

國立交通大學

電信工程學系

博士論文

正交分頻多工系統之通道估計與資料檢測
技術研究

A Study on Channel Estimation and Data
Detection Techniques for OFDM Systems

研究生：古孟霖

指導教授：黃家齊

中華民國九十八年七月

正交分頻多工系統之通道估計與資料檢測
技術研究

A Study on Channel Estimation and Data
Detection Techniques for OFDM Systems

研究生：古孟霖

Student: Meng-Lin Ku

指導教授：黃家齊 博士

Advisor: Dr. Chia-Chi Huang



A Dissertation

Submitted to Institute of Communication Engineering
College of Electrical and Computer Engineering
National Chiao Tung University
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in
Communication Engineering
Hsinchu, Taiwan

2009 年 7 月

正交分頻多工系統之通道估計與資料檢測 技術研究

研究生：古孟霖

指導教授：黃家齊 博士

國立交通大學

電信工程學系



摘要

正交分頻多工是可在無線通道下達高速資料傳輸的有效技術，應用多輸入多輸出技術於正交分頻多工系統被視為下一代無線通訊提升系統效能的熱門方式。於本論文中，吾人研究正交分頻多工系統中通道估計與資料檢測技術，研究主題包括多輸入多輸出通道之領航訊號設計、慢速時變通道之通道估計與追蹤以及快速時變通道之資料檢測。

本論文可分為四部分，第一部分提出以互補碼領航訊號為基礎之空時區塊碼-正交分頻多工系統。於此系統，一組預先定義順序之互補碼與資料訊號同時傳送，做為兩根天線傳送分集系統之領航訊號，使用於接收機端估計通道以達最佳資料檢測。吾人設計完整接收機架構，分析理論系統效能，同時利用電腦模擬來驗證系統於行動無線電衰退通道下之效能。

於論文第二部分中，吾人由牛頓法推導於空時區塊碼-正交分頻多工系統中以決策回饋離散傅立葉轉換為基礎之通道估計方法，藉由導證過程，證明牛頓法與以決策回饋離散傅立葉轉換為基礎方法之間的等效。吾人亦使用電腦模擬驗證位元錯誤率及正規化方差效能來展現兩方法之間的等效，此結論於傳統正交分頻多工系統亦成立。

於論文第三部分中，吾人研究於行動無線通道下正交分頻多工系統採用空時區塊碼之通道估計，根據典型以離散傅立葉轉換為基礎之通道估計方法，提出一涵蓋兩階段處理方法。於初始階段，吾人使用多重路徑干擾消除技術估得多重路徑延遲與複數增益，於追縱階段，吾人發展一種改善的以決策迴饋離散傅立葉轉換為基礎之通道估計方法，此方法應用少量相嵌於正交分頻多工資料符元之領航載波，在第一次迭代時，形成最佳梯度向量來減緩錯誤蔓延效應，並且利用近似的權重矩陣以降低反矩陣計算複雜度。吾人經由電腦模擬兩發射天線一接收天線之空時區塊碼-正交分頻多工系統以驗證所提出的方法，結果顯示所提出方法不僅優於典型以離散傅立葉轉換為基礎之方法，亦優於以空時區塊碼為基礎之最小均方差方法及卡爾曼濾波方法。模擬結果亦證明所提出方法可達成顯著的訊號雜訊比效能改善，尤其在使用高階調變方式（例如：十六點正交振幅調變）於高車速環境下。

於本論文最後部分中，吾人憑藉最大期望值演算法，來處理時變多路徑通道對於正交分頻多工系統以及位元交錯調變碼-正交分頻多工系統所造成的載波間干擾問題。吾人首先在頻域上分析載波間干擾以便使用減少的參數集，根據此分析，導證最大期望值演算法用於最大似然資料檢測。吾人又針對正交分頻多工系統提出最大似然-最大期望值接收機及位元交錯調變碼-正交分頻多工系統提出渦輪-最大期望值接收機，其主要概念在於將所提出之最大期望值演算法與群式載波間干擾消除方法結合，用以減少計算複雜度及獲得時間分集益處。不同於最大似然-最大期望值接收機，渦輪-最大期望值接收機藉由渦輪原理，進一步利用軟輸出維特比演算法與最大後驗之最大期望值檢測器交換訊息。電腦模擬證實所提出之二個接收機顯然勝於傳統一階等化器，且渦輪-最大期望值接收機的效能在正規化最大都卜勒頻率為 0.1 時，能逼近匹配濾波器界限。

A Study on Channel Estimation and Data Detection Techniques for OFDM Systems

Student: Meng-Lin Ku

Advisor: Dr. Chia-Chi Huang

Department of Communication Engineering

National Chiao Tung University



Abstract

Orthogonal frequency division multiplexing (OFDM) is an effective technique for high data rate transmission over wireless channels. Employing multiple-input multiple-output (MIMO) techniques in OFDM systems is viewed as a popular way to improve system performance for the next generation wireless communications. In this dissertation, we investigate channel estimation and data detection techniques for OFDM systems, covering the research topics of pilot signal designs for MIMO channels, channel estimation and tracking for slowly time-varying channels, and data detection for fast time-varying channels.

This dissertation is divided into four parts. The first part presents a complementary codes (CC) pilot-based space-time block code (STBC)-OFDM system. In this system, a pair of complementary codes transmitted in a pre-defined order with the OFDM data signals is used as the pilot signals in a

two-antenna transmit diversity system, and used to estimate the channels for optimal data detection at the receiver side. A complete receiver architecture has been designed, the theoretical system performance has been analyzed, and computer simulations have been used to verify the performance of the system in mobile radio fading channels.

In the second part, we derive the decision-feedback (DF) discrete Fourier transform (DFT)-based channel estimation method from Newton's method for STBC-OFDM systems. Through our derivation, the equivalence between Newton's method and the DF DFT-based method is established. Computer simulations are also used to demonstrate the equivalence of the two methods in terms of bit error rate (BER) and normalized square error (NSE) performance. Finally, the results presented in this part also hold for conventional OFDM systems.

In the third part, we investigate channel estimation for OFDM systems with STBC in mobile wireless channels. Our proposed method consists of two-stage processing and is developed on the basis of the classical DFT-based channel estimation method. In the initialization stage, we employ a multipath interference cancellation (MPIC) technique to estimate multipath delays and multipath complex gains. In the tracking stage, we develop a refined DF DFT-based channel estimation method in which a few pilot tones inserted in OFDM data symbols are applied to form an optimal gradient vector at the first iteration such that the error propagation effect is mitigated. In order to reduce computational complexity, an approximate weighting matrix is adopted to avoid matrix inversion. We demonstrate the proposed method

through computer simulation of an STBC-OFDM system with two transmit antennas and a single receive antenna. The results show that our method outperforms the classical DFT-based method, the STBC-based minimum mean square error (MMSE) method, and the Kalman filtering method as well, and that significant signal-to-noise ratio (SNR) performance improvement can be achieved, especially when a high-level modulation scheme, e.g. 16-quadrature amplitude modulation (QAM), is adopted in low-mobility environments.

In the final part, we resort to the expectation-maximization (EM) algorithm to tackle the inter-carrier interference (ICI) problem, caused by time-variant multipath channels, for both the OFDM systems and the bit-interleaved coded modulation (BICM)-OFDM systems. We first analyze the ICI in frequency domain with a reduced set of parameters, and following this analysis, we derive an EM algorithm for maximum likelihood (ML) data detection. An ML-EM receiver for OFDM systems and a TURBO-EM receiver for BICM-OFDM systems are then developed to reduce computational complexity and to exploit temporal diversity, the main idea of which is to integrate the proposed EM algorithm with a groupwise ICI cancellation method. Compared with the ML-EM receiver, the TURBO-EM receiver further employs a soft-output Viterbi algorithm (SOVA) decoder to exchange information with a maximum a posteriori (MAP) EM detector through the turbo principle. Computer simulation demonstrates that the two proposed receivers clearly outperform the conventional one-tap equalizer, and the performance of the TURBO-EM receiver is close to the matched-filter bound even at a normalized maximum Doppler frequency up to 0.1.

Acknowledgements

I would like to express sincere gratitude to my advisor, Dr. Chia-Chi Huang, for his constant support and gracious guidance throughout my research works. His constructive suggestion and persistent help enables me to complete my dissertation successfully. I would also thank him for his encouragement to me in many times of frustration.

A debt of gratitude is owed to my colleagues and friends in Wireless Communication Laboratory at National Chiao Tung University. Many thanks are due to Dr. Chih-Min Yu, Dr. Shin-Yuan Wang, Dr. Shiang-Jiun Lin, Yong-Ting Chen, Chia-Hui Lin, Shin-Yi Tu as well as other colleagues for providing their invaluable help and consultation in many aspects during my PhD course. A special thanks to Yunn-Wen Chen, Ching-Kai Li, Che-Fang Yeh, Wen-Chuan Chen and Siao-Yi Jhong for their active contribution to my research works. Working with them has been a pleasant and memorable experience.

I would like to express my appreciation to Professors Wen-Rong Wu, Chin-Liang Wang, Shyh-Jye Jou, Sau-Gee Chen, Jeng-Kuang Hwang and Jia-Chin Lin, for their time and very useful comments on improving this dissertation.

Finally, yet importantly, this dissertation is dedicated to my lovely family. I am deeply indebted to my parents, my brother and my sister for their understanding, patience, and unwavering love.

Meng-Lin Ku
Hsinchu, Taiwan

Contents

Chinese Abstract	i
Abstract	iii
Acknowledgements	vi
Contents	vii
List of Tables	xii
List of Figures	xiii
Acronyms	xvii
Symbols	xxi
1 Introduction	1
1.1 Problem and Motivation	4
1.2 Organization of the Dissertation	10
2 A Complementary Codes Pilot-Based Transmit Diversity Tech- nique for OFDM Systems	12



2.1	Literature Survey and Motivation	12
2.2	CC Pilot Signals for Two Transmit Antenna Systems	14
2.3	CC Pilot-Based STBC-OFDM Systems: Transmitter Architecture	17
2.4	CC Pilot-Based STBC-OFDM Systems: Receiver Architecture	18
2.4.1	Fine Data Detection	21
2.4.2	Channel Estimation	22
2.4.3	Computational Complexity	24
2.5	Performance Analysis	24
2.5.1	Time-Varying Effect of Two-Path Channels	24
2.5.2	Performance Analysis of Coarse Data Detection	26
2.5.3	Performance Analysis of Channel Estimation and Fine Data Detection	28
2.6	Computer Simulation	32
2.6.1	Effect of Vehicle Speed	34
2.6.2	Effect of Path Selection	34
2.6.3	Analytic and Simulated BER Performance	35
2.7	Summary	39
3	On the Equivalence between DF DFT-based Channel Estimation Method and Newton's Method in OFDM Systems	40
3.1	Literature Survey and Motivation	40
3.2	STBC-OFDM Systems	41
3.3	DF DFT- Based Method and Newton's Method	43
3.3.1	DF DFT-Based Channel Estimation Method	43

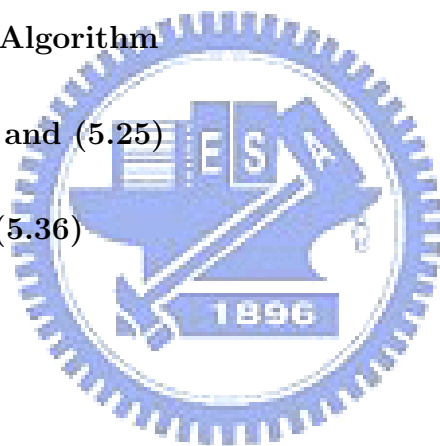
3.3.2	Channel Estimation via Newton’s Method	44
3.3.3	Equivalence between Newton’s Method and DF DFT- Based Method	49
3.4	Computer Simulation	54
3.4.1	BER Performance	54
3.4.2	NSE Performance	54
3.5	Summary	58

4 A Refined Channel Estimation Method for STBC-OFDM

	Systems in Low-Mobility Wireless Channels	59
4.1	Literature Survey and Motivation	59
4.2	STBC-OFDM Systems	62
4.2.1	Transmitted Signals	62
4.2.2	Channel Model	65
4.2.3	Received Signals	65
4.3	Proposed Channel Estimation Method	66
4.3.1	Initialization Stage: The MPIC-Based Decorrelation Method	66
4.3.2	Equivalence between DF DFT-Based Method and New- ton’s Method	69
4.3.3	Refined DF DFT-Based Channel Estimation	71
4.4	Computational Complexity	75
4.5	Computer Simulation	76
4.5.1	NSE Performance of MPIC-Based Decorrelation Method	80
4.5.2	BER Performance	81

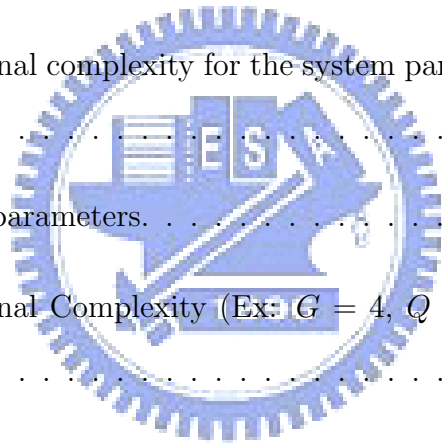
4.5.3	Effect of Normalized Maximum Doppler Frequency . . .	82
4.5.4	Effect of Number of Pilot Tones	82
4.5.5	Effect of Number of Data Subcarriers Used	83
4.5.6	Average Number of Iterations	83
4.6	Summary	93
5	EM-based Iterative Receivers for OFDM and BICM-OFDM	
	Systems in Doubly Selective Channels	94
5.1	Literature Survey and Motivation	94
5.2	System Model	98
5.2.1	Transmitted and Received Signals	98
5.2.2	Modeling of ICI in Frequency Domain	100
5.3	EM-based Data Detection Method	102
5.4	Implementation: EM-based Iterative Receivers	106
5.4.1	ML-EM Receiver for OFDM Systems	106
5.4.2	TURBO-EM Receiver for BICM-OFDM Systems . . .	112
5.4.3	Initial Setting and Channel Estimation Update	114
5.4.4	Computational Complexity	115
5.5	Computer Simulation	116
5.5.1	BER Performance of ML-EM Receiver	118
5.5.2	Effect of Group Size	120
5.5.3	BER Performance of TURBO-EM Receiver	120
5.5.4	FER Performance of TURBO-EM Receiver	121
5.6	Summary	128
6	Conclusions	129

Bibliography	131
Appendices	143
A Complementary Sequences	143
B Proof of (2.37) and (2.38)	146
C Proof of (3.21)	148
D Explanation of Hessian Matrix	149
E Review of EM Algorithm	151
F Proof of (5.24) and (5.25)	154
G Calculation of (5.36)	156
Vita	157
Publication List	158



List of Tables

2.1	Simulation parameters.	33
3.1	Simulation parameters.	55
4.1	Computational complexity for the system parameters given in Section 4.5.	77
4.2	Simulation parameters.	78
5.1	Computational Complexity (Ex: $G = 4$, $Q = 4$, $\gamma = 1$, and $L = 6$).	117
5.2	Simulation parameters.	119



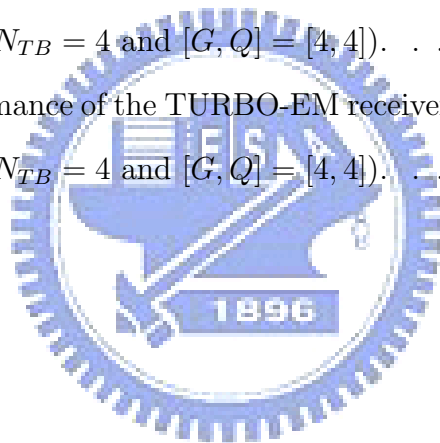
List of Figures

1.1	Transmitter of basic OFDM systems.	5
1.2	Receiver of basic OFDM systems.	7
2.1	Transmitter architecture of CC pilot-based STBC-OFDM systems.	19
2.2	Receiver architecture of CC pilot-based STBC-OFDM systems.	19
2.3	Fine data detection functional block.	21
2.4	Channel estimation functional block.	22
2.5	BER performance versus number of iterations at $E_b/\sigma_n^2=24dB$ and $f_D T_s=0.0111$	31
2.6	BER performance of CC pilot-based STBC-OFDM systems in the two-path fading channel with vehicular speed as a parameter ($N_p = 2$).	36
2.7	BER performance of CC pilot-based STBC-OFDM systems in the UMTS system defined fading channel with N_p as a parameter ($V = 120km/hr$).	37

2.8	The analytic and simulated BER performance of CC pilot-based STBC-OFDM systems in the two-path fading channel with vehicular speed as a parameter.	38
3.1	STBC-OFDM systems.	42
3.2	The block diagram of the DF DFT-based channel estimation method. (D is a delay component.)	44
3.3	Equivalence between (a) Newton's method and (b) the DF DFT-based method.	53
3.4	BER performance of the two methods.	56
3.5	NSE performance of the two methods ($f_d = 0.05$).	57
4.1	(a) STBC-OFDM system (b) OFDM frame format.	63
4.2	The MPIC-based decorrelation method in the initialization stage. (G is the ratio of the CP length to the useful OFDM symbol time, and $IDFT\{\cdot\}$ is a K -point IDFT operation.)	67
4.3	Block diagram of the refined DF DFT-based channel estimation method in the tracking stage. (The subscript " p " is to indicate that the calculation is only associated with the pilot subcarrier set.)	72
4.4	NSE performance of the MPIC-based decorrelation method in the initialization stage ($v_e = 240km/hr$).	84
4.5	BER performance for QPSK modulation in the two-path channel at $v_e = 240km/hr$ ($ \mathbf{J} = 8$ and $ \mathbf{\Theta} = \mathbf{Q} = 192$).	85
4.6	BER performance for 16QAM modulation in the two-path channel at $v_e = 240km/hr$ ($ \mathbf{J} = 8$ and $ \mathbf{\Theta} = \mathbf{Q} = 192$).	86

4.7	BER performance for 16QAM modulation in the ITU Veh-A channel at $v_e = 240km/hr$ ($ \mathbf{J} = 8$ and $ \Theta = \mathbf{Q} = 192$). . . .	87
4.8	BER versus normalized maximum Doppler frequency in the ITU Veh-A channel for QPSK modulation ($ \mathbf{J} = 8$, $ \Theta = \mathbf{Q} = 192$, and $E_b/N_o=16dB$).	88
4.9	BER versus normalized maximum Doppler frequency in the ITU Veh-A channel for 16QAM modulation ($ \mathbf{J} = 8$, $ \Theta = \mathbf{Q} = 192$, and $E_b/N_o=22dB$).	89
4.10	BER versus number of pilot tones used in the ITU Veh-A channel for 16QAM modulation ($v_e = 240km/hr$, $ \Theta = \mathbf{Q} $, and $E_b/N_o=20dB$).	90
4.11	BER versus number of data subcarriers used in the ITU Veh-A channel for 16QAM modulation ($v_e = 240km/hr$, $ \mathbf{J} = 8$, and $E_b/N_o=20dB$).	91
4.12	Average number of iterations versus $ \Theta $ in the ITU Veh-A channel for 16QAM modulation ($v_e = 240km/hr$, $ \mathbf{J} = 8$, and $E_b/N_o=20dB$).	92
5.1	BICM-OFDM systems.	98
5.2	EM-based data detection method.	107
5.3	(a) ML-EM receiver for OFDM systems (b) An illustration for group detection.	108
5.4	TURBO-EM receiver for BICM-OFDM systems.	111
5.5	Initialization procedure for ML-EM and TURBO-EM receivers.	114

