon, and A. M. Sengupta, "MIMO capacity through correlated channels in the presence of correlated interferers and noise: A (not so) large $N$ analysis," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2545–2561, Oct. 2003.

[31] S. M. Alamouti, "A simple transmit diversity techniques for wireless communications," *IEEE J. Selected Areas Commun.*, vol. 16, pp. 1451–1458, Oct. 1998.

[32] B. Hassibi and B. M. Hochwald, "High-rate codes that are linear is space and time," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1804–1824, July 2002.

[33] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal design," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1456–1467, July 1999.

[34] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block coding for wireless communications: performance results," *IEEE J. Selected Areas Commun.*, vol. 17, pp. 451–460, March 1999.

[35] G. Caire, G. Taricoo and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inform. Theory*, vol. 44, pp. 927–946, May 1998.

[36] B. Vucetic and J. Yuan, *Space-time coding*. New York: John Wiley & Sons, 2003.

[37] G. D. Forney, Jr., "Geometrically uniform codes," *IEEE Trans. Inform. Theory*, vol. 37, pp. 1241–1260, Sept. 1991.

[38] H. Jafarkhani, "A quasi-orthogonal space-time block code," *IEEE Trans. Commun.*, vol. 49, pp. 1–4, Jan. 2001.

[39] W. Sun and X. G. Xia, "Quasiorthogonal space-time block codes with full diversity," in *IEEE Global TeleCommunications Conference*, vol. 2, pp. 1098–1102, 2002.

[40] N. Sharma and C. B. Papadias, "Improved quasi-orthogonal space-time codes through constellation rotation," *IEEE Trans. Commun.*, vol. 51, pp. 332–335, March 2003.

[41] C. Yuen, Y. L. Guan, and T. T. Tjhung, “New high-rate STBC with good dispersion property,” in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, (Berlin), Sept. 2005.

[42] R. Heath Jr. and A. Paulraj, “Swithching between multiplexing and diversity based on constellation distance,” in *Allerton Conf. Communication, Control and Computing*, Oct. 2000.

[43] I.-M. Kim and V. Tarokh, “Variable-Rate Space-Time Block Codes in M-ary PSK Systems,” *IEEE J. Selected Areas Commun.*, vol. 21, pp. 362–373, April 2003.

[44] S. Sun, T. T. Tjhung, and P. H. W. Fung, “Soft-decision Based Iterative Interference Cancellation (IIC) for Group-Wise STBC MIMO Systems,” in *IEEE Vehicular Technology Conference-Spring*, vol. 2, pp. 984–988, April 2003.

[45] V. Tarokh, A. Naguib, N. Seshadri, and A. R. Calderbank, “Combined array processing and space-time coding,” *IEEE Trans. Inform. Theory*, vol. 45, pp. 1121–1128, May 1999.

[46] Z. Lei, Y. Dai and S. Sun, “Ordered maximum SNR array processing for space time coded systems,” *IEE Electron. Lett.*, vol. 39, pp. 561–562, March 2003.

[47] R. W. Heath and A. J. Paulraj, “Linear dispersion codes for MIMO systems based on frame theory,” *IEEE Trans. Signal Processing*, vol. 50, pp. 2429–2441, Oct. 2002.

[48] L. Zheng and D. N. C. Tse, “Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels,” *IEEE Trans. Inform. Theory*, vol. 49, pp. 1073–1096, May 2003.

[49] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1998.

[50] Helmut Lütkepohl, *Handbook of Matrices*. John Wiley & Sons, 1996.

[51] D. Gesbert, H. Bölcskei, D. A. Gore, and A. J. Paulraj, “Outdoor MIMO wireless channels: Models and performance prediction,” *IEEE Trans. Commun.*, vol. 50, pp. 1926–1934, Dec. 2002.

[52] H. Bölcskei, D. Gesbert, and A. J. Paulraj, “On the capacity of OFDM-based spatial multiplexing systems,” *IEEE Trans. Commun.*, vol. 50, pp. 225–234, Feb. 2002.

[53] Xiaodong Li and J. A. Ritcey, “Trellis coded modulation with bit interleaving and iterative decoding,” *IEEE J. Selected Areas Commun.*, vol. 17, pp. 715–724, April 1999.

[54] X. Li and J. A. Ritcey, “Bit-interleaved coded modulation with iterative decoding,” *IEEE Commun. Lett.*, vol. 1, pp. 169–171, Nov. 1997.

[55] A. Chindapol and J. A. Ritcey, “Design, analysis, and performance evaluation for BICM-ID with square QAM constellations in Rayleigh fading channels,” *IEEE J. Selected Areas Commun.*, vol. 19, pp. 944–957, May 2001.