5.6	BER performance of the ML-EM receiver in the two-path channel ($N_{ML} = 3$ and $[G, Q] = [4, 4]$).	122
5.7	BER performance of the ML-EM receiver in the ITU Veh-A channel ($N_{ML} = 3$ and $[G, Q] = [4, 4]$).	123
5.8	BER performance of the ML-EM receiver with channel estimation update for various $[G, Q]$ ($N_{ML} = 3$).	124
5.9	BER performance of the TURBO-EM receiver in the two-path channel ($N_{TB} = 3$ and $[G, Q] = [4, 4]$).	125
5.10	BER performance of the TURBO-EM receiver in the ITU Veh-A channel ($N_{TB} = 4$ and $[G, Q] = [4, 4]$).	126
5.11	FER performance of the TURBO-EM receiver in the ITU Veh-A channel ($N_{TB} = 4$ and $[G, Q] = [4, 4]$).	127



Acronyms

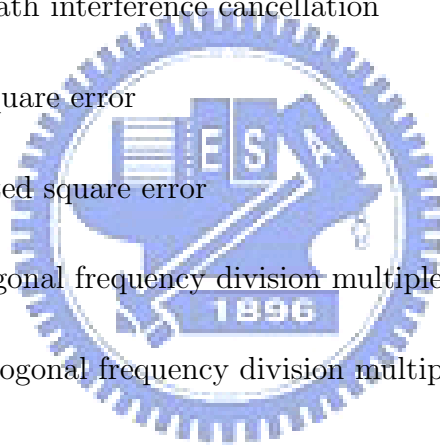
- 1G : first-generation
- 2G : second-generation
- 3G : third-generation
- 3GPP : Third Generation Partnership Project
- 4G : fourth-generation
- AR : autoregressive
- AWGN : additive white Gaussian noise
- BER : bit error rate
- BICM : bit-interleaved coded modulation
- BPSK : binary phase-shift keying
- CC : complementary codes
- CDMA : code division multiple access
- CIR : channel impulse response



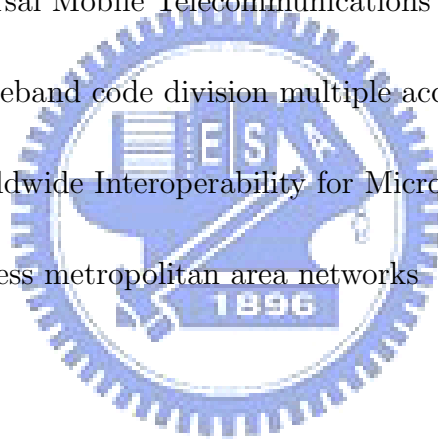
- CP : cyclic prefix
- CSI : channel state information
- DF : decision-feedback
- DFT : discrete Fourier transform
- ECLL : expected complete log-likelihood
- EM : expectation-maximization
- FER : frame error rate
- FFT : fast Fourier transform
- GI : guard interval
- GSM : Global System for Mobile Communications
- IAI : inter-antenna interference
- ICI : inter-carrier interference
- IDFT : inverse discrete Fourier transform
- IFFT : inverse fast Fourier transform
- ISI : inter-symbol interference
- ITU : International Telecommunication Union
- LLR : log-likelihood ratio
- LS : least-square



- LTE : Long-Term Evolution
- MAP : maximum *a posteriori*
- MC : multicarrier modulation
- MIMO : multiple-input multiple-output
- ML : maximum likelihood
- MMSE : minimum mean square error
- MPIC : multipath interference cancellation
- MSE : mean square error
- NSE : normalized square error
- OFDM : orthogonal frequency division multiplexing
- OFDMA : orthogonal frequency division multiple access
- P/S : parallel-to-serial
- PAPR : peak-to-average power ratio
- QAM : quadrature amplitude modulation
- QPSK : quadratic phase shift keying
- S/P : serial-to-parallel
- SC : single-carrier
- SISO : single-input single-output



- SM : spatial multiplexing
- SNR : signal-to-noise ratio
- SOVA : soft-output Viterbi algorithm
- STBC : space-time block code
- STC : space-time coding
- TDMA : time division multiple access
- UMTS : Universal Mobile Telecommunications System
- WCDMA : wideband code division multiple access
- WiMAX : Worldwide Interoperability for Microwave Access
- WMAN : wireless metropolitan area networks



Symbols

By convention, boldface letters are used for matrices, vectors, and sets. The superscripts $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$ and $(\cdot)^{-1}$ stand for complex conjugate, transpose, Hermitian, matrix inversion, respectively. The notation $\Re(\cdot)$ takes the real part of (\cdot) . The notation \mathbf{I}_N presents an $N \times N$ identity matrix. We denote the Kronecker delta function as $\delta[n]$. Besides, we denote the dimension of the vector \mathbf{x} as $|\mathbf{x}|$. The notation $\{\cdot\}$ denotes a set, e.g. a set $\mathbf{x} = \{x_1, \dots, x_{|\mathbf{x}|}\}$, where $|\mathbf{x}|$ is cardinality of the set \mathbf{x} . Further, $diag\{\mathbf{x}\}$ denotes the diagonal matrix with vector \mathbf{x} on the diagonal, and $diag\{\mathbf{X}_1, \dots, \mathbf{X}_M\}$ denotes the block diagonal matrix with the submatrices $\mathbf{X}_1, \dots, \mathbf{X}_M$ on the diagonal. The notations $\mathbb{E}[\cdot]$ and $\text{Var}[\cdot]$ denote the calculation of expectation and variance, respectively. Finally, $j = \sqrt{-1}$.

Chapter 1

Introduction

The second-generation (2G) wireless systems were introduced in the early 1990s, and considered as the prominent evolution of cellular systems in telecommunication. Because 2G systems feature the implementation of digital technology, they offer better voice quality and elementary data service, as compared with analog communication provided in the first-generation (1G) radio systems. Today, the European digital cellular system, referred to as Global System for Mobile Communications (GSM), is undoubtedly the most successful 2G systems in the world, and still dominates the cellular service market with over 250 million subscribers in about 120 countries. With the increasing demands for multimedia-level data rates in mobile communication, the third-generation (3G) systems, based on wideband code division multiple access (WCDMA) radio technology, are now being progressively deployed on a large scale all over the world. The 3G systems upgrade the existing 2G network to provide data rates up to 144kb/s for high-mobility environments, 284kb/s for low-mobility environments and 2Mb/s for stationary environ-

ments, and they also aim to greatly expand cell coverage. A comprehensive introduction of air interface evolution for cellular communication systems can be referred to [1] for details. A desire for high data rate transmission motivates the development of the fourth-generation (4G) systems. In order to get richer multimedia services in the future, the 4G systems attempt to provide a peak data rate up to a level of 100Mb/s for wide area coverage with high mobility and 1Gb/s for local area coverage with low mobility. As data rate increases, the effect of time dispersion in wireless channels becomes much more significant and will cause more severe inter-symbol interference (ISI). As a result, a relatively complicate time-domain equalizer in single-carrier (SC) systems, such as GSM, is needed, which will raise the cost of implementation. In 3G systems, although a rake receiver can be used to combat dispersive fading by resolving and combining the multipath, the receiver is not suitable over a broadband channel due to the appearance of excessive multipath interference. On the other hand, orthogonal frequency division multiplexing (OFDM) is an attractive choice to meet the requirement for high-data-rate transmission in future 4G systems due to its inherent ability to compensate for multipath fading [2]. OFDM was originally proposed in the 1970s, and it is a special form of multicarrier modulation (MC) scheme in which a large number of narrowband and orthogonal subcarriers are used to transmit data symbols in parallel through the use of inexpensive inverse fast Fourier transform (IFFT) and fast Fourier transform (FFT). As a consequence, it converts a frequency selective fading channel into several flat fading channels, and allows for a simpler one-tap equalizer at the receiver side.

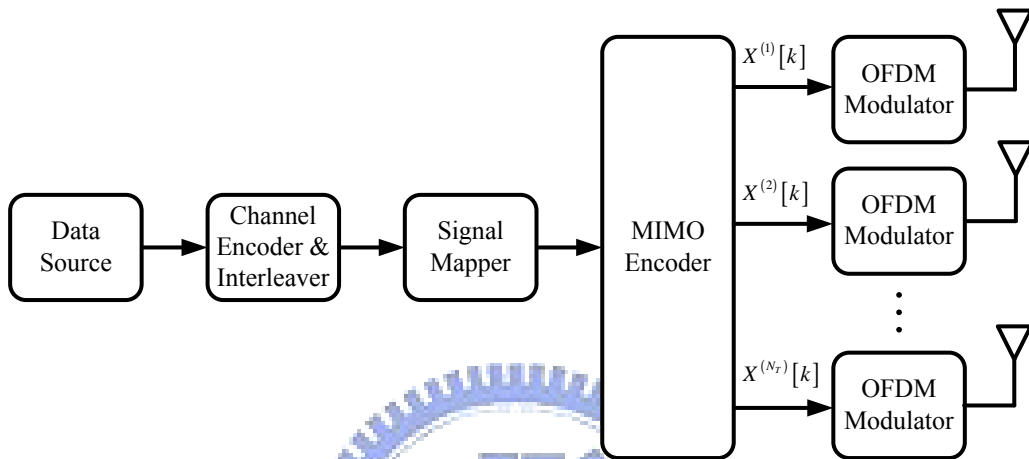
Besides, the flexibility of OFDM enables the use of frequency-domain adaptation and multi-antenna solutions. Based on all these advantages, OFDM has been chosen as a potential air interface candidate for two popular 4G cellular systems [3–5]:

- Worldwide Interoperability for Microwave Access (WiMAX) systems in IEEE 802.16 standard.
- Long-Term Evolution (LTE) of 3G systems in Third Generation Partnership Project (3GPP).

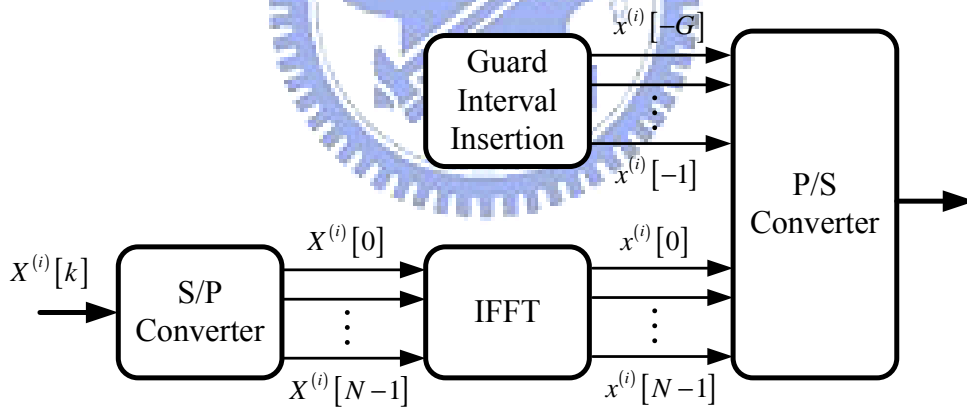
The capacity of wireless communication can be substantially boosted if a communication system employs multiple antennas at the transmitter and receiver, also known as multiple-input multiple-output (MIMO) system. It is proved in [6] that, the capacity of MIMO systems is linearly increased with the minimum number of transmit and receive antennas in a flat fading channel. In order to improve cell coverage and data rate, various MIMO technologies should be supported as a well-integrated part of 4G systems, rather than just an add-on to the specifications [7,8]. Since OFDM is well suited for MIMO processing, the combination of MIMO and OFDM is a nature choice for wideband transmission to obtain diversity gains or multiplexing gains in the spatial dimension. With this popularity, MIMO-OFDM technology has been adopted in several standards, and it is the most attractive air interface for high performance 4G broadband wireless communications [9, 10]. The primary goal of this dissertation is to deal with channel estimation and data detection for OFDM systems, and detailed descriptions of problem and motivation are given in the next section.

1.1 Problem and Motivation

Figure 1.1 shows a transmitter of basic OFDM systems with N_T transmit antennas. The information bits are first encoded by a channel encoder and fed into an interleaver in order to achieve good coding gains as well as diversity gains. Afterwards, the coded bit stream is mapped onto constellation points, e.g. quadratic phase shift keying (QPSK) or quadrature amplitude modulation (QAM), using a signal mapper, and then processed by a MIMO encoder. The MIMO encoder transforms a symbol stream into N_T parallel substreams, and it can be implemented in a number of different ways, depending on application requirements. Herein, we focus on two key schemes: space-time coding (STC) and spatial multiplexing (SM). More details on these two schemes can be found in [11]. In the STC scheme, the same symbol stream is encoded into different substreams across multiple transmit antennas to achieve spatial diversity gains. Of particular interest is Alamouti's space-time block code (STBC) in two transmit-antenna systems [12]. On the other hand, the SM scheme aims at increasing data rate by spatially multiplexing a symbol stream into independent substreams across transmit antennas. After the MIMO processing, these substreams are simultaneously transmitted using the corresponding antennas through the same process of OFDM transmission. As illustrated in Figure 1.1(b), in an OFDM modulator at the i th transmit antenna, a symbol substream $\{X^{(i)}[k]\}$ is first passed through a serial-to-parallel (S/P) converter, and modulated by an inverse fast Fourier transform (IFFT) to produce time domain samples $\{x^{(i)}[n]\}$. Then, in order to avoid the ISI caused by multipath channels, a guard inter-



(a) Transmitter architecture



(b) OFDM modulator

Figure 1.1: Transmitter of basic OFDM systems.

val (GI) $x^{(i)} [n]$, for $n = -1, \dots, -G$, with length longer than delay spread, is attached at the beginning of these samples through a parallel-to-serial (P/S) converter to generate an OFDM symbol. Although a silent GI can be used to take care of ISI, inter-carrier interference (ICI) remains a critical issue due to the loss of orthogonality among subcarriers. If a cyclic prefix (CP) is used for the GI, ICI can also be avoided. In this regard, the CP is more widely adopted in current standards, and an OFDM symbol transmitted from the i th antenna is given by

$$x^{(i)} [n] = \frac{1}{N} \sum_{k=0}^{N-1} X^{(i)} [k] e^{j \frac{2\pi k n}{N}} \quad (1.1)$$

for $n = -G, \dots, N - 1$. The impulse response of wireless channels between the i th transmit antenna and the j th receive antenna can be expressed as

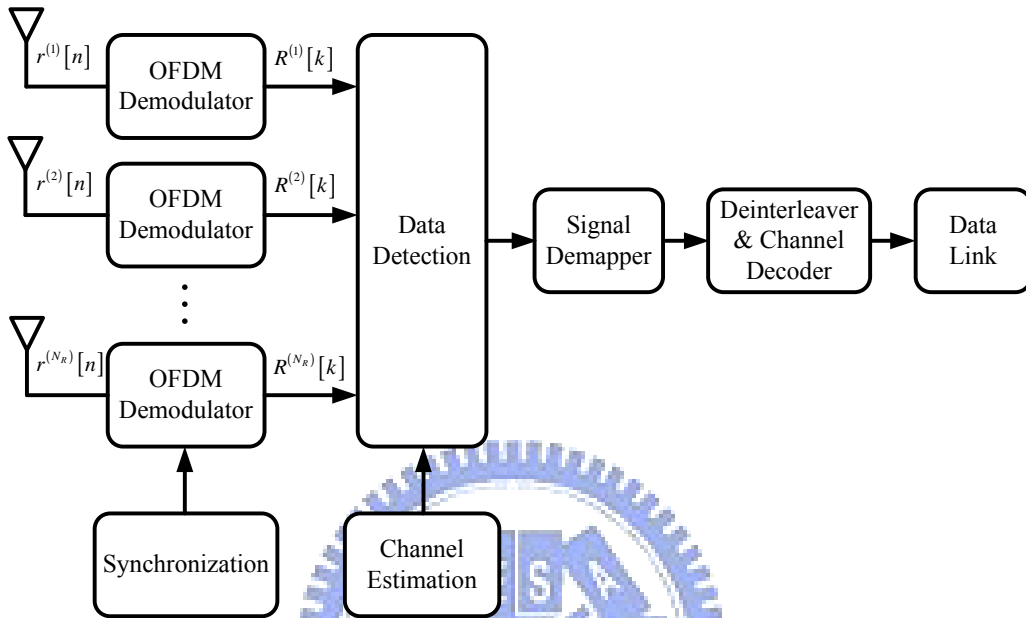
$$h^{(j,i)} [n, \tau] = \sum_{l=0}^{L-1} h^{(j,i)} [l, n] \delta [\tau - l] \quad (1.2)$$

where $h^{(j,i)} [l, n]$ is complex gain of the l th path, associated with path delay l , and L is the number of propagation paths. Note that $h^{(j,i)} [l, n]$ changes with time index n when Doppler spread exists.