[56] Y. Huang and A. Ritcey, “16-QAM BICM-ID in fading channels with imperfect channel state information,” *IEEE Trans. Wireless Commun.*, vol. 2, pp. 1000–1007, Sept. 2003.

[57] S. Schreckenbach, N. Görtz, J. Hagenauer, and G. Bauch, “Optimization of symbol mappings for bit-interleaved coded modulation with iterative decoding,” *IEEE Commun. Lett.*, vol. 7, pp. 593–595, Dec. 2003.

[58] S. Schreckenbach, N. Görtz, J. Hagenauer, and G. Bauch, “Optimized symbol mappings for bit-interleaved coded modulation with iterative decoding,” in *IEEE Global TeleCommunications Conference*, vol. 6, pp. 3316–3320, Dec. 2003.

[59] V. Tarokh, A. Naguib, N. Seshadri, and A. R. Calderbank, “Space-time codes for high data rate wireless communication: performance criteria in the presence of channel estimation errors, mobility, and multiple paths,” *IEEE Trans. Commun.*, vol. 47, pp. 199–207, Feb. 1999.

[60] Y. Wu, Z. Lei, and S. Sun, “Performance of Walsh-Hadamard Transformed STBC OFDM System,” in *IEEE Vehicular Technology Conference-Spring*, (Milan, Italy), May 2004.

[61] A. Dammann and S. Kaiser, “Standard conformable antenna diversity techniques for OFDM and its application to the DVB-T system,” in *IEEE Global TeleCommunications Conference*, vol. 5, pp. 3100–3105, Nov. 2001.

[62] S. Kaiser, “Spatial transmit diversity techniques for broadband OFDM systems,” in *IEEE Global TeleCommunications Conference*, vol. 3, pp. 1824–1828, Nov. 2000.

[63] J. Tan and G. L. Stüber, "Multicarrier delay diversity modulation for MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 1756–1763, Sept. 2004.

[64] A. Dammann, R. Raulefs and S. Kaiser, "Beamforming in Combination with Space-Time Diversity for Broadband OFDM Systems," in *IEEE International Conference on Communications*, vol. 1, pp. 165–171, April 2002.

[65] S. Sun, Y, Wu, Y, Li, and T. T. Tjhung, "A Novel Iterative Receiver for Coded MIMO OFDM Systems," in *IEEE International Conference on Communications*, vol. 4, pp. 2473–2477, June 2004.

[66] B. Hassibi and B. M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1804–1824, July 2002.

[67] Y.-C. Liang, R. Zhang, and J. M. Cioffi, "Sub-channel grouping and statistical water-filling for MIMO-OFDM Systems," in *36th Asilomar Conference on Signals, Systems, and Computers*, vol. 1, (Pacific Grove, CA), pp. 997–1001, 2003.

[68] Y. Li, Y.-C. Liang, S. Sun, and R. Zhang, "Adaptive Trellis and Bit-Interleaved Coded Modulation for Ordered MIMO-OFDM Channels," in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, (Berlin), Sept. 2005.

[69] Wei Yu and J. M. Cioffi, "Sum Capacity of Gaussian Vector Broadcast Channels," *IEEE Trans. Inform. Theory*, vol. 50, pp. 1875–1892, Sept. 2004.

[70] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel," in *IEEE ISSSE-98*, (Pisa, Italy), pp. 295–300, Sept. 1998.

[71] B. M. Hochwald and S. Brink, "Achieving Near-Capacity on a Multiple-Antenna Channel," *IEEE Trans. Commun.*, vol. 51, pp. 389–399, March 2003.

[72] B. Lu, G. Yue, and X. Wang, "Performance analysis and design optimization of LDPC-coded MIMO OFDM systems," *IEEE Trans. Signal Processing*, vol. 52, pp. 348–361, Feb. 2004.

[73] S. Sun, T. T. Tjhung, and Y, Li, " An Iterative Receiver for Groupwise Bit-Interleaved Coded QAM STBC OFDM," in *IEEE Vehicular Technology Conference-Spring*, vol. 3, pp. 1256–1260, May 2004.

[74] A. Molisch, M. Z. Win, Y.-S. Choi, and J. H. Winters, "MIMO systems with antenna selection," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 2759–2773, July 2005.

[75] D. A. Gore, R. W. Heath Jr. and A. J. Paulraj, "Transmit antenna selection in spatial multiplexing systems," *IEEE Commun. Lett.*, vol. 6, pp. 491–493, Nov. 2002.

[76] X. Shao, J. Yuan, and P. Rapajic, "Antenna selection for MIMO-OFDM spatial multiplexing systems," in *Proc. IEEE Int. Symp. on Information Theory*, (Yokohama, Japan), p. 90, 2003.

[77] A. Molisch and M. Z. Win, "MIMO systems with antenna selection," *IEEE Microwave Magazine*, vol. 5, pp. 46–56, March 2004.