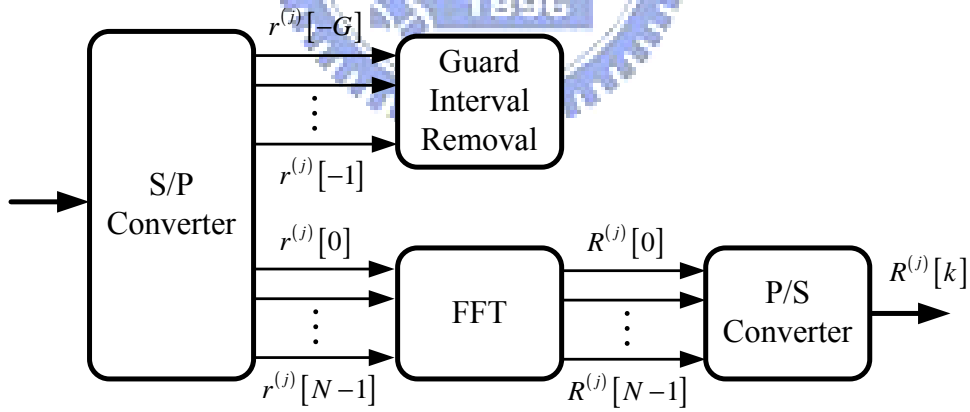
The receiver of basic OFDM systems with N_R receive antennas is shown in Figure 1.2. The received signal at the j th receive antenna can be represented as

$$r^{(j)} [n] = \sum_{i=1}^{N_T} \sum_{l=0}^{L-1} h^{(j,i)} [l, n] x^{(i)} [((n - l))_N] + z^{(j)} [n] \quad (1.3)$$

where $((\cdot))_N$ represents modulo N operation and $z^{(j)} [n]$ is complex additive white Gaussian noise (AWGN) at the j th antenna. At the receiver side, after S/P conversion, GI removal and FFT operation, the received signal of each



(a) Receiver architecture



(b) OFDM demodulator

Figure 1.2: Receiver of basic OFDM systems.

antenna arm is transformed into frequency domain as follows:

$$R^{(j)} [k] = \sum_{n=0}^{N-1} r^{(j)} [n] e^{-j\frac{2\pi kn}{N}} \quad (1.4)$$

for $k = 0, \dots, N-1$. By substituting (1.1) and (1.3) into (1.4), we can obtain

$$R^{(j)} [k] = \sum_{i=1}^{N_T} \sum_{l=0}^{L-1} \sum_{m=0}^{N-1} X^{(i)} [m] \alpha^{(j,i)} [k, m, l] e^{-j\frac{2\pi ml}{N}} + Z^{(j)} [k] \quad (1.5)$$

where $Z^{(j)} [k]$ and $\alpha^{(j,i)} [k, m, l]$ can be evaluated by

$$Z^{(j)} [k] = \sum_{n=0}^{N-1} z^{(j)} [n] e^{-j\frac{2\pi kn}{N}} \quad (1.6)$$

$$\alpha^{(j,i)} [k, m, l] = \frac{1}{N} \sum_{n=0}^{N-1} h^{(j,i)} [l, n] e^{-j\frac{2\pi n((k-m)N)}{N}} \quad (1.7)$$

We then define

$$H^{(j,i)} [k, m] = \sum_{l=0}^{L-1} \alpha^{(j,i)} [k, m, l] e^{-j\frac{2\pi ml}{N}} \quad (1.8)$$

and rewrite (1.5) as follows:

$$\begin{aligned} R^{(j)} [k] &= \underbrace{H^{(j,q)} [k, k] X^{(q)} [k]}_{\text{desired term}} \\ &+ \underbrace{\sum_{i=1, i \neq q}^{N_T} H^{(j,i)} [k, k] X^{(i)} [k]}_{\text{IAI term}} \\ &+ \underbrace{\sum_{i=1}^{N_T} \sum_{m=0, m \neq k}^{N-1} H^{(j,i)} [k, m] X^{(i)} [m]}_{\text{ICI term}} \\ &+ \underbrace{Z^{(j)} [k]}_{\text{noise term}} \end{aligned} \quad (1.9)$$

It is observed from (1.9) that for detecting the q th transmitted signal $X^{(q)} [k]$, the demodulated signal $R^{(j)} [k]$ suffers from not only noise but also inter-carrier interference (ICI) as well as inter-antenna interference (IAI). Essentially, the ICI arises from Doppler spread, while the IAI is due to multiple antennas transmission, and these two unwanted signals will severely degrade the performance of a mobile OFDM receiver. When wireless channels are quasi-static, i.e. channel gains are time-invariant, (1.7) and (1.8) become as

$$\alpha^{(j,i)} [k, m, l] = \begin{cases} h^{(j,i)} [l], & \text{for } k = m \\ 0, & \text{for } k \neq m \end{cases} \quad (1.10)$$

$$H^{(j,i)} [k, m] = \begin{cases} \sum_{l=0}^{L-1} h^{(j,i)} [l] e^{-j \frac{2\pi ml}{N}}, & \text{for } k = m \\ 0, & \text{for } k \neq m \end{cases} \quad (1.11)$$

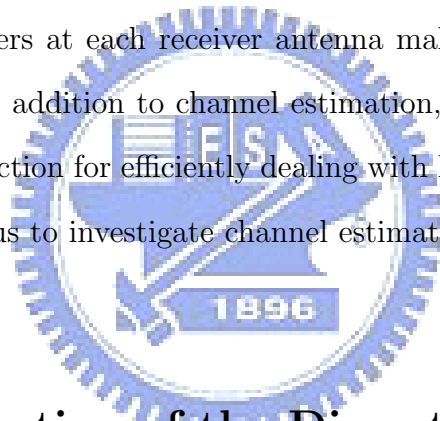
where $h^{(j,i)} [l, n]$ is a constant value of $h^{(j,i)} [l]$, for $n = 0, \dots, N - 1$. Then, (1.9) is reduced to

$$\begin{aligned} R^{(j)} [k] = & \underbrace{H^{(j,q)} [k, k] X^{(q)} [k]}_{\text{desired term}} \\ & + \underbrace{\sum_{i=1, i \neq q}^{N_T} H^{(j,i)} [k, k] X^{(i)} [k]}_{\text{IAI term}} \\ & + \underbrace{Z^{(j)} [k]}_{\text{noise term}} \end{aligned} \quad (1.12)$$

Moreover, the extension of (1.9) and (1.12) to single-input single-output (SISO)-OFDM systems is also straightforward by setting $N_T = 1$ and $N_R = 1$. As expected, there is no IAI problem in this special case.

There are several challenges in attempts to design an OFDM system. As depicted in Figure 1.2, the success of implementing a receiver hinges on

several basic issues, consisting of synchronization, channel estimation, data detection, and channel decoding, etc. In this dissertation, we will focus on channel estimation and data detection problem. From the aforementioned discussion, the knowledge of channel state information (CSI) is required for coherent data detection in (1.9) and (1.12). Hence, it is crucial to have accurate estimates of CSI, which is in general difficult to achieve over fast fading channels and MIMO channels. In fast fading channels, a more sophisticated receiver is needed to track rapid channel variation; otherwise, performance deterioration may occur. In MIMO channels, the intent to identify multiple channel parameters at each receiver antenna makes channel estimation more challenging. In addition to channel estimation, it is also necessary to investigate data detection for efficiently dealing with ICI and IAI. These observations motivate us to investigate channel estimation and data detection for OFDM systems.



1.2 Organization of the Dissertation

The rest of this dissertation is organized as follows. In Chapter 2, we use complementary codes (CC) to design pilot signals with minimum peak-to-average power ratio (PAPR) for channel estimation in MIMO systems. Then, we present a CC pilot-based STBC-OFDM system which transmits CC pilot signals together with OFDM data signals in time domain without sacrificing bandwidth efficiency. A complete receiver architecture for channel estimation and data detection is proposed and analyzed. Chapter 3 introduces a classical decision-feedback (DF) discrete Fourier transform (DFT)-based channel es-

timization method for STBC-OFDM systems. We then prove the equivalence between the DF DFT-based method and the Newton's method, and establish their relationship. In Chapter 4, we propose a two-stage channel estimation method for STBC-OFDM systems. In the initialization stage, a multipath interference cancellation (MPIC)-based decorrelation method is employed to identify significant channel taps. Through the equivalence discussed in Chapter 3, we develop a refined DF DFT-based channel estimation method in the tracking stage to improve bit error rate (BER) performance. Chapter 5 presents two expectation-maximization (EM)-based iterative receivers to mitigate ICI, introduced by Doppler effect, for both OFDM systems and bit-interleaved coded modulation (BICM)-OFDM systems. We derive an EM algorithm for maximum likelihood (ML) data detection. Towards the goal of reducing computational complexity, we then develop an ML-EM receiver for OFDM systems and a TURBO-EM receiver for BICM-OFDM systems. Chapter 6 draws some conclusions.

Chapter 2

A Complementary Codes

Pilot-Based Transmit Diversity

Technique for OFDM Systems

2.1 Literature Survey and Motivation

In a recent paper, STBC [12,13] has been suggested to improve the performance of an OFDM system [14]. Garg studied the degradation in the performance of the STC with imperfect channel estimation [15]. In general, channel estimation is an important issue in realizing a successful STBC-OFDM system. We all know that channel estimation can be performed at a receiver by inserting pilot signals into transmitted signals. There are several factors which have to be considered for the practical use of pilot signals in multiple transmit antenna systems. First, pilot signals should be sent in the same frequency band and at the same time with data signals to ensure the accuracy of channel estimation. Second, multiple CSIs must be derived

from the received signal in a system with multiple transmit antennas. To estimate multiple channels, antennas at the transmitter side can alternately transmit a single pilot signal or simultaneously transmit different pilot signals which have impulse-like auto-correlation and zero cross-correlation properties [16–19]. Third, the length of pilot signals should not be too short in order to achieve accurate channel estimation. In some special cases, pilot signals take the length of 2^P , where P is an integer, to meet with the requirement of fast signal processing algorithms. Finally, the values of pilot signals need to be taken from some predetermined signal sets such as $\{1, -1, j, -j\}$ to maintain constant amplitude and avoid the non-linear effect of an amplifier. The design of optimal pilot signals is an open problem [16–23]. Computer simulations are used in [16–19] to exhaustively search the pilot signals which are able to achieve MMSE channel estimation. [20] proposed a channel estimator based on the correlation of channel frequency response at adjacent frequencies. However, a large size of matrix inverse is required in this scheme. [21–23] extended the work in [20] to reduce the complexity of channel estimation. Besides, some channel estimation methods based on superimposed training sequences are proposed for SISO-OFDM systems [24–26]. At the expense of reducing power efficiency, these methods can not only significantly save bandwidth but also effectively track time-variant channels. Recently, channel estimation using superimposed training for MIMO-OFDM systems has been investigated [27, 28]. The PAPR of the OFDM signals with superimposed training is analyzed in [29], and it is demonstrated that the constant magnitude pilot sequences result in the best BER performance due to their

ability to lower the PAPR of the transmitted OFDM signal.

In this chapter, we suggest that CC can be used as pilot signals and transmitted in time domain for the purpose of channel estimation in a two transmit antenna system. The CC pilot signals satisfy the requirements of pilot signals we mentioned above, and at the same time, have the minimum PAPR. We will describe the functional block diagrams and simulate the performance of an STBC-OFDM system with CC pilot signals. The rest of this chapter is organized as follows. In Section 2.2, we will describe the CC pilot signals for a two transmit antenna system. In Section 2.3, We will introduce the transmitter architecture of a CC pilot-based STBC-OFDM system. The details of the receiver operation such as data detection, channel estimation, etc., are described in Section 2.4. The performance of the CC pilot-based STBC-OFDM system is then analyzed in Section 2.5. In Section 2.6, we show our computer simulation and performance evaluation results. Finally, concluding remarks are drawn in Section 2.7.

2.2 CC Pilot Signals for Two Transmit Antenna Systems

Binary CC was originally conceived by Golay for infrared multi-slit spectrometry applications [30,31]. More recently, these codes were also used in OFDM systems to reduce PAPR [32]. For details, please refer to Appendix A. We can define CC as follows. Let us consider a pair of equally long sequences $\{\alpha[n]\}$ and $\{\beta[n]\}$, for $n = 0, \dots, N - 1$, where N is the length of the two sequences. These sequences are called CC if their auto-correlations

satisfy the relationship:

$$\begin{aligned}
\Gamma [n] &\equiv \sum_{m=0}^{N-1} \{ \alpha [m] \alpha^* [((m-n))_N] \\
&\quad + \beta [m] \beta^* [((m-n))_N] \} \\
&= 2N \cdot \delta [n] \\
&= \begin{cases} 2N, & \text{for } n = 0 \\ 0, & \text{for } n \neq 0 \end{cases} \quad (2.1)
\end{aligned}$$

where $(\cdot)^*$ denotes the complex conjugate operation, $((\cdot))_N$ denotes the modulo N operation, and $\delta[n]$ is the Kronecker delta function.

The pilot signals of a two transmit antenna system can be constructed from a pair of CC. For simplicity, we assume that CC sequences $\{\alpha [n]\}$ and $\{\beta [n]\}$ are normalized such that their combined auto-correlation value $\Gamma [n] = \delta [n]$ as shown in (2.1). Consider a system with two transmit antennas and one receive antenna. Two signals $\{\alpha [n]\}$ and $\{-\beta [n]\}$ are simultaneously transmitted from the two transmit antennas over two independent frequency selective fading channels in the first time slot; the other two signals $\{\beta^* [((-n))_N]\}$ and $\{\alpha^* [((-n))_N]\}$ are then simultaneously transmitted over the same two channels in the second time slot. Furthermore, a CP is added before each transmitted pilot signal to avoid ISI and to preserve the circular convolution between the pilot signal and the channel impulse response in the time domain. Here, we assume that the two channels are quasi-static over the two transmission time slots. Hence, the received pilot signals in the frequency domain in the first and second time slot, $R_{T1} [k]$ and $R_{T2} [k]$, can be

expressed as

$$\begin{aligned} \begin{bmatrix} R_{T1}[k] \\ R_{T2}[k] \end{bmatrix} &= \begin{bmatrix} P_\alpha[k] - P_\beta[k] \\ P_\beta^*[k] \ P_\alpha^*[k] \end{bmatrix} \begin{bmatrix} H_1[k] \\ H_2[k] \end{bmatrix} + \mathbf{Z}[k] \\ &= \mathbf{P}[k] \begin{bmatrix} H_1[k] \\ H_2[k] \end{bmatrix} + \mathbf{Z}[k] \end{aligned} \quad (2.2)$$

for $k = 0, \dots, N - 1$, where $H_1[k]$ and $H_2[k]$ are the channel frequency responses from the two transmit antennas to the receive antenna, $P_\alpha[k]$ and $P_\beta[k]$ are the N -point DFT of the CC sequences $\{\alpha[n]\}$ and $\{\beta[n]\}$, $\mathbf{P}[k]$ is called a pilot matrix which is a unitary matrix, and $\mathbf{Z}[k]$ is an AWGN vector with zero mean and covariance matrix $\sigma_n^2 \mathbf{I}_2$, where \mathbf{I}_K is a $K \times K$ identity matrix. The received pilot signals are then multiplied by a matrix $\mathbf{P}^H[k]$, where $(\cdot)^H$ denotes the complex conjugate transpose operation, to obtain estimated CSIs in the frequency domain

$$\begin{aligned} \begin{bmatrix} \hat{H}_1[k] \\ \hat{H}_2[k] \end{bmatrix} &= \mathbf{P}^H[k] \begin{bmatrix} R_{T1}[k] \\ R_{T2}[k] \end{bmatrix} \\ &= \begin{bmatrix} H_1[k] \\ H_2[k] \end{bmatrix} + \mathbf{P}^H[k] \mathbf{Z}[k] \end{aligned} \quad (2.3)$$

for $k = 0, \dots, N - 1$. As a result, the mean square error (MSE) of channel estimation is given by

$$\begin{aligned} MSE &= E \left[\left| \hat{H}_j[k] - H_j[k] \right|^2 \right] \\ &= \sigma_n^2, \text{ for } j = 1, 2 \end{aligned} \quad (2.4)$$

Generally speaking, the pilot matrix can be any unitary matrix which satisfies the power constraint of $|P_\alpha[k]|^2 + |P_\beta[k]|^2 = 1$ for all k . For ex-

ample, $|P_\alpha[k]|^2 = |P_\beta[k]|^2 = 1/2$ for all k also satisfies the unitary matrix requirement. However, the CC pilot signals used in this dissertation have the minimum PAPR ($= 0dB$) in time domain, thus improve radio frequency power amplifier efficiency. In the case of $|P_\alpha[k]|^2 = |P_\beta[k]|^2 = 1/2$ for all k , the PAPR is $10\log N(dB)$. Moreover, we can easily extend the design of the pilot matrix to four transmit antenna systems by using a 2×2 unitary matrix as follows:

$$\mathbf{P}_4[k] = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{P}[k] & \mathbf{P}[k] \\ -\mathbf{P}[k] & \mathbf{P}[k] \end{bmatrix} \quad (2.5)$$

2.3 CC Pilot-Based STBC-OFDM Systems: Transmitter Architecture

The transmitter block diagram of a CC pilot-based STBC-OFDM system is shown in Figure 2.1. The block diagram shows two transmit antennas. The transmitted signal from each antenna consists of a data signal and a pilot signal. Now we describe the generation of the data signals. At the output of the signal mapper, the i th block of $2N$ data symbols $d^{(i)}[k]$ are separated into two data sub-blocks and represented as

$$X_F^{(i)}[k] = d^{(i)}[k] \quad (2.6)$$

and

$$X_S^{(i)}[k] = d^{(i)}[N+k] \quad (2.7)$$

for $k = 0, \dots, N - 1$, where N denotes number of subcarriers in an OFDM symbol, $X_F^{(i)} [k]$ and $X_S^{(i)} [k]$ are the k th data symbol of the first and second data sub-block, respectively. Here, we use Alamouti's STBC encoding method [12] to encode the two data sub-blocks, $X_F^{(i)} [k]$ and $X_S^{(i)} [k]$, in the sequence as described in [14]. An N -point inverse discrete Fourier transform (IDFT) unit is used in each arm of Figure 2.1 to transform the frequency domain data symbols into a time domain data signal. Afterward, we add the CC pilot signal as described in Section 2.2 to the time domain data signal. Both the data signal and the pilot signal are assumed to be of the same length and they are added synchronously to become an effective OFDM symbol with symbol duration T . The cyclic extension with time duration T_g of an effective OFDM symbol is then inserted as a GI to combat the ISI effect. Finally, a complete OFDM symbol with symbol duration T_s is converted into an analog signal with a digital-to-analog converter, filtered by a low-pass filter, up converted to radio frequency band, and transmitted in air with a pre-selected antenna, where we set $T_s = T + T_g$.

2.4 CC Pilot-Based STBC-OFDM Systems: Receiver Architecture

The receiver block diagram of the CC pilot-based STBC-OFDM system is shown in Figure 2.2. This receiver architecture consists mainly of a fine data detection functional block and a channel estimation functional block, along with other common blocks. After a radio frequency signal is received from an antenna, it is down converted to the equivalent baseband, low-pass filtered,

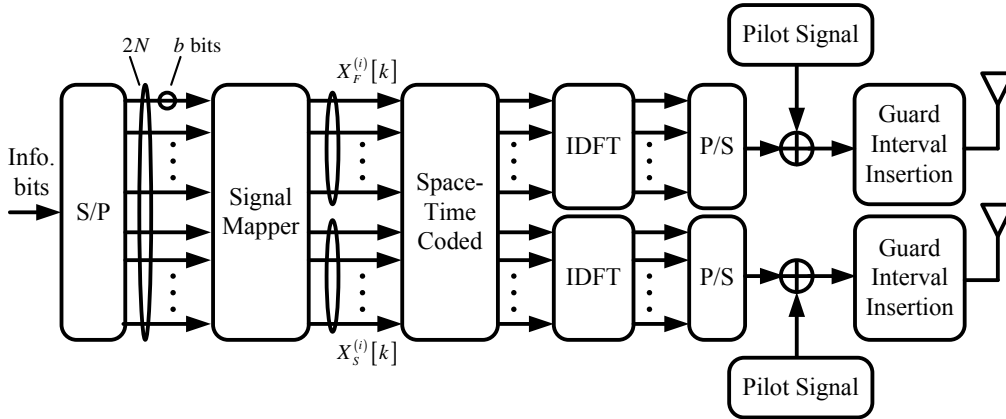


Figure 2.1: Transmitter architecture of CC pilot-based STBC-OFDM systems.

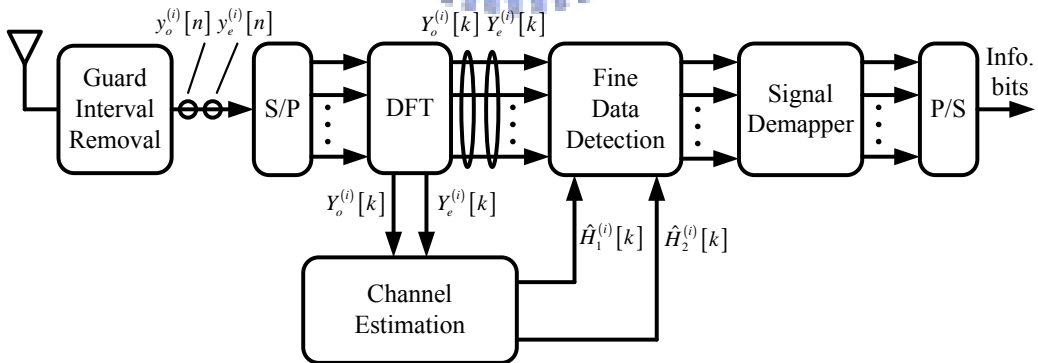
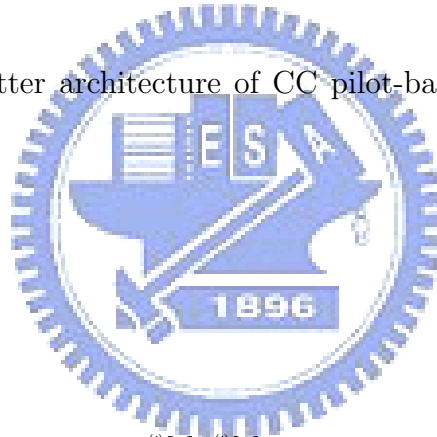


Figure 2.2: Receiver architecture of CC pilot-based STBC-OFDM systems.

and digitized. We assume that timing and carrier frequency synchronization are perfect and the length of the channel impulse response is not longer than the length of the GI. The channels are assumed to be quasi-static during any two successive OFDM symbols duration. The frequency response of the propagation channel between the first transmit antenna and the receive antenna is denoted by $H_1^{(i)} [k]$, and the other one is denoted by $H_2^{(i)} [k]$, where index i is used to indicate the corresponding $(2i)th$ and $(2i + 1)th$ OFDM symbol duration. Hence, after the GI removal, S/P conversion and N -point DFT computation, the successively received signals $Y_e^{(i)} [k]$ and $Y_o^{(i)} [k]$ on the kth subcarrier in the $(2i)th$ and $(2i + 1)th$ OFDM symbol duration can be represented as

$$Y_e^{(i)} [k] = H_1^{(i)} [k] \left(X_F^{(i)} [k] + P_\alpha [k] \right) + H_2^{(i)} [k] \left(X_S^{(i)} [k] - P_\beta [k] \right) + Z_e^{(i)} [k] \quad (2.8)$$

and

$$Y_o^{(i)} [k] = H_1^{(i)} [k] \left(-X_S^{(i)*} [k] + P_\beta^* [k] \right) + H_2^{(i)} [k] \left(X_F^{(i)*} [k] + P_\alpha^* [k] \right) + Z_o^{(i)} [k] \quad (2.9)$$

for $k = 0, \dots, N - 1$, where $Z_e^{(i)} [k]$ and $Z_o^{(i)} [k]$ are the AWGN in the $(2i)th$ and $(2i + 1)th$ OFDM symbol duration, respectively. The noise is modeled as an independent complex Gaussian random variable with zero-mean and variance σ_n^2 .

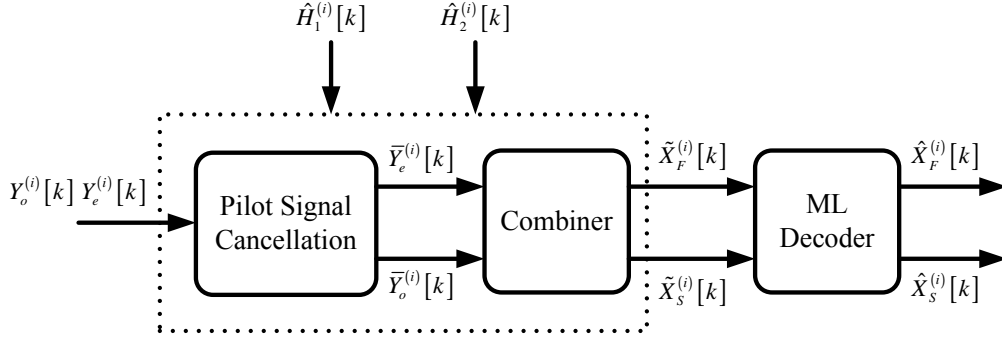


Figure 2.3: Fine data detection functional block.

2.4.1 Fine Data Detection

Figure 2.3 shows the details of the fine data detection functional block which include a pilot cancellation unit, a combiner unit and an ML decoder unit. We assume that the estimated CSIs are accurate, i.e. $\hat{H}_1^{(i)}[k] = H_1^{(i)}[k]$; $\hat{H}_2^{(i)}[k] = H_2^{(i)}[k]$, and the pilot interference signals can thus be reconstructed and be completely subtracted from the received signals $Y_e^{(i)}[k]$ and $Y_o^{(i)}[k]$. The output of the pilot signal cancellation unit can be expressed as

$$\bar{Y}_e^{(i)}[k] = H_1^{(i)}[k] X_F^{(i)}[k] + H_2^{(i)}[k] X_S^{(i)}[k] + Z_e^{(i)}[k] \quad (2.10)$$

and

$$\bar{Y}_o^{(i)}[k] = -H_1^{(i)}[k] X_S^{(i)*}[k] + H_2^{(i)}[k] X_F^{(i)*}[k] + Z_o^{(i)}[k] \quad (2.11)$$

for $k = 0, \dots, N - 1$. We then use the combiner unit to combine the received signals from different transmit antennas according to the method revealed in [14]. Finally, in the ML decoder unit, the ML decision rule can be used to detect the transmitted data symbol on each subchannel [12].

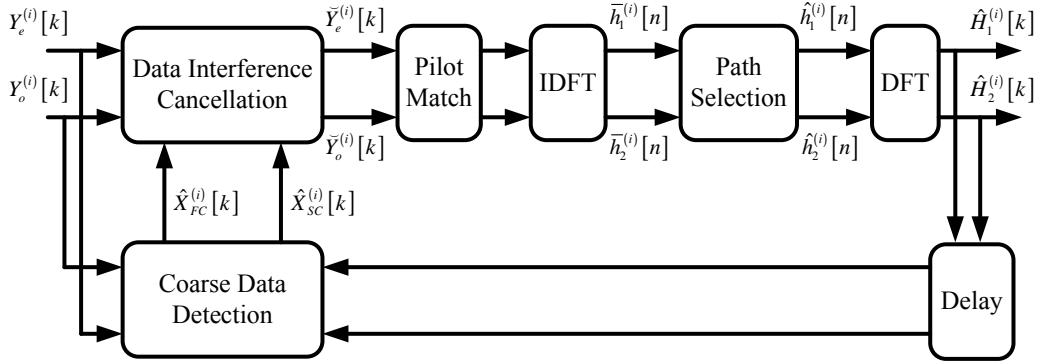


Figure 2.4: Channel estimation functional block.

2.4.2 Channel Estimation

Figure 2.4 depicts the detailed structure of the channel estimation functional block. First, we use a coarse data detection unit to obtain coarsely estimated data symbols $\hat{X}_{FC}^{(i)}[k]$ and $\hat{X}_{SC}^{(i)}[k]$ by using the estimated CSIs in the previous two OFDM symbols, i.e., the $(2i-2)$ th and $(2i-1)$ th OFDM symbol. The structure of the coarse data detection unit is similar to the fine data detection functional block as shown in Figure 2.3. The data interference signals are then reconstructed and subtracted from the received signals $Y_e^{(i)}[k]$ and $Y_o^{(i)}[k]$ in a data interference cancellation unit. We assume that the data interference signals can be cancelled perfectly and the refined signals can be represented as

$$\tilde{Y}_e^{(i)}[k] = H_1^{(i)}[k] P_\alpha[k] - H_2^{(i)}[k] P_\beta[k] + Z_e^{(i)}[k] \quad (2.12)$$

and

$$\tilde{Y}_o^{(i)}[k] = H_1^{(i)}[k] P_\beta^*[k] + H_2^{(i)}[k] P_\alpha^*[k] + Z_o^{(i)}[k] \quad (2.13)$$

for $k = 0, \dots, N - 1$. Next, the refined signals are multiplied by the complex conjugate transpose of the pilot matrix $\mathbf{P}[k]$ in the pilot matching unit in Figure 2.4, to obtain more accurately estimated CSIs in the frequency domain. We then use an N -point IDFT unit to obtain the estimated CSIs in the time domain, i.e.,

$$\bar{h}_1^{(i)}[n] = h_1^{(i)}[n] + IDFT \{ Z_e^{(i)}[k] P_\alpha^*[k] + Z_o^{(i)}[k] P_\beta[k] \} \quad (2.14)$$

and

$$\bar{h}_2^{(i)}[n] = h_2^{(i)}[n] + IDFT \{ -Z_e^{(i)}[k] P_\beta^*[k] + Z_o^{(i)}[k] P_\alpha[k] \} \quad (2.15)$$

for $n = 0, \dots, N - 1$. Finally, a path selection unit is used to suppress the noise effect and to refine the estimated CSIs. For path selection, we first define a parameter N_p , which is the desired number of paths to be selected. Only the N_p paths with larger amplitudes in $\bar{h}_1^{(i)}[n]$ (or $\bar{h}_2^{(i)}[n]$) are preserved and all the other paths are discarded. As a result, we obtain $\hat{h}_1^{(i)}[n]$ and $\hat{h}_2^{(i)}[n]$ in the following way:

$$\hat{h}_j^{(i)}[n] = \begin{cases} \bar{h}_j^{(i)}[n], & \text{if } |\bar{h}_j^{(i)}[n]| \text{ is one of the } N_p \text{ larger values} \\ 0, & \text{otherwise} \end{cases} \quad (2.16)$$

for $n = 0, \dots, N - 1$ and $j = 1, 2$.

To initialize the channel estimator, pilot preambles without data signals added are transmitted in the first two OFDM symbols. The received signals are passed only through the pilot matching and the path selection unit to generate the preliminary channel estimations.

2.4.3 Computational Complexity

In this subsection, the complexity of the CC pilot-based STBC-OFDM receiver will be calculated in terms of the number of complex multiplications. Either N -point DFT or IDFT needs $DFTN$ multiplications per OFDM symbol. At the receiver, three DFT or IDFT operations are needed for every OFDM symbol (see Figure 2.2 to Figure 2.4). According to Figure 2.3, the pilot signal cancellation unit and the combiner unit need $8N$ multiplications for every two OFDM symbols. In Figure 2.4, the data interference cancellation unit and the pilot matching unit also need $8N$ multiplications for every two OFDM symbols. Furthermore, the coarse data detection unit requires the same number of multiplications as the fine data detection functional block, i.e. $8N$ per two OFDM symbols. Hence, the receiver has the complexity order of $3DFTN + 12N$ multiplications per OFDM symbol.

2.5 Performance Analysis

In this section, we include an analysis of the BER performance of the proposed system in a two-path fading channel. The analysis method is general and can be easily extended to a mobile radio channel with more paths.

2.5.1 Time-Varying Effect of Two-Path Channels

The equivalent baseband impulse response of the two-path fading channel is represented by

$$h[n] = a_1\delta[n] + a_2\delta[n - \tau] \quad (2.17)$$

where $\delta[n]$ denotes a delta function, τ is the excess delay of the second path, and a_l is the complex gain of the l th path. Assume that the complex gain of the l th path, for $l = 1$ and 2 , during the $(2i)$ th and $(2i + 1)$ th OFDM symbol is a complex Gaussian random variable and denoted as $a_l^{(i)} = a_{l,I}^{(i)} + ja_{l,Q}^{(i)}$, where $a_{l,I}^{(i)}$ and $a_{l,Q}^{(i)}$ are the real and imaginary part of $a_l^{(i)}$, respectively. According to the Jakes fading channel model [33], the random variables $a_{l,I}^{(i-1)}$, $a_{l,Q}^{(i-1)}$, $a_{l,I}^{(i)}$ and $a_{l,Q}^{(i)}$ have the following correlations:

$$\mathbb{E} \left[a_{l,I}^{(i-1)} a_{l,I}^{(i)} \right] = \mathbb{E} \left[a_{l,Q}^{(i-1)} a_{l,Q}^{(i)} \right] = \frac{\varepsilon_l}{2} J_0 (2\pi f_D (2T_s)) \quad (2.18)$$

and

$$\mathbb{E} \left[a_{l,I}^{(i-1)} a_{l,Q}^{(i)} \right] = \mathbb{E} \left[a_{l,Q}^{(i-1)} a_{l,I}^{(i)} \right] = 0 \quad (2.19)$$

where $\mathbb{E}[\cdot]$ is the operation of taking expectation, $\varepsilon_l = \mathbb{E} \left[\left| a_l^{(i)} \right|^2 \right]$ is the power of the l th path, f_D is the maximum Doppler frequency, and $J_0(\cdot)$ is the Bessel function of the first kind. In order to model the time-varying effect of the l th path between $a_l^{(i-1)}$ and $a_l^{(i)}$, we have the following equation:

$$a_l^{(i)} = a_l^{(i-1)} + \Delta_l^{(i)} \quad (2.20)$$

where $\Delta_l^{(i)}$ is a complex Gaussian random variable with zero mean. Therefore, the normalized error power of the l th path becomes

$$\frac{\mathbb{E} \left[\left| \Delta_l^{(i)} \right|^2 \right]}{\varepsilon_l} = 2 (1 - J_0 (4\pi f_D T_s)) \quad (2.21)$$

Assume the corresponding channel frequency response is denoted by $H^{(i)}[k]$, and as a result, the channel variation in frequency domain, $\Omega^{(i)}[k]$, can be described by the following equation:

$$H^{(i)}[k] = H^{(i-1)}[k] + \Omega^{(i)}[k] \quad (2.22)$$

The two fading paths are assumed to be independent of each other, and it is obtained from (2.21) that $\Omega^{(i)} [k]$ is a zero-mean complex Gaussian random variable with variance

$$\sigma_{\Omega}^2 = \mathbb{E} \left[\left| \Omega^{(i)} [k] \right|^2 \right] = 2(1 - J_0(4\pi f_D T_s)) \sum_{l=1}^2 \varepsilon_l \quad (2.23)$$

Without loss of generality, we assume that the channel power is normalized to one, i.e. $\sum_{l=1}^2 \varepsilon_l = 1$; hence, we get $\sigma_{\Omega}^2 = 2(1 - J_0(4\pi f_D T_s))$.

2.5.2 Performance Analysis of Coarse Data Detection

Assume that binary phase-shift keying (BPSK) modulation is used, and data signals $X_F^{(i)} [k]$ and $X_S^{(i)} [k]$ in each subcarrier k are independently and identically distributed (i.i.d.) random variables with zero mean and variance $E_b/2$. For the simplicity of analysis, it is assumed that the CC sequences are ideal random binary sequences with zero mean, and the power of CC pilot signals is the same as data signals. Finally, we assume that the estimated CSI $\hat{H}_1^{(i-1)} [k]$ used by the coarse data detection unit can be modeled as $H_1^{(i)} [k] - \Lambda_1^{(i)} [k]$, where the term $\Lambda_1^{(i)} [k]$ includes both the effects of channel variation and channel estimation error, and its mean and variance are given by zero and σ_{Λ}^2 . The estimated CSI $\hat{H}_2^{(i-1)} [k]$ is also assumed in a like manner. For simplicity, the indices i and k are omitted in the following derivation.

From (2.8) and (2.9), the outputs of the pilot signal cancellation unit, in the coarse data detection unit in Figure 2.4, are given by

$$\bar{Y}_{eC} = \underbrace{H_1 X_F + H_2 X_S}_{D_e} + \underbrace{P_{\alpha} \Lambda_1 - P_{\beta} \Lambda_2}_{\tilde{I}_e} + Z_e \quad (2.24)$$

and

$$\bar{Y}_{oC} = \underbrace{-H_1 X_S^* + H_2 X_F^*}_{D_o} + \underbrace{P_\beta^* \Lambda_1 + P_\alpha^* \Lambda_2}_{\tilde{I}_o} + Z_o \quad (2.25)$$

where \bar{Y}_{eC} and \bar{Y}_{oC} consist of three mutually independent components:

- desired data signal components D_e and D_o ;
- residual pilot interference signal components \tilde{I}_e and \tilde{I}_o ;
- AWGN components Z_e and Z_o .

Since the random variables P_α , P_β , Λ_1 , and Λ_2 are mutually independent of each other, the mean and variance of \tilde{I}_e and \tilde{I}_o are given by

$$\mathbb{E} [\tilde{I}_e] = \mathbb{E} [\tilde{I}_o] = 0 \quad (2.26)$$

and

$$\text{Var} [\tilde{I}_e] = \text{Var} [\tilde{I}_o] = E_b \sigma_\Lambda^2 \quad (2.27)$$

where $\text{Var} [\cdot]$ is the operation of taking variance. Given the channel gains H_1 and H_2 , and the transmitted data signals X_F and X_S , we have the mean

$$\mathbb{E} [\bar{Y}_{eC}] = H_1 X_F + H_2 X_S \quad (2.28)$$

$$\mathbb{E} [\bar{Y}_{oC}^*] = -H_1^* X_S + H_2^* X_F \quad (2.29)$$

and the variance

$$\text{Var} [\bar{Y}_{eC}] = \text{Var} [\bar{Y}_{oC}^*] = E_b \sigma_\Lambda^2 + \sigma_n^2 \quad (2.30)$$

The outputs of the combiner unit in the coarse data detection unit are [14]

$$\tilde{X}_{FC} = \hat{H}_1^* \bar{Y}_{eC} + \hat{H}_2 \bar{Y}_{oC}^* \quad (2.31)$$

and

$$\tilde{X}_{SC} = \hat{H}_2^* \bar{Y}_{eC} - \hat{H}_1 \bar{Y}_{oC}^* \quad (2.32)$$

Since \hat{H}_1^* , \bar{Y}_{eC} , \hat{H}_2 and \bar{Y}_{oC}^* are mutually uncorrelated of each other, the mean and variance of \tilde{X}_{FC} are given by

$$\mathbb{E} [\tilde{X}_{FC}] = \mathbb{E} [\hat{H}_1^*] \mathbb{E} [\bar{Y}_{eC}] + \mathbb{E} [\hat{H}_2] \mathbb{E} [\bar{Y}_{oC}^*] = \zeta X_F \quad (2.33)$$

and

$$\begin{aligned} \text{Var} [\tilde{X}_{FC}] &= \mathbb{E}^2 [\hat{H}_1^*] \text{Var} [\bar{Y}_{eC}] + \mathbb{E}^2 [\bar{Y}_{eC}] \text{Var} [\hat{H}_1^*] \\ &\quad + \mathbb{E}^2 [\hat{H}_2] \text{Var} [\bar{Y}_{oC}^*] + \mathbb{E}^2 [\bar{Y}_{oC}^*] \text{Var} [\hat{H}_2] \\ &\quad + \text{Var} [\hat{H}_1^*] \text{Var} [\bar{Y}_{eC}] + \text{Var} [\hat{H}_2] \text{Var} [\bar{Y}_{oC}^*] \\ &= \zeta (2E_b \sigma_\Lambda^2 + \sigma_n^2) + 2E_b \sigma_\Lambda^4 + 2\sigma_\Lambda^2 \sigma_n^2 \end{aligned} \quad (2.34)$$

where $\zeta = |H_1|^2 + |H_2|^2$. Similarly, the other combiner output \tilde{X}_{SC} has the same variance as \tilde{X}_{FC} , but different mean ζX_S . Therefore, conditioned on the combined channel gain ζ , the BER is given by

$$\text{BER}_C(\zeta) = Q\left(\frac{\zeta}{\sqrt{a\zeta + b}}\right) \quad (2.35)$$

where $Q(x) = \int_x^\infty (1/\sqrt{2\pi}) e^{-y^2/2} dy$, $a = 2\sigma_\Lambda^2 + \sigma_n^2/E_b$, $b = 2\sigma_\Lambda^4 + 2\sigma_\Lambda^2 \sigma_n^2/E_b$.