[78] Y. Li, Y.-C. Liang, S. Sun, and R. Zhang, "SCCR LDPC Code for Ordered MIMO-OFDM Channels," in *38th Asilomar Conference on Signals, Systems, and Computers*, (Pacific Grove, CA), Nov. 2005.

[79] Y.-C. Liang, S. Sun, and C. K. Ho, "Signal detection for large MIMO systems using block-iterative generalized decision feedback equalizers (BI-GDFE)," in *IEEE International Conference on Communications*, vol. 3, (Seoul, Korea), pp. 2112–2116, May 2005.

[80] Y.-C. Liang, S. Sun, and C. K. Ho, "Block-iterative generalized decision feedback equalizers (BI-GDFE) for large MIMO systems: algorithm design and asymptotic performance analysis," *IEEE Trans. Signal Processing*, to appear.

[81] R. Kalbasi, R. Dinis, D. D. Falconer, and A. H. Banihashemi, "Layered space-time receivers for single carrier transmission with iterative frequency domain equalization," in *IEEE Vehicular Technology Conference-Spring*, vol. 1, pp. 575–579, May 2004.

[82] C. K. Ho, Z. Lei, S. Sun, and Y. Wu, "Iterative detection for pretransformed OFDM by subcarrier reconstruction," *IEEE Trans. Signal Processing*, vol. 53, pp. 2842–2854, Aug. 2005.

[83] M. Witzke, S. Bäro, and J. Hagenauer, "Iterative detection of generalized coded MIMO signals using a widely linear detector," in *IEEE Global TeleCommunications Conference*, vol. 4, pp. 1821–1825, Dec. 2003.

[84] C. K. Ho, Z. Lei, S. Sun, and Y. Wu, "Performance analysis of iterative detection for pre-transformed OFDM," in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 2, (Barcelona, Spain), pp. 1332–1336, Sept. 2004.

[85] X. Wang and H. Vincent Poor, "Iterative (Turbo) Soft Interference Cancellation and Decoding for Coded CDMA," *IEEE Trans. Commun.*, vol. 47, pp. 1046–1060, July 1999.

[86] P. D. Alexander, A. J. Grant,and M. C. Reed, "Iterative Detection in Code-Division Multiple-Access with Error Control Coding," *European Trans. Telecommun. Special Issue CDMA Techniques Wireless Commun. Syst.*, vol. 9, pp. 419–425, Sept./Oct. 1998.

[87] M. C. Reed, S. B. Schlegel, P. D. Alexander, and J. A. Asenstorfer, "Iterative multiuser detection for CDMA with FEC: near-single-user performance," *IEEE Trans. Commun.*, vol. 46, pp. 1693–1699, Dec. 1998.

[88] P. D. Alexander, M. C. Reed, J. A. Asenstorfer,and C. B. Schlegel, "Iterative Multiuser Interference Reduction: Turbo CDMA," *IEEE Trans. Commun.*, vol. 47, pp. 1008–1014, July 1999.

[89] B. Lu and X. Wang, "Iterative Receivers for Multiuser Space-Time Coding Systems," *IEEE J. Selected Areas Commun.*, vol. 18, pp. 2322–2335, Nov. 2000.

[90] M. Sellathurai and S. Haykin, "Turbo-BLAST for Wireless Communications: Theory and Experiments," *IEEE Trans. Signal Processing*, vol. 50, pp. 2538–2546, Oct. 2002.

[91] M. Sellathurai and S. Haykin, "Turbo-BLAST: Performance Evaluation in Correlated Rayleigh-Fading Environment," *IEEE J. Selected Areas Commun.*, vol. 21, pp. 340–349, April 2003.

[92] M. Sellathurai and S. Haykin, "T-BLAST for Wireless Communications: First Experimental Results," *IEEE Trans. Veh. Technol.*, vol. 52, pp. 530–535, May 2003.

[93] J. Boutros and G. Caire, “Iterative multiuser joint decoding: Unified framework and asymptotic analysis,” *IEEE Trans. Inform. Theory*, vol. 48, pp. 1772–1793, July 2002.

[94] F. Kschischang, B. Frey, and H.-A. Loeliger, “Factor graphs and the sum-product algorithm,” *IEEE Trans. Inform. Theory*, vol. 47, pp. 498–519, Feb. 2001.

[95] S. M. Kay, *Fundamentals of statistical signal processing, Volume I: Estimation theory*. Prentice Hall PTR, 1993.

[96] E. Biglieri, A. Nordio, and G. Taricco, “Iterative receiver interfaces for coded multiple-antenna signaling,” in *2nd International Workshop on Signal Processing for Wireless Communications*, (King’s College London), June 2004.