2.5.3 Performance Analysis of Channel Estimation and Fine Data Detection

As shown in Figure 2.4, in order to reconstruct the data interference signals, tentative decisions of data symbols are made and denoted as \hat{X}_{FC} and \hat{X}_{SC} .

From (2.8) and (2.9), notice that after the data interference cancellation, the output signal \check{Y}_e is given by

$$\begin{aligned} \check{Y}_e &= \underbrace{H_1 P_\alpha - H_2 P_\beta}_{I_e} \\ &\quad + \underbrace{H_1 (X_F - \hat{X}_{FC}) + \Lambda_1 \hat{X}_{FC} + H_2 (X_S - \hat{X}_{SC}) + \Lambda_2 \hat{X}_{SC}}_{\check{D}_e} \\ &\quad + Z_e \end{aligned} \quad (2.36)$$

and consists of three parts. The first part is the desired pilot signal I_e . The second part is the residual data interference signal \check{D}_e , which is caused by two factors. One is the channel estimation error, and the other is the tentative data decision error from the coarse data detection unit. The third part is AWGN Z_e . From (2.35) and (2.36), the mean and the second moment of \check{D}_e can be calculated as (see Appendix B)

$$\mathbb{E} [\check{D}_e] = 2BER_C(\zeta) (H_1 X_F + H_2 X_S) \quad (2.37)$$

and

$$\begin{aligned} \mathbb{E} \left[\left| \check{D}_e \right|^2 \right] &= 2BER_C(\zeta) \zeta E_b + \sigma_\Lambda^2 E_b \\ &\quad + 4BER_C^2(\zeta) (H_1 H_2^* X_F X_S^* + H_1^* H_2 X_F^* X_S) \end{aligned} \quad (2.38)$$

Similarly, for the other output \check{Y}_o , we have the residual data interference signal \check{D}_o as follows

$$\mathbb{E} [\check{D}_o] = 2BER_C(\zeta) (-H_1 X_S^* + H_2 X_F^*) \quad (2.39)$$

and

$$\begin{aligned} \mathbb{E} \left[\left| \check{D}_o \right|^2 \right] &= 2BER_C(\zeta) \zeta E_b + \sigma_\Lambda^2 E_b \\ &\quad - 4BER_C^2(\zeta) (H_1 H_2^* X_F X_S^* + H_1^* H_2 X_F^* X_S) \end{aligned} \quad (2.40)$$

According to (2.3) and (2.36), the estimated CSIs at the output of the pilot matching unit can be written as

$$\bar{H}_1 = H_1 + \frac{1}{E_b} \left(P_\alpha^* \tilde{D}_e + P_\alpha^* Z_e + P_\beta \tilde{D}_o + P_\beta Z_o \right) \quad (2.41)$$

and

$$\bar{H}_2 = H_2 + \frac{1}{E_b} \left(-P_\beta^* \tilde{D}_e - P_\beta^* Z_e + P_\alpha \tilde{D}_o + P_\alpha Z_o \right) \quad (2.42)$$

As can be seen in (2.41) and (2.42), the estimated CSIs \bar{H}_1 and \bar{H}_2 have different means H_1 and H_2 , respectively, but equal variance $2BER_C(\zeta)\zeta + \sigma_\Lambda^2 + \sigma_n^2/E_b$. In a two-path fading channel, the variance of \bar{H}_1 and \bar{H}_2 can be further reduced by a factor of $2/N$ if we assume the path selection process is perfect, i.e., we have

$$\bar{\sigma}_\Lambda^2 = \text{Var} [\bar{H}_1] = \text{Var} [\bar{H}_2] = \frac{2}{N} \left(2BER_C(\zeta)\zeta + \sigma_\Lambda^2 + \frac{\sigma_n^2}{E_b} \right) \quad (2.43)$$

We can then replace σ_Λ^2 with $\bar{\sigma}_\Lambda^2$ in (2.35) to obtain the BER, denoted as $BER(\zeta)$, for the fine data detection unit. With the two-path Rayleigh fading channel model, ζ is a random variable with probability density function $p(\zeta) = \zeta e^{-\zeta}$, where $\zeta \geq 0$. The averaged BER for the fine data detection unit can be obtained by averaging $BER(\zeta)$ over ζ , i.e.

$$BER = \int_0^\infty BER(\zeta) p(\zeta) d\zeta \quad (2.44)$$

In general, BER can be computed iteratively using software like MATLAB with the following procedure:

1. Initially, we set the variance of channel estimation error in the pilot preambles as $\sigma_U^2 = \sigma_{init}^2 = 2\sigma_n^2/NE_b$

2. Due to the Doppler effect, the variance of channel estimation error for the next two OFDM symbols is given by $\sigma_{\Lambda}^2 = \sigma_U^2 + \sigma_{\Omega}^2$
3. Using (2.35), (2.43) and (2.44), we can calculate $BER_C(\zeta)$, $\bar{\sigma}_{\Lambda}^2$, $BER(\zeta)$, and BER , respectively.

For simple analysis of BER performance, we average $\bar{\sigma}_{\Lambda}^2$ over ζ and use the averaged $\bar{\sigma}_{\Lambda}^2$ instead of σ_U^2 for the next iteration (return to the procedure 2). With 100 iterations, i.e., 200 OFDM symbols are processed, Figure 2.5 shows that BER almost stayed at the same value.

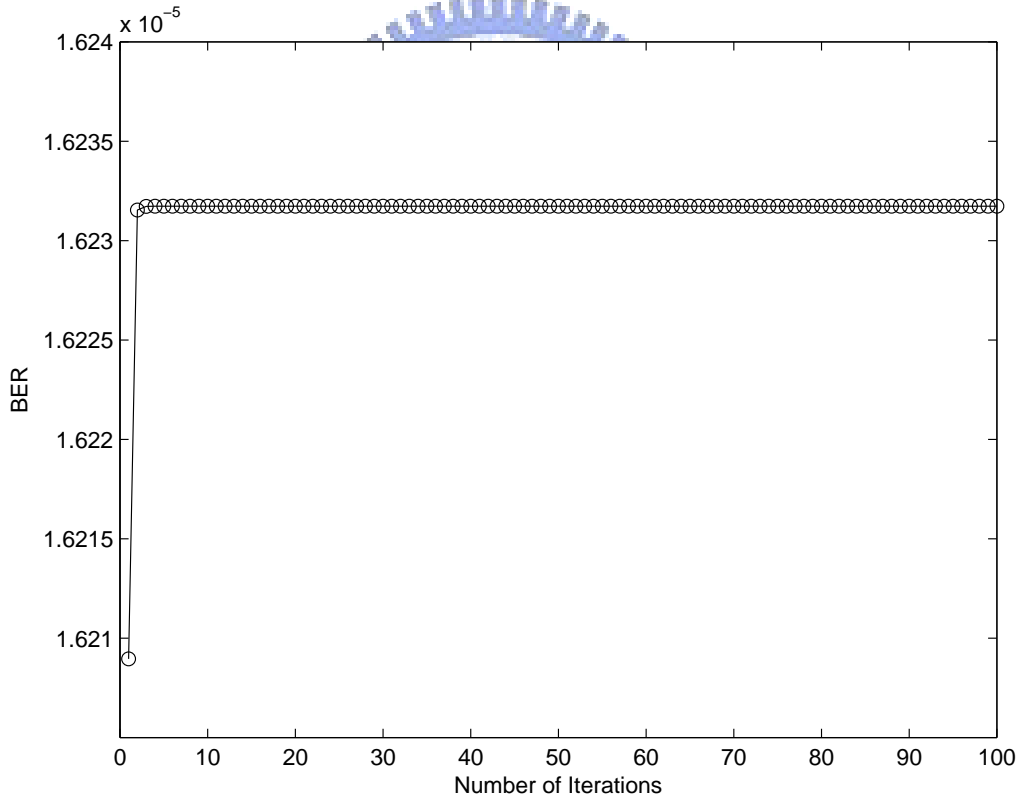


Figure 2.5: BER performance versus number of iterations at $E_b/\sigma_n^2=24dB$ and $f_D T_s=0.0111$.

2.6 Computer Simulation

We utilize computer simulations to verify the performance of the proposed CC pilot-based STBC-OFDM system in a two-path fading channel and a Universal Mobile Telecommunications System (UMTS) defined channel. The complex gain of each path is independently generated from the Jakes fading channel model [33]. The relative path power profiles of the two-path channel is 0, 0 (*dB*). Besides, the channel selected for evaluating the third generation UMTS European systems with relative path power profiles: -2.5, 0, -12.8, -10, -25.2, -16 (*dB*), is also used to simulate the system performance [34]. We also assume the two transmit antennas are spatially separated far enough and the two channels from the two transmitters to the receiver are uncorrelated in our simulation.

The system parameters for the CC pilot-based STBC-OFDM system simulation are listed in Table 2.1. In our simulation, a single-user at a time scenario, i.e. time division multiple access (TDMA) can be used as a multiple access scheme, is assumed. The entire simulations are conducted in the equivalent baseband. We assume both symbol synchronization and carrier synchronization are perfect. Golay's binary CC [31] is directly used to generate the CC pilot signals. Each antenna transmits both the data signal and the pilot signal at an equal power level of 0.5 per sample. The excess delay of the paths is uniformly distributed between $0.19\mu s$ and $9.77\mu s$. Finally, throughout the simulation, the parameter E_b/N_o is defined as the received bit power to the noise power ratio.

Table 2.1: Simulation parameters.

Parameter	Value
Carrier frequency	2 GHz
Bandwidth	5.12 MHz
FFT size	256
Length of CP	64
Modulation	BPSK
Number of receive antennas	1
Transmitted data signal energy/ antenna	0.5
Pilot signal energy/ antenna	0.5
Channel power profiles	ITU Veh-B channel
Channel delay profiles	0.19 μs ~ 9.77 μs
Vehicle speed	8, 30, 120 km/hr

2.6.1 Effect of Vehicle Speed

Figure 2.6 shows the BER performance of the proposed system in the two-path fading channel with a relative path power profile: 0, 0 (dB) at different vehicle speeds (V). Here, we set the number of selected paths $N_p = 2$. The system performance is approximately only $1dB$ poorer (in E_b/N_o) at a BER= 10^{-3} as compared with the theoretic BPSK case with a second-order diversity, even when the motor vehicle moves at a high speed. The simulated performance curves under the perfect CSIs assumption are also included for comparison purpose. These curves indicate that CC pilot-based channel estimation method works very well.

2.6.2 Effect of Path Selection

Figure 2.7 shows the BER performance of the proposed system in the UMTS defined multipath fading channel (with six paths) at a vehicle speed of 120 km/hr , with N_p is chosen as a simulation parameter. It is observed that the system performance have a small degradation of about $2dB$ at a BER= 10^{-3} as compared with the previous two-path fading channel case as shown in Figure 2.6. When the parameter N_p is chosen appropriately, i.e. close to the actual number of paths in the mobile radio environment, the system performance is almost the same. In other words, the parameter N_p can be set a little bit larger than the number of available paths in a practical system design.

2.6.3 Analytic and Simulated BER Performance

Figure 2.8 shows the analytic and simulated BER performance of the proposed system in the two-path fading channel at different vehicle speeds. Here, we set the number of selected paths $N_p = 2$. The results of our analysis and simulation both show that the fine data detection unit reduces the BER further as compared with the coarse data detection unit at a vehicle speed of 120 *km/hr*. Our analytic result shows that the BER performance is very close to the theoretic BPSK case with a second-order diversity, even when the motor vehicle moves at 120 *km/hr*. In our analysis, we assume that CC sequences are ideal binary random sequences with zero mean. Nevertheless, these sequences have non-zero mean, and it leads to a biased channel estimation. Furthermore, for simple analysis, we average out the error propagation effect in calculating the analytic BER performance. As a result, there is an observable discrepancy between the results of analysis and simulation only in the case of high vehicle speeds and high E_b/N_o values. As we can see, these curves indicate that the CC pilot-based channel estimation method works very effectively.

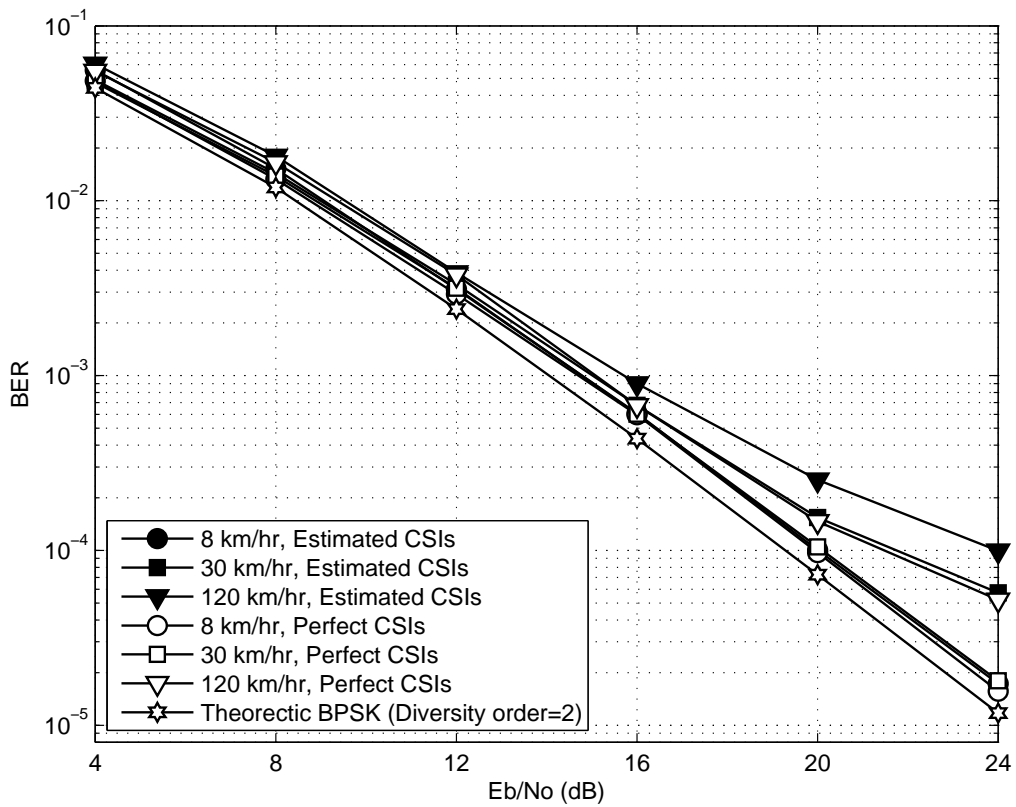


Figure 2.6: BER performance of CC pilot-based STBC-OFDM systems in the two-path fading channel with vehicular speed as a parameter ($N_p = 2$).

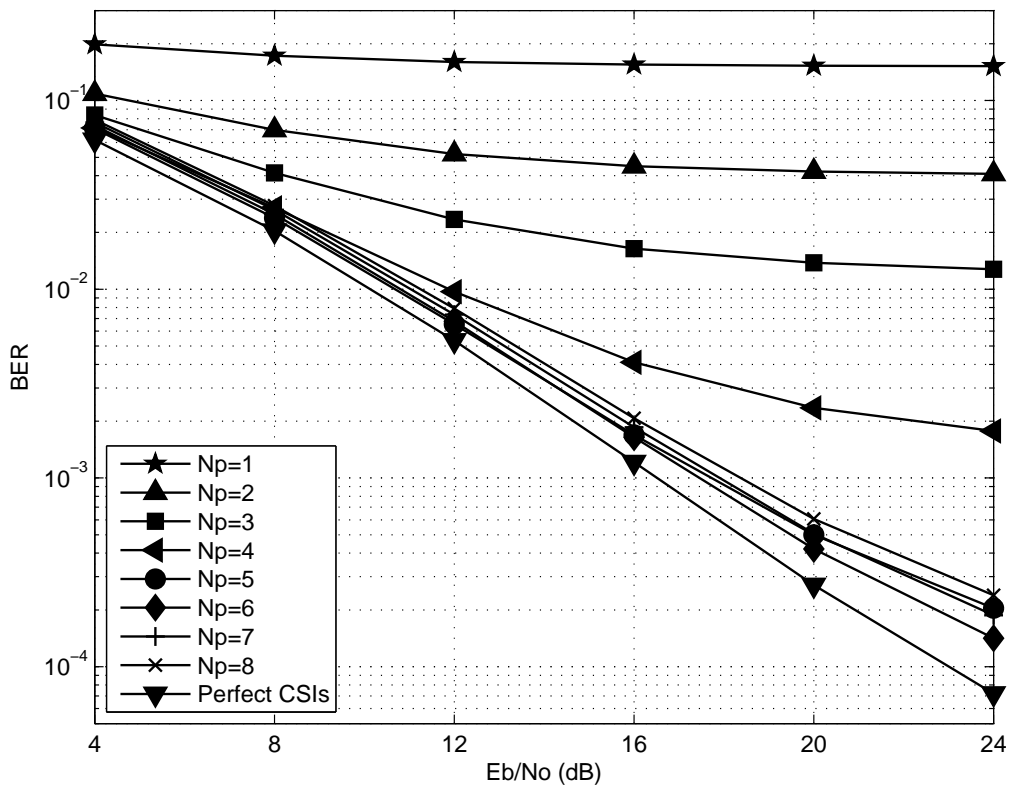


Figure 2.7: BER performance of CC pilot-based STBC-OFDM systems in the UMTS system defined fading channel with N_p as a parameter ($V = 120km/hr$).

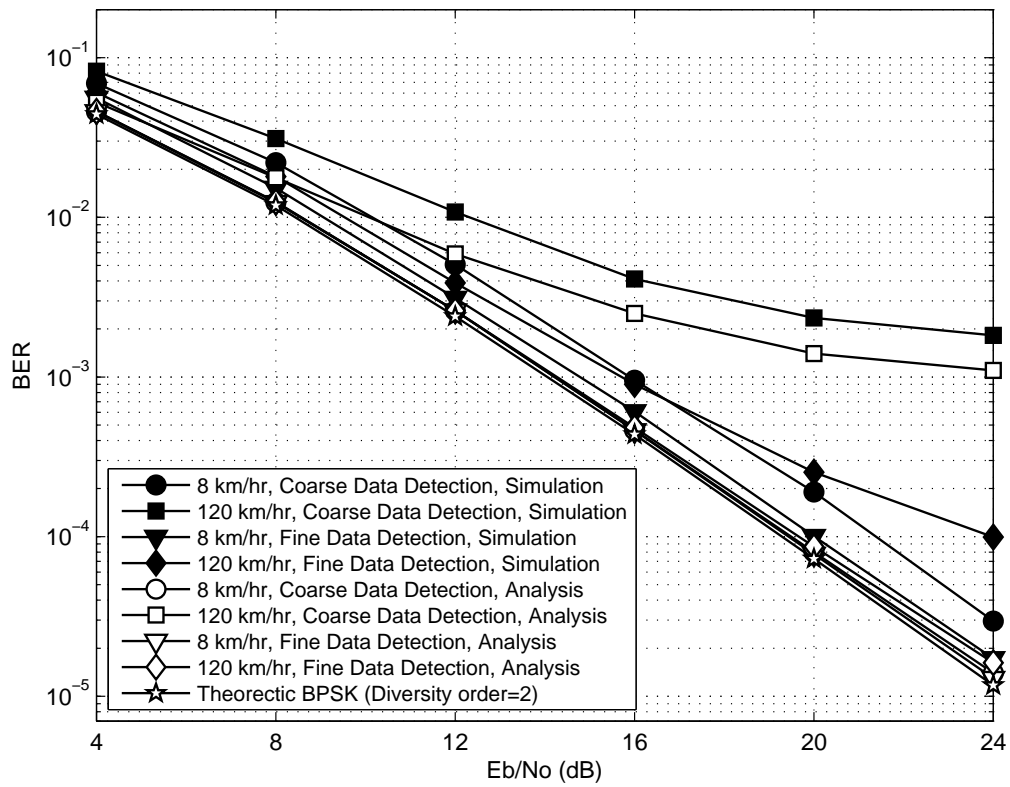
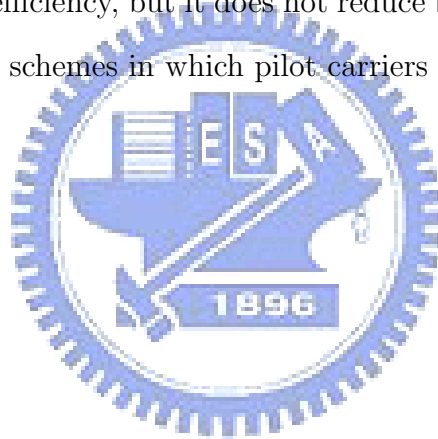


Figure 2.8: The analytic and simulated BER performance of CC pilot-based STBC-OFDM systems in the two-path fading channel with vehicular speed as a parameter.

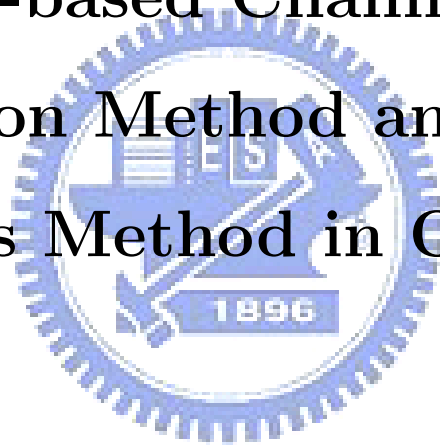
2.7 Summary

In this chapter, we designed a pilot signal structure for an STBC-OFDM system by utilizing a pair of complementary codes, which can be systematically generated from well-understood rules. We have also described the detailed functional block diagrams of the CC pilot-based STBC-OFDM system and simulated its performance. The CC pilot signals are transmitted simultaneously along with data signals in the time domain and used to estimate CSIs optimally at the receiver end. Although this approach has a drawback of lower power efficiency, but it does not reduce bandwidth efficiency as compared with other schemes in which pilot carriers are added in frequency domain.



Chapter 3

On the Equivalence between DF DFT-based Channel Estimation Method and Newton's Method in OFDM Systems



3.1 Literature Survey and Motivation

The DFT-based channel estimation method derived from the ML criterion is originally proposed for OFDM systems with pilot preambles [35–39]. In order to save bandwidth and improve system performance, DF data symbols are usually exploited to track channel variations in subsequent OFDM data symbols, and this method is called DF DFT-based channel estimation [35–37]. However, the working principle of this empirical method has

not been explored from the viewpoint of Newton's method in previous studies. This chapter derives the DF DFT-based channel estimation via Newton's method for STBC-OFDM systems. In this way, the equivalence between the two methods is established. Our results indicate that both methods can be implemented through the same four components: a least-square (LS) estimator, an IDFT matrix, a weighting matrix, and a DFT matrix, but with different connections. On one hand, the gradient vector in Newton's method can be found by calculating the difference between an estimated channel frequency response and an LS estimate, followed by the IDFT operation. On the other hand, the inverse of the Hessian matrix in Newton's method is just the weighting matrix operation in the DF DFT-based method.

The rest of this chapter is organized as follows. In Section 3.2, we briefly describe an STBC-OFDM system. In Section 3.3, a classical DF DFT-based channel estimation method is introduced, and a channel estimation method using the ML criterion is derived from Newton's method. The equivalence between the DF DFT-based method and Newton's method is then discussed in this section. In Section 3.4, we show our computer simulation results. Finally, some concluding remarks are drawn in Section 3.5.

3.2 STBC-OFDM Systems

Consider an STBC-OFDM system in Figure 3.1 with N_T transmit and N_R receive antennas, employing K subcarriers among which M subcarriers are used to transmit data symbols and the other $K - M$ subcarriers are used as either a DC subcarrier or virtual subcarriers. Assume that the set of

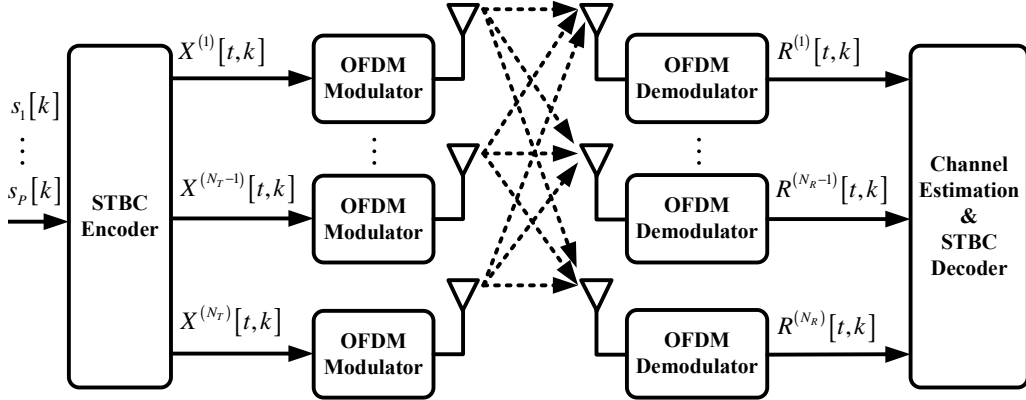


Figure 3.1: STBC-OFDM systems.

data subcarrier indices is denoted as $\mathbf{Q} \subseteq \{1, \dots, K\}$. At subcarrier $k \in \mathbf{Q}$ and after symbol mapping, P modulated data symbols $\{s_1[k], \dots, s_P[k]\}$ are encoded by an $N_T \times N_L$ STBC encoder $\mathbf{X}[k]$ to generate N_T signal sequences of length N_L , denoted by $\{X^{(i)}[1, k], \dots, X^{(i)}[N_L, k]\}$, for $i = 1, \dots, N_T$ [12, 13, 40]. As a simple example, for a 2×2 Alamouti's STBC, we have $X^{(1)}[1, k] = s_1[k]$, $X^{(2)}[1, k] = s_2[k]$, $X^{(1)}[2, k] = -s_2^*[k]$, and $X^{(2)}[2, k] = s_1^*[k]$. It is noted that these signal sequences possess the orthogonal property, given by $\mathbf{X}^*[k] \mathbf{X}^T[k] = C[k] \mathbf{I}_{N_T}$, where $(\cdot)^*$ and $(\cdot)^T$ represent complex conjugate and transpose, respectively, \mathbf{I}_N is an $N \times N$ identity matrix, and $C[k] = \sum_{t=1}^{N_L} |X^{(i)}[t, k]|^2$. After insertion of $K - M$ zeros for DC and virtual subcarriers, the STBC encoded data symbols $X^{(i)}[t, k]$ are modulated onto M subcarriers via a K -point IDFT unit to produce time domain samples, for $t = 1, \dots, N_L$ and $i = 1, \dots, N_T$. The time domain samples are then appended with CP of length G and transmitted through N_T transmit antennas within the duration of N_L OFDM data symbols.

We assume that both timing and carrier frequency synchronization are

perfect, and that the length of channel impulse response is always smaller than the length of the CP. Another assumption here is that the channel is quasi-static over the duration of a time slot, including N_L OFDM data symbols, but it varies from one time slot to another. Hence, at the output of the OFDM demodulator in Figure 3.1, the N_L successively received OFDM data symbols at the j th receive antenna are given by

$$R^{(j)} [t, k] = \sum_{i=1}^{N_T} H^{(j,i)} [k] X^{(i)} [t, k] + Z^{(j)} [t, k] \quad (3.1)$$

for $t = 1, \dots, N_L$ and $k \in \mathbf{Q}$, where $H^{(j,i)} [k]$ is the channel frequency response for the (j, i) th antenna pair, and $Z^{(j)} [t, k]$ is uncorrelated additive white Gaussian noise (AWGN) on the j th receive antenna with zero-mean and variance σ_Z^2 .

3.3 DF DFT-Based Method and Newton's Method

3.3.1 DF DFT-Based Channel Estimation Method

As shown in Figure 3.2, the block diagram of the DF DFT-based channel estimation method is composed of an LS estimator, an IDFT matrix, a weighting matrix, and a DFT matrix [36–39]. The LS estimator exploits DF data symbols to produce an LS estimate, which is a noisy estimation of channel frequency response. After taking the IDFT to transform the estimate to time domain, we can improve this estimate by using a weighting matrix which depends on the performance criterion chosen, either ML or MMSE [37, 39].

Finally, the enhanced estimate is transformed back to frequency domain to obtain a new estimate of channel frequency response.

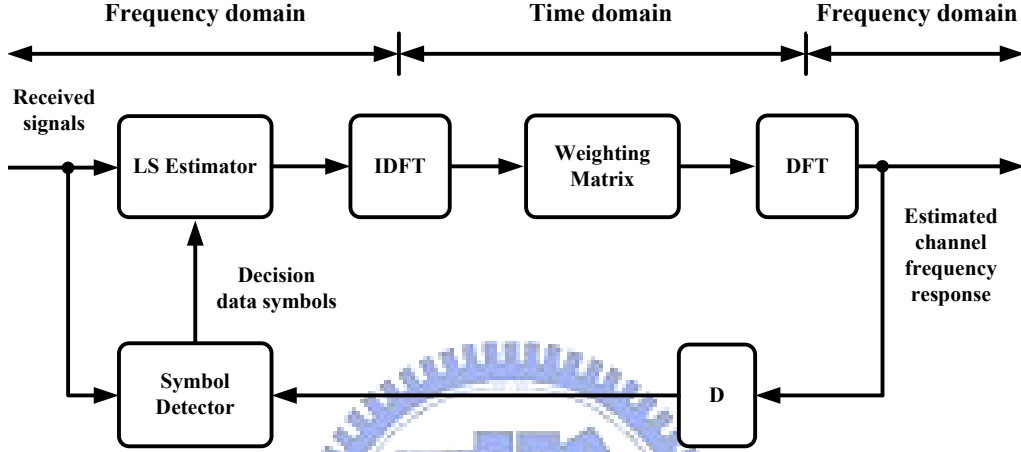


Figure 3.2: The block diagram of the DF DFT-based channel estimation method. (D is a delay component.)

3.3.2 Channel Estimation via Newton's Method

A parametric channel model $M^{(j,i)}[k]$ of the channel frequency response $H^{(j,i)}[k]$ is first formed by a summation of G complex sinusoids as follows:

$$M^{(j,i)}[k] = \sum_{l=1}^G \mu_l^{(j,i)} e^{-j \frac{2\pi(k-1)(l-1)}{K}} \quad (3.2)$$

where $\mu_l^{(j,i)} = \alpha_l^{(j,i)} + j\beta_l^{(j,i)}$ is a complex fading gain to be tracked in subsequent time slots. From (3.1) and (3.2), the joint channel estimation and data detection problem can be formulated in an ML estimation framework

as follows:

$$(\hat{\mathbf{s}}, \hat{\mathbf{y}}) = \arg \min_{\mathbf{s}, \mathbf{y}} \sum_{j=1}^{N_R} \sum_{t=1}^{N_L} \sum_{k \in \Theta} \left| R^{(j)} [t, k] - \sum_{i=1}^{N_T} M^{(j,i)} [k] X^{(i)} [t, k] \right|^2 \quad (3.3)$$

where $\Theta = \{\Theta_1, \dots, \Theta_{N_S}\}$ is a subset of \mathbf{Q} over which we execute the summation, \mathbf{s} denotes the data symbols which are STBC encoded and transmitted over subcarriers Θ , and N_S denotes the cardinality of Θ . In addition, we define $\boldsymbol{\mu}_I^{(j,i)} = [\alpha_1^{(j,i)}, \dots, \alpha_G^{(j,i)}]^T$, $\boldsymbol{\mu}_Q^{(j,i)} = [\beta_1^{(j,i)}, \dots, \beta_G^{(j,i)}]^T$, $\mathbf{y}^{(j,i)} = [\boldsymbol{\mu}_I^{(j,i)T}, \boldsymbol{\mu}_Q^{(j,i)T}]^T$, $\mathbf{y}^{(j)} = [\mathbf{y}^{(j,1)T}, \dots, \mathbf{y}^{(j,N_T)T}]^T$, and $\mathbf{y} = [\mathbf{y}^{(1)T}, \dots, \mathbf{y}^{(N_R)T}]^T$. Because it is hard to solve (3.3) directly, we yield a simplified optimization problem by relaxing (3.3) as follows:

$$(\hat{\mathbf{s}}, \hat{\mathbf{y}}) = \arg \min_{\mathbf{y}} \min_{\mathbf{s}} \sum_{j=1}^{N_R} \sum_{t=1}^{N_L} \sum_{k \in \Theta} \left| R^{(j)} [t, k] - \sum_{i=1}^{N_T} M^{(j,i)} [k] X^{(i)} [t, k] \right|^2 \quad (3.4)$$

Assuming that $M^{(j,i)} [k]$ is known, it is straightforward to solve the minimization problem with respect to \mathbf{s} first by applying the STBC decoding algorithm [12, 40], and we have

$$\begin{aligned} \hat{\mathbf{y}} &= \arg \min_{\mathbf{y}} \sum_{j=1}^{N_R} \sum_{t=1}^{N_L} \sum_{k \in \Theta} \left| R^{(j)} [t, k] - \sum_{i=1}^{N_T} M^{(j,i)} [k] \hat{X}^{(i)} [t, k] \right|^2 \\ &\triangleq \arg \min_{\mathbf{y}} \sum_{j=1}^{N_R} \sum_{t=1}^{N_L} \sum_{k \in \Theta} \left| \Psi^{(j)} [t, k] \right|^2 \end{aligned} \quad (3.5)$$

where $\hat{X}^{(i)} [t, k]$ is the signal (corresponding to $X^{(i)} [t, k]$) obtained by re-encoding the decision symbols $\hat{s}_p [k] = \Phi (\tilde{s}_p [k])$, $\Phi (\cdot)$ is a symbol decision

function, and $\tilde{s}_p[k]$ is the signal after diversity combining [12, 40]. Notice that (3.5) might converge to local minima, leading to BER performance loss, as compared with (3.3), particularly when the initial choice of $M^{(j,i)}[k]$ is not accurate enough. By rewriting (3.5), we have

$$\begin{aligned}\hat{\mathbf{y}} &= \arg \min_{\mathbf{y}} \sum_{j=1}^{N_R} \sum_{t=1}^{N_L} \sum_{k \in \Theta} \Psi_I^{(j)^2}[t, k] + \Psi_Q^{(j)^2}[t, k] \\ &\triangleq \arg \min_{\mathbf{y}} D(\mathbf{y})\end{aligned}\quad (3.6)$$

where notations $\Upsilon_I(\cdot)$ and $\Upsilon_Q(\cdot)$ denote the real and imaginary part of the notation $\Upsilon(\cdot)$, respectively. For simplification, we drop the variable notation " \mathbf{y} " in $D(\mathbf{y})$ hereafter except otherwise stated. Now we use Newton's method to find the minimum of (3.6), and the well-known iterative formula of Newton's method is provided in the following [41]:

$$\hat{\mathbf{y}}_v = \hat{\mathbf{y}}_{v-1} - \mathbf{g}_v \quad (3.7)$$

where v is the iteration index and $v = 1, \dots, V$, $\hat{\mathbf{y}}_v$ is the estimated CSI obtained at the v th iteration, \mathbf{g}_v is a search vector associated with $\mathbf{g} = \mathbf{E}^{-1}\mathbf{q}$ at $\mathbf{y} = \hat{\mathbf{y}}_{v-1}$ in which \mathbf{E} and \mathbf{q} are the Hessian matrix and the gradient vector of D , respectively, and $(\cdot)^{-1}$ represents the matrix inverse. Thus, the u th entry of \mathbf{q} is calculated as

$$\begin{aligned}(\mathbf{q})_u &\triangleq \frac{\partial D}{\partial (\mathbf{y})_u} \\ &= 2 \sum_{j=1}^{N_R} \sum_{t=1}^{N_L} \sum_{k \in \Theta} \Psi_I^{(j)}[t, k] \frac{\partial \Psi_I^{(j)}[t, k]}{\partial (\mathbf{y})_u} \\ &\quad + \Psi_Q^{(j)}[t, k] \frac{\partial \Psi_Q^{(j)}[t, k]}{\partial (\mathbf{y})_u}\end{aligned}\quad (3.8)$$

where $(\mathbf{y})_u$ is the u th entry of \mathbf{y} . The partial derivative of $\partial\Psi_I^{(j)}[t, k]/\partial(\mathbf{y})_u$ and $\partial\Psi_Q^{(j)}[t, k]/\partial(\mathbf{y})_u$ can be derived in the following way. First, we assume that the probabilities of $\tilde{s}_{p,I}[k] = 0$ or $\tilde{s}_{p,Q}[k] = 0$ are zero; thus, it is reasonable to take the terms involving the partial derivative of the function $\Phi(\cdot)$ as zero. Since the variable $(\mathbf{y})_u$ in \mathbf{y} is either $\alpha_l^{(j,i)}$ or $\beta_l^{(j,i)}$, straightforward calculation using (3.6) shows that for $j = j'$, we have

$$\begin{aligned}\frac{\partial\Psi_I^{(j')}[t, k]}{\partial\alpha_l^{(j,i)}} &= -\cos\left(\frac{2\pi(k-1)(l-1)}{K}\right)\hat{X}_I^{(i)}[t, k] \\ &\quad -\sin\left(\frac{2\pi(k-1)(l-1)}{K}\right)\hat{X}_Q^{(i)}[t, k]\end{aligned}\quad (3.9)$$

$$\begin{aligned}\frac{\partial\Psi_Q^{(j')}[t, k]}{\partial\alpha_l^{(j,i)}} &= -\cos\left(\frac{2\pi(k-1)(l-1)}{K}\right)\hat{X}_Q^{(i)}[t, k] \\ &\quad +\sin\left(\frac{2\pi(k-1)(l-1)}{K}\right)\hat{X}_I^{(i)}[t, k]\end{aligned}\quad (3.10)$$

$$\begin{aligned}\frac{\partial\Psi_I^{(j')}[t, k]}{\partial\beta_l^{(j,i)}} &= -\sin\left(\frac{2\pi(k-1)(l-1)}{K}\right)\hat{X}_I^{(i)}[t, k] \\ &\quad +\cos\left(\frac{2\pi(k-1)(l-1)}{K}\right)\hat{X}_Q^{(i)}[t, k]\end{aligned}\quad (3.11)$$

$$\begin{aligned}\frac{\partial\Psi_Q^{(j')}[t, k]}{\partial\beta_l^{(j,i)}} &= -\sin\left(\frac{2\pi(k-1)(l-1)}{K}\right)\hat{X}_Q^{(i)}[t, k] \\ &\quad -\cos\left(\frac{2\pi(k-1)(l-1)}{K}\right)\hat{X}_I^{(i)}[t, k]\end{aligned}\quad (3.12)$$