[97] J. M. Mendel, *Lessons in estimation theory for signal processing, communications, and control*. Prentice Hall PTR, 1995.

[98] H. V. Poor and S. Verdú, “Probability of error in MMSE multiuser detection,” *IEEE Trans. Inform. Theory*, vol. 43, pp. 858–871, May 1997.

[99] G. Begin and D. Haccoun, “High-rate punctured convolutional codes: structure properties and construction technique,” *IEEE Trans. Commun.*, vol. 37, pp. 1381–1385, Dec. 1989.

[100] G. Begin, D. Haccoun, and C. Paquin, “Further results on high-rate punctured convolutional codes for Viterbi and sequential decoding,” *IEEE Trans. Commun.*, vol. 38, pp. 1922–1928, Nov. 1990.

[101] L. B. Nelson and H. V. Poor, “Iterative Multiuser Receivers for CDMA Channels: An EM-Based Approach,” *IEEE Trans. Commun.*, vol. 44, pp. 1700–1710, Dec. 1996.

[102] I. S. Reed, “On a moment theorem for complex Gaussian processes,” *IEEE Trans. Inform. Theory*, vol. 8, pp. 194–195, April 1962.

[103] S. ten Brink, “Convergence behaviour of iteratively decoded parallel concatenated codes,” *IEEE Trans. Commun.*, vol. 49, pp. 1727–1737, Oct. 2001.

[104] M. Tüechler, R. Koetter, and A. Singer, “Turbo-equalization: principles and new results,” *IEEE Trans. Commun.*, vol. 50, pp. 754–767, May 2002.

[105] C. Hermosilla and L. Szczeciński, “EXIT charts for turbo receivers in MIMO systems,” in *Proceedings 7th International Symposium on Signal Processing and its Applications (ISSPA 2003)*, pp. 209 – 212, July 2003.

[106] J. Hagenauer, “The EXIT chart - Introduction to Extrinsic Information Transfer In Iterative Processing,” in *Proceedings 12th European Signal Processing Conference (EUSIPCO 2004)*, (Vienna, Austria), pp. 1541 – 1548, Sept. 2004.

[107] R. G. Gallager, *Low density parity check codes*. MIT Press, 1963.

[108] D. J. C. Mackay, “Good Error-Correcting Codes Based on Very Sparse Matrices,” *IEEE Trans. Inform. Theory*, vol. 45, pp. 399–431, March 1999.

[109] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, “Serial concatenation of interleaved codes: Performance analysis, design, and iterative decoding,” *IEEE Trans. Inform. Theory*, vol. 44, pp. 909–926, May 1998.

[110] D. Divsalar, S. Dolinar and F. Pollara, “Iterative turbo decoder analysis based on density evolution,” *IEEE J. Selected Areas Commun.*, vol. 19, pp. 891–907, May 2001.

[111] T. Richardson and R. Urbanke, “The capacity of low density parity check codes under message passing decoding,” *IEEE Trans. Inform. Theory*, vol. 47, pp. 599–618, Feb. 2001.

[112] T. Richardson, A. Shokrollahi, and R. Urbanke, “Design of capacity-approaching irregular low-density parity check codes,” *IEEE Trans. Inform. Theory*, vol. 47, pp. 619–637, Feb. 2001.

[113] N. Wiberg, *Codes and Decoding on General Graphs*. Dept. Electrical Engineering, Linkoping University, Sweden, 1996.

[114] M. Tüchler and J. Hagenauer, “EXIT charts of irregular codes,” in *2002 Conference on Information Sciences and Systems*, March 2002.

[115] H. El Gamal and A. R. Hammons, “Analyzing the turbo decoder using the Gaussian approximation,” *IEEE Trans. Inform. Theory*, vol. 47, pp. 671–686, Feb. 2001.

[116] Y. Li, “Simplified channel estimation for OFDM systems with multiple transmit antennas,” *IEEE Trans. Wireless Commun.*, vol. 1, pp. 67–75, Jan. 2002.

[117] O. Edfors, M. Sandell, J.-J. van de Beek ans Sarah Kate Wilson, and P. O. Börjesson, “OFDM Channel Estimation by Singular Value Decomposition,” *IEEE Trans. Commun.*, vol. 46, pp. 931–939, July 1998.

[118] W. C. Jakes, *Microwave Mobile Communications*. New York: Wiley, 1974.

[119] Y. Li, N. Seshadri, and S. Ariyavisitakul, “Channel estimation for OFDM systems with transmit diversity in mobile wireless channels,” *IEEE J. Selected Areas Commun.*, vol. 17, pp. 461–471, March 1999.

[120] L. Zhou and M. Nakamura, “Channel estimation of multiple transmit antennas for OFDM systems with cyclic delay preamble,” in *VTC2005-Fall*, vol. 1, pp. 583–587, Sept. 2005.