Otherwise, i.e. if $j \neq j'$, we have

$$\begin{aligned}\frac{\partial\Psi_I^{(j')}[t, k]}{\partial\alpha_l^{(j,i)}} &= \frac{\partial\Psi_Q^{(j')}[t, k]}{\partial\alpha_l^{(j,i)}} \\ &= \frac{\partial\Psi_I^{(j')}[t, k]}{\partial\beta_l^{(j,i)}} = \frac{\partial\Psi_Q^{(j')}[t, k]}{\partial\beta_l^{(j,i)}} = 0\end{aligned}\quad (3.13)$$

Next, we compute the (m, u) th entry of \mathbf{E} as

$$\begin{aligned}
(\mathbf{E})_{m,u} &\triangleq \frac{\partial^2 D}{\partial(\mathbf{y})_m \partial(\mathbf{y})_u} \\
&= 2 \sum_{j=1}^{N_R} \sum_{t=1}^{N_L} \sum_{k \in \Theta} \frac{\partial \Psi_I^{(j)} [t, k]}{\partial(\mathbf{y})_m} \frac{\partial \Psi_I^{(j)} [t, k]}{\partial(\mathbf{y})_u} \\
&\quad + \frac{\partial \Psi_Q^{(j)} [t, k]}{\partial(\mathbf{y})_m} \frac{\partial \Psi_Q^{(j)} [t, k]}{\partial(\mathbf{y})_u}
\end{aligned} \tag{3.14}$$

where the terms involving the second derivative of $\Psi_I^{(j)} [t, k]$ (or $\Psi_Q^{(j)} [t, k]$) are all equal to zero. Since the variable in \mathbf{y} is either $\alpha_l^{(j,i)}$ or $\beta_l^{(j,i)}$, the calculation of $\partial^2 D / \partial(\mathbf{y})_m \partial(\mathbf{y})_u$ is equivalent to finding $\partial^2 D / \partial \alpha_l^{(j,i)} \partial \alpha_{l'}^{(j',i')}$, $\partial^2 D / \partial \beta_l^{(j,i)} \partial \beta_{l'}^{(j',i')}$, $\partial^2 D / \partial \alpha_l^{(j,i)} \partial \beta_{l'}^{(j',i')}$, and $\partial^2 D / \partial \beta_l^{(j,i)} \partial \alpha_{l'}^{(j',i')}$ in turn. By using (3.9)–(3.14) and the orthogonal property of STBC, as described in Section 3.2, it follows that

$$\begin{aligned}
\frac{\partial^2 D}{\partial \alpha_l^{(j,i)} \partial \alpha_{l'}^{(j',i')}} &= \frac{\partial^2 D}{\partial \beta_l^{(j,i)} \partial \beta_{l'}^{(j',i')}} \\
&= \begin{cases} 0, & \text{if } i \neq i' \text{ or } j \neq j' \\ 2 \sum_{k \in \Theta} \hat{C} [k] \cos \left(\frac{2\pi(k-1)(l-l')}{K} \right), & \text{o.w.} \end{cases}
\end{aligned} \tag{3.15}$$

$$\begin{aligned}
\frac{-\partial^2 D}{\partial \alpha_l^{(j,i)} \partial \beta_{l'}^{(j',i')}} &= \frac{\partial^2 D}{\partial \beta_l^{(j,i)} \partial \alpha_{l'}^{(j',i')}} \\
&= \begin{cases} 0, & \text{if } i \neq i' \text{ or } j \neq j' \\ 2 \sum_{k \in \Theta} \hat{C} [k] \sin \left(\frac{2\pi(k-1)(l-l')}{K} \right), & \text{o.w.} \end{cases}
\end{aligned} \tag{3.16}$$

where $\hat{C} [k] = \sum_{t=1}^{N_L} \left| \hat{X}^{(i)} [t, k] \right|^2$. According to (3.14)–(3.16), we can make two observations. One is that the matrix \mathbf{E} is related not only to the multi-path delay l but also to the estimate of the total transmitted signal energy

$\hat{C}[k]$ at the k th subcarrier for each transmit antenna. The other observation is that the matrix \mathbf{E} is a block diagonal matrix. Owing to the second observation, the iterative channel estimation method in (3.7) can be further simplified to [†]

$$\hat{\mathbf{y}}_v^{(j,i)} = \hat{\mathbf{y}}_{v-1}^{(j,i)} - \mathbf{g}_v^{(j,i)} \quad (3.17)$$

where $\mathbf{g}_v^{(j,i)}$ is obtained by computing $\mathbf{g}^{(j,i)} = \mathbf{E}^{(j,i)^{-1}} \mathbf{q}^{(j,i)}$ at $\mathbf{y}^{(j,i)} = \hat{\mathbf{y}}_{v-1}^{(j,i)}$ in which $\mathbf{E}^{(j,i)}$ is the truncated matrix obtained from the $((j-1)N_T + i)$ th diagonal block of \mathbf{E} , and $\mathbf{q}^{(j,i)}$ is the truncated vector merely containing the partial derivate of $\partial D / \partial \alpha_l^{(j,i)}$ and $\partial D / \partial \beta_l^{(j,i)}$, for $l = 1, \dots, N_T$ and $j = 1, \dots, N_R$.

3.3.3 Equivalence between Newton's Method and DF DFT-Based Method

We now turn our attention to deriving the equivalence between Newton's method and the DF DFT-based channel estimation method. By using (3.8)–(3.12) and defining $\tilde{\mathbf{q}}^{(j,i)} = \partial D / \partial \boldsymbol{\mu}_I^{(j,i)} + j \partial D / \partial \boldsymbol{\mu}_Q^{(j,i)}$, the gradient vector $\mathbf{q}^{(j,i)}$ in (3.17) is rewritten in a complex vector form as follows [†]

$$\tilde{\mathbf{q}}^{(j,i)} = \mathbf{F}^{(j,i)H} \boldsymbol{\Delta}^{(j,i)} \quad (3.18)$$

[†]It is noted that in this paper, $\mathbf{E}^{(j,i)}$ is a constant matrix for all antenna pairs. This is also the case for $\mathbf{F}^{(j,i)}$ and $\tilde{\mathbf{E}}^{(j,i)}$, but in practice, these matrices are specific for transceiver antenna pairs, depending upon path delays of the corresponding channel.

where $\mathbf{\Delta}^{(j,i)} = [\Delta^{(j,i)}[\Theta_1], \dots, \Delta^{(j,i)}[\Theta_{N_S}]]^T$, each element of which is calculated by

$$\Delta^{(j,i)}[k] = -2 \sum_{t=1}^{N_L} \Psi^{(j)}[t, k] \hat{X}^{(i)*}[t, k] \quad (3.19)$$

Moreover, $\mathbf{F}^{(j,i)}$ is an $N_S \times G$ truncated DFT matrix, with the (m, l) th element given by $\exp\{-j2\pi(\Theta_m - 1)(l - 1)/K\}$, and $(\cdot)^H$ is the Hermitian matrix of (\cdot) . By substituting $\Psi^{(j)}[t, k]$ of (3.5) into $\Delta^{(j,i)}[k]$ and applying the orthogonal property as described in Section 3.2, a more meaningful expression is provided by rewriting (3.19) in a column vector form:

$$\begin{aligned} \boldsymbol{\delta}^{(j)}[k] &= [\Delta^{(j,1)}[k], \dots, \Delta^{(j,N_T)}[k]]^T \\ &= 2\hat{\mathbf{X}}^*[k] \left(\hat{\mathbf{X}}^T[k] \mathbf{M}^{(j)}[k] - \mathbf{R}^{(j)}[k] \right) \\ &= 2 \left(\hat{C}[k] \mathbf{M}^{(j)}[k] - \hat{\mathbf{X}}^*[k] \mathbf{R}^{(j)}[k] \right) \end{aligned} \quad (3.20)$$

for $k \in \Theta$, and where $\mathbf{R}^{(j)}[k] = [R^{(j)}[1, k], \dots, R^{(j)}[N_L, k]]^T$, $\mathbf{M}^{(j)}[k] = [M^{(j,1)}[k], \dots, M^{(j,N_T)}[k]]^T$, and $\hat{\mathbf{X}}[k]$ is the re-encoded STBC matrix with $\hat{X}^{(i)}[t, k]$ as its element. Here, we observe that $\hat{\mathbf{X}}^*[k] \mathbf{R}^{(j)}[k]$ is an LS estimate for $[H^{(j,1)}[k], \dots, H^{(j,N_T)}[k]]^T$, and that $\boldsymbol{\delta}^{(j)}[k]$ represents the difference between the two channel frequency responses, $\hat{C}[k] \mathbf{M}^{(j)}[k]$ and $\hat{\mathbf{X}}^*[k] \mathbf{R}^{(j)}[k]$. From (3.18) and rewriting $\mathbf{E}^{(j,i)}$ in a complex matrix, $\tilde{\mathbf{E}}^{(j,i)}$, by using Appendix C, we have a complex-form representation of (3.17):

$$\hat{\boldsymbol{\mu}}_v^{(j,i)} = \hat{\boldsymbol{\mu}}_{v-1}^{(j,i)} - \tilde{\mathbf{g}}_v^{(j,i)} \quad (3.21)$$

where $\hat{\boldsymbol{\mu}}_v^{(j,i)}$ and $\tilde{\mathbf{g}}_v^{(j,i)}$ are the calculation associated with $\boldsymbol{\mu}^{(j,i)} = \boldsymbol{\mu}_I^{(j,i)} + j\boldsymbol{\mu}_Q^{(j,i)}$ and $\tilde{\mathbf{g}}^{(j,i)} = \tilde{\mathbf{E}}^{(j,i)-1} \tilde{\mathbf{q}}^{(j,i)}$, respectively, at $\boldsymbol{\mu}^{(j,i)} = \hat{\boldsymbol{\mu}}_{v-1}^{(j,i)}$ in which the (l, l') th entry of $\tilde{\mathbf{E}}^{(j,i)}$ is given by $2 \sum_{k \in \Theta} \hat{C}[k] \exp\{j2\pi(k - 1)(l - l')/K\}$. As shown

in Appendix D, the matrix $\tilde{\mathbf{E}}^{(j,i)^{-1}}$ in effect acts as a path decorrelator to decorrelate inter-path interference. It is desirable for the path decorrelator to be independent of $\hat{C}[k]$; therefore, $\tilde{\mathbf{E}}^{(j,i)^{-1}}$ only needs to be calculated once in each OFDM frame, containing several time slots. One way to achieve this is to normalize $\boldsymbol{\delta}^{(j)}[k]$ in (3.20) by $2\hat{C}[k]$ and to modify $\tilde{\mathbf{E}}^{(j,i)}$ as follows [†]

$$\boldsymbol{\delta}^{(j)}[k] = \mathbf{M}^{(j)}[k] - \frac{1}{\hat{C}[k]} \hat{\mathbf{X}}^*[k] \mathbf{R}^{(j)}[k] \quad (3.22)$$

$$\left(\tilde{\mathbf{E}}^{(j,i)}\right)_{l,l'} = \sum_{k \in \Theta} e^{j \frac{2\pi(k-1)(l-l')}{K}} \quad (3.23)$$

To be precise, $\tilde{\mathbf{E}}^{(j,i)}$ in (3.23) can be equivalently expressed as $\mathbf{F}^{(j,i)H} \mathbf{F}^{(j,i)}$. Finally, a truncated DFT matrix is applied to (3.21) to extrapolate the overall channel frequency response as follows

$$\hat{\mathbf{M}}_v^{(j,i)} = \hat{\mathbf{M}}_{v-1}^{(j,i)} - \mathbf{F}^{(j,i)} \hat{\mathbf{g}}_v^{(j,i)} \quad (3.24)$$

where $\hat{\mathbf{M}}_v^{(j,i)} = \mathbf{F}^{(j,i)} \hat{\boldsymbol{\mu}}_v^{(j,i)}$ is the estimated channel frequency response at the v th iteration, with respect to the channel frequency response $\mathbf{M}^{(j,i)} = [M^{(j,i)}[\Theta_1], \dots, M^{(j,i)}[\Theta_{N_S}]]^T$. It is clear that $\hat{\mathbf{M}}_{v-1}^{(j,i)}$ belongs to the subspace spanned by $\mathbf{F}^{(j,i)}$, and $\mathbf{F}^{(j,i)} \tilde{\mathbf{E}}^{(j,i)^{-1}} \mathbf{F}^{(j,i)H}$ is an orthogonal projection onto this subspace. From matrix theory, these two observations imply that

$$\begin{aligned} & \mathbf{F}^{(j,i)} \tilde{\mathbf{E}}^{(j,i)^{-1}} \mathbf{F}^{(j,i)H} \hat{\mathbf{M}}_{v-1}^{(j,i)} \\ &= \mathbf{F}^{(j,i)} \tilde{\mathbf{E}}^{(j,i)^{-1}} \mathbf{F}^{(j,i)H} \mathbf{F}^{(j,i)} \hat{\boldsymbol{\mu}}_{v-1}^{(j,i)} \\ &= \mathbf{F}^{(j,i)} \hat{\boldsymbol{\mu}}_{v-1}^{(j,i)} \\ &= \hat{\mathbf{M}}_{v-1}^{(j,i)} \end{aligned} \quad (3.25)$$

Hence, the vector $\hat{\mathbf{M}}_{v-1}^{(j,i)} - \mathbf{F}^{(j,i)} \tilde{\mathbf{E}}^{(j,i)-1} \mathbf{F}^{(j,i)H} \hat{\mathbf{M}}_{v-1}^{(j,i)}$ implicitly contained in the right side of (3.24) is zero. As a result, (3.24) is reduced to the DF DFT-based channel estimation method and the equivalence can be expressed as ††

$$\hat{\mathbf{M}}_v^{(j,i)} = \mathbf{F}^{(j,i)} \bar{\mathbf{g}}_v^{(j,i)} \quad (3.26a)$$

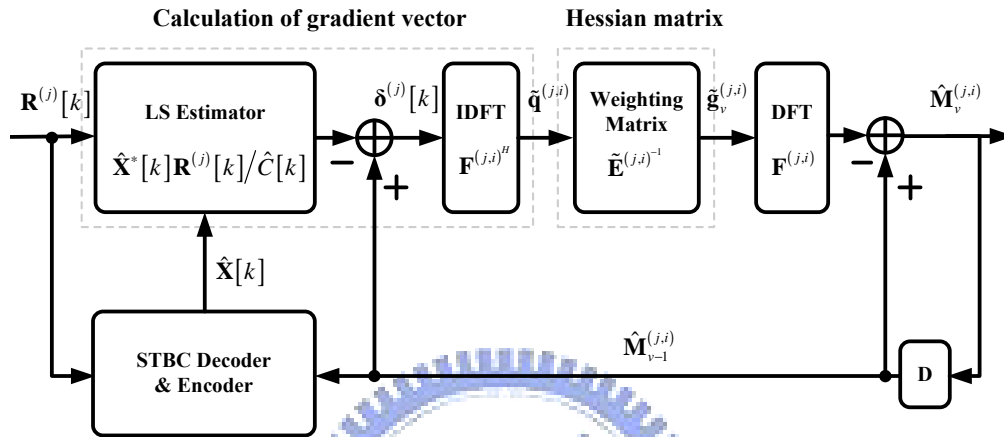
$$\bar{\mathbf{g}}^{(j,i)} = \tilde{\mathbf{E}}^{(j,i)-1} \bar{\mathbf{q}}^{(j,i)} \quad (3.26b)$$

$$\bar{\mathbf{q}}^{(j,i)} = \mathbf{F}^{(j,i)H} \bar{\Delta}^{(j,i)} \quad (3.26c)$$

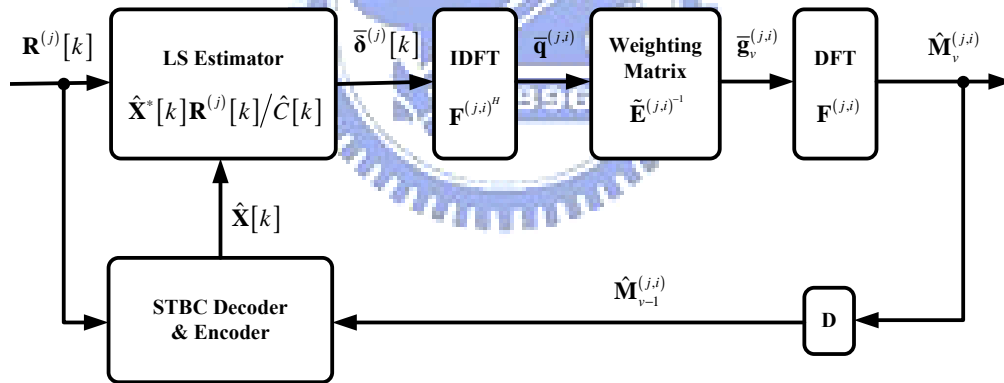
$$\bar{\delta}^{(j)}[k] = \frac{1}{\hat{C}[k]} \hat{\mathbf{X}}^*[k] \mathbf{R}^{(j)}[k], \text{ for } k \in \Theta \quad (3.26d)$$

where $\bar{\mathbf{g}}_v^{(j,i)}$ is the calculation with respect to $\bar{\mathbf{g}}^{(j,i)}$ at $\mathbf{M}^{(j,i)} = \hat{\mathbf{M}}_{v-1}^{(j,i)}$, and moreover, we define $\bar{\Delta}^{(j,i)} = [\bar{\Delta}^{(j,i)}[\Theta_1], \dots, \bar{\Delta}^{(j,i)}[\Theta_{N_S}]]^T$ and $\bar{\delta}^{(j)}[k] = [\bar{\Delta}^{(j,1)}[k], \dots, \bar{\Delta}^{(j,N_T)}[k]]^T$.

The above derivation clearly establishes the mathematical equivalence between Newton's method of (3.24) shown in Figure 3.3(a) and the DF DFT-based method of (3.26) shown in Figure 3.3(b). Our results indicate that both Newton's method and the DF DFT-based method for channel estimation in STBC-OFDM systems can be implemented through four components: an LS estimator, an IDFT matrix, a weighting matrix, and a DFT matrix. According to (3.18) and (3.22), we can also observe that the process of calculating the difference between the estimated channel frequency response $\mathbf{M}^{(j)}[k]$ and the LS estimate $\hat{\mathbf{X}}^*[k] \mathbf{R}^{(j)}[k] / \hat{C}[k]$, followed by the IDFT matrix $\mathbf{F}^{(j,i)H}$, is equivalent to forming the gradient vector in Newton's method. Moreover, the weighting matrix $\tilde{\mathbf{E}}^{(j,i)-1}$ in (3.26) is in fact the inverse of the Hessian matrix in Newton's method as observed in (3.21) and (3.23).



(a) Newton's method



(b) the DF DFT-based method

Figure 3.3: Equivalence between (a) Newton's method and (b) the DF DFT-based method.

3.4 Computer Simulation

The equivalence of the two methods is also verified by simulation in this section, with parameters listed in Table 3.1. We assume that the CSI in the previous time slot is known and utilized to initialize channel estimators. The parameter f_d denotes the maximum Doppler frequency, normalized to subcarrier spacing.

3.4.1 BER Performance

Figure 3.4 shows the equivalence of the two methods in terms of BER performance. The BER curve for ideal CSI is also provided for the purpose of calibration.

3.4.2 NSE Performance

As observed in Figure 3.5, the equivalence of the two methods is demonstrated in terms of the performance of normalized square error (NSE) between true and estimated CSI.

^{††}Thanks to the orthogonal property of STBC, extending the results in SISO-OFDM systems [36–38] to the STBC-OFDM systems should be straightforward.

Table 3.1: Simulation parameters.

Parameter	Value
Carrier frequency	2.3GHz
Bandwidth	5MHz
FFT size (K)	256
Length of CP (G)	64
Number of data subcarriers (M)	200
Number of data subcarriers used (N_S)	200
Modulation	QPSK
Number of transmit antennas (N_T)	2
Number of receive antennas (N_R)	1
Channel power profiles	ITU Veh-A [34] and Jakes model [33]
Channel delay profiles	0 ~ 63 (Uniform distribution)
Normalized maximum Doppler frequency (f_d)	0.01, 0.05

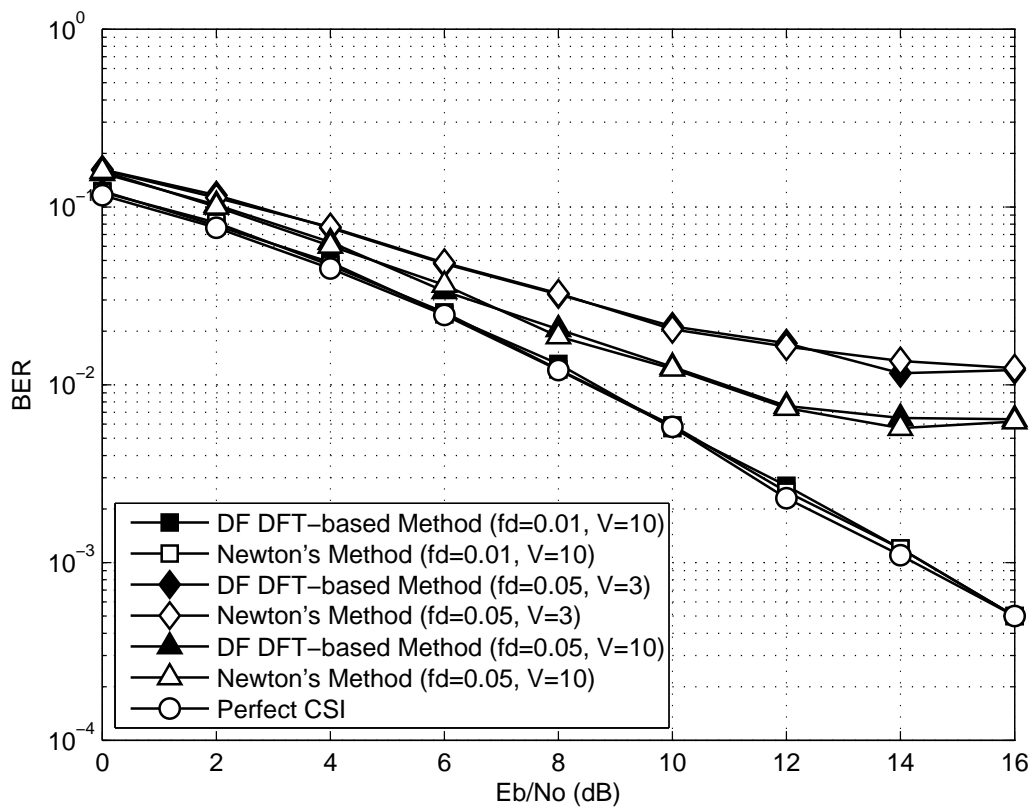


Figure 3.4: BER performance of the two methods.

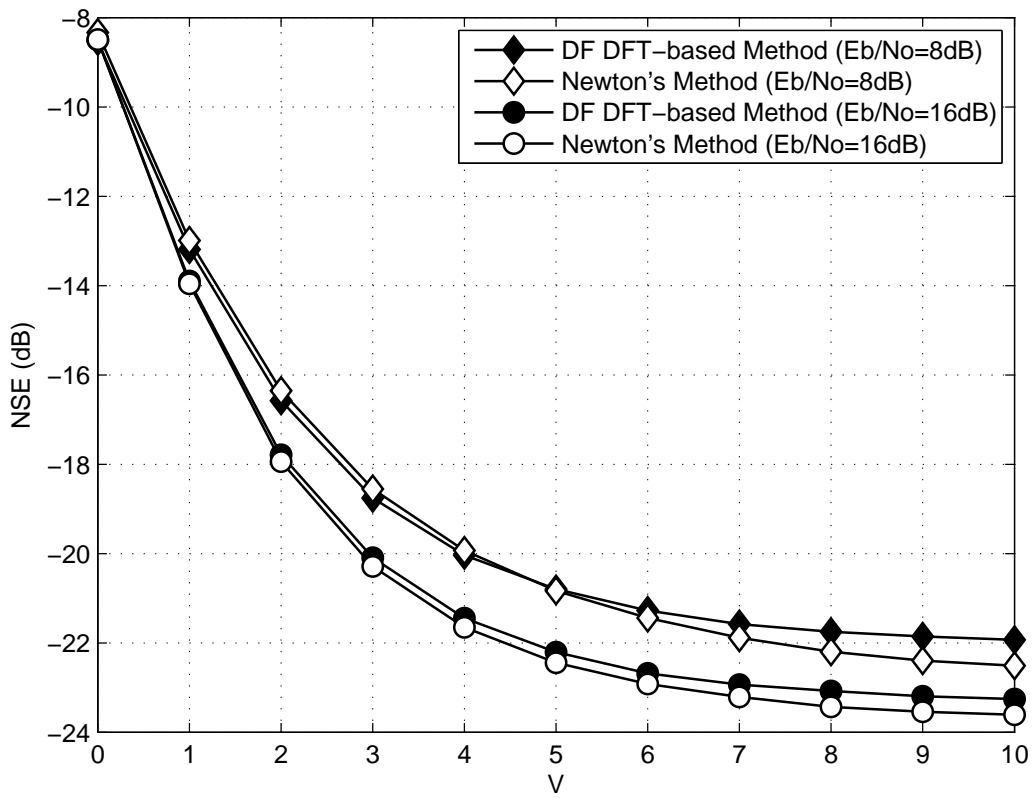
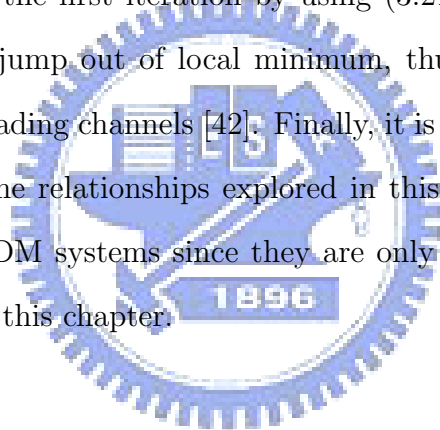


Figure 3.5: NSE performance of the two methods ($f_d = 0.05$).

3.5 Summary

In this chapter, we present a derivation on the equivalence between Newton's method and the DF DFT-based method for channel estimation in STBC-OFDM systems. The results could provide useful insights for the development of new algorithms. For example, extending the DF DFT-based method to the Levenberg-Marquardt method is quite simple through this equivalence, which is particularly helpful when the inverse for the weighting matrix does not exist [41]. As another example, a few pilot tones can be applied to form a gradient vector at the first iteration by using (3.21) and to help the DF DFT-based method jump out of local minimum, thus improving the BER performance in fast fading channels [42]. Finally, it is worth mentioning that the derivation and the relationships explored in this chapter are also valid for conventional OFDM systems since they are only simplified cases of the systems discussed in this chapter.



Chapter 4

A Refined Channel Estimation Method for STBC-OFDM Systems in Low-Mobility Wireless Channels



4.1 Literature Survey and Motivation

OFDM has been widely applied in wireless communication systems in recent years due to its capability of high-rate transmission and low-complexity implementation over frequency-selective fading channels. STC is another promising technique to provide diversity gain through the use of multiple transmit antennas, especially when receive diversity is too expensive to deploy. In particular, STBC has received a lot of attention because a simple linear decoder can be used at the receiver side [12, 13, 40]. These advantages make OFDM combined with STBC, known as STBC-OFDM, an ideal

choice for several applications such as wireless metropolitan area networks (WMANs) 802.16e [43], etc. However, a high-rate STBC-OFDM system employing multi-level modulation with non-constant envelope (e.g. 16QAM) generally requires accurate CSI to perform coherent detection. This in turn implies that dynamic channel estimation is a crucial factor in realizing a successful STBC-OFDM system over doubly selective channels.

Blind channel estimation, which merely relies on the received signals, is very attractive due to its bandwidth-saving advantage. Nevertheless, it requires a long data record, involves high computational complexity, and only applies to slowly time-varying channels. On the contrary, pilot-aided channel estimation, using pilot tones known to the receiver, shows great promise for applications in mobile wireless communication, even though the use of pilot tones ends up with lower data rate. DF channel estimation offers an alternative way to track channel variations; nevertheless, it is vulnerable to decision error propagation in fast time-varying channels [35–37, 44–46]. As a high quality channel estimator with low training overhead is needed for successful implementation of STBC-OFDM systems, we restrict our attention to the category of pilot-aided plus DF channel estimation methods in this chapter.

Among a wide variety of pilot-aided plus DF channel estimation methods, the DFT-based channel estimation method, derived from either MMSE criterion or ML criterion, has been intensively studied for OFDM systems with preambles [36–39]. It is shown in [35–38] that the DFT-based channel estimation method using the ML criterion is simpler to implement because it

requires neither channel statistics nor operating signal-to-noise ratio (SNR). Furthermore, as presented in [38], the performance of the ML estimator is comparable to that of the MMSE estimator at intermediate or high SNR values when the number of pilot tones is sufficiently larger than the maximal channel length (in samples). Thus, we will focus on the ML estimator in this chapter. In order to save bandwidth and improve system performance, DF data symbols are also used as pilots to track channel variations in subsequent OFDM data symbols, and this method is called the DF DFT-based channel estimation method. Recently, the mathematical equivalence between the DF DFT-based method and Newton's method has been studied in [47], and it is concluded that even though a global solution for CSI is given as the initial value in the preamble, the DF DFT-based method is only applicable to very slowly time-varying channels because of the local optimization capability of Newton's method.

Most mobile wireless channels are characterized by channel impulse response (CIR) consisting of a few dominant paths. The multipath delays are usually slowly time-varying. The amplitude and the phase of each path, however, can vary relatively fast. In this chapter, we propose a two-stage channel estimation method by utilizing these channel characteristics. In the initialization stage, we employ an MPIC-based decorrelation method to identify significant paths. In the following tracking stage, we develop a refined DF DFT-based channel estimation method, in which we use a few pilot tones inserted in OFDM data symbols to form an optimal gradient vector at the first iteration. This optimal gradient vector helps the classical method jump

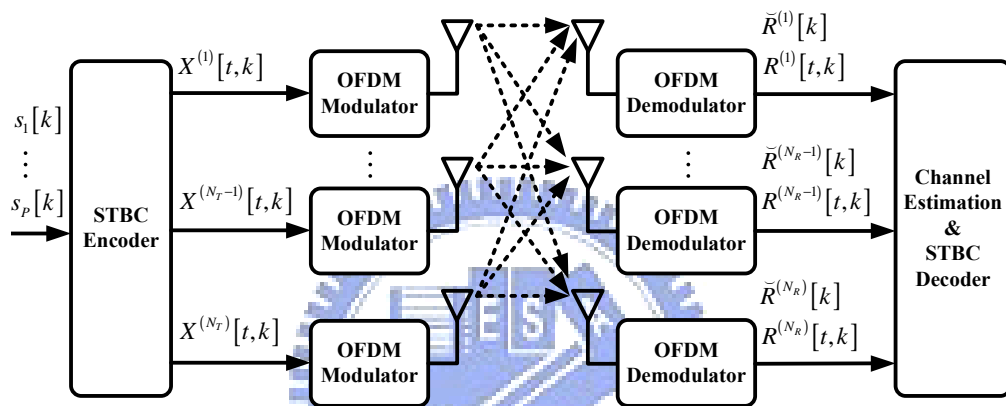
out of the local optimum, thus reducing the error propagation effect. The classical DF DFT-based channel estimation method is then used at the following iterations. In addition, an approximate weighting matrix is adopted to reduce the computational complexity associated with the matrix inversion operation of the weighting matrix in the DF DFT-based channel estimation method.

The rest of this chapter is organized as follows. In Section 4.2, we describe the STBC-OFDM system. In Section 4.3, we present the MPIC-based decorrelation method in the initialization stage. Next, the equivalence between the DF DFT-based channel estimation method and Newton's method is briefly reviewed, and we propose a refined DF DFT-based channel estimation method in the tracking stage. We then discuss the computational complexity of the proposed two-stage channel estimation method in Section 4.4. In Section 4.5, we present our computer simulation and performance evaluation results. Finally, some concluding remarks are drawn in Section 5.6.

4.2 STBC-OFDM Systems

4.2.1 Transmitted Signals

As shown in Figure 4.1(a), we consider an STBC-OFDM system with N_T transmit and N_R receive antennas, employing K subcarriers among which M subcarriers are used to transmit data symbols plus pilot tones and the other $K - M$ subcarriers are used as either a DC subcarrier or virtual sub-



(a) STBC-OEDM system



(b) OFDM frame format

Figure 4.1: (a) STBC-OFDM system (b) OFDM frame format.

carriers at the two edges to avoid the spectrum overlapping problem at the receiver. Assume that \mathbf{Q} and \mathbf{J} denote the sets of subcarrier indices assigned to transmit data symbols and pilot tones, respectively. At subcarrier $k \in \mathbf{Q}$ (or $k \in \mathbf{J}$) and after symbol mapping, P modulated data symbols (or pilot tones) $\{s_1[k], \dots, s_P[k]\}$ are encoded by a $T \times N_T$ STBC encoder $\mathbf{X}[k]$ to generate N_T signal sequences of length T , each of which is denoted by $\{X^{(i)}[1, k], \dots, X^{(i)}[T, k]\}$, for $i = 1, \dots, N_T$ [12, 13, 40]. These N_T signal sequences possess the orthogonal property given by $\mathbf{X}^H[k] \mathbf{X}[k] = C[k] \mathbf{I}_{N_T}$, where $C[k] = C^{(i)}[k] = \sum_{t=1}^T |X^{(i)}[t, k]|^2$ is the total transmitted signal energy at the k th subcarrier for each transmit antenna, and they are simultaneously transmitted by the N_T transmit antennas within the duration of T OFDM data symbols. The notation $(\cdot)^H$ represents the operation of taking Hermitian and the notation \mathbf{I}_N represents an $N \times N$ identity matrix. At the t th OFDM data symbol and the i th transmit antenna, after insertion of $K - M$ zeros for DC and virtual subcarriers, STBC encoded data symbols and pilot tones $X^{(i)}[t, k]$, for $k \in \mathbf{Q}$ and $k \in \mathbf{J}$, respectively, are modulated onto M subcarriers via a K -point IDFT unit to produce time domain samples. A CP is then added in front of each OFDM data symbol to eliminate intersymbol interference caused by multipath channels. As shown in Figure 4.1(b), each OFDM frame starts with a CP-added preamble which occupies one OFDM symbol and is followed by N_D consecutive OFDM data symbols. In order to avoid IAI at the receiver side, pilot tones are alternatively inserted into the available subcarriers of the preamble with antenna-specific subcarrier shifts [43], and the preamble in frequency domain is denoted by

$P^{(i)} [k]$, for $k \in \mathbf{Q} \cup \mathbf{J}$ and $i = 1, \dots, N_T$.

4.2.2 Channel Model

The complex baseband representation of impulse response for a mobile wireless channel (in discrete expression) between the i th transmit antenna and the j th receive antenna can be described by

$$h^{(j,i)} [t, \tau] = \sum_{l=1}^{L^{(j,i)}} \mu_l^{(j,i)} (t) \delta [\tau - \tau_l^{(j,i)}] \quad (4.1)$$

where $L^{(j,i)}$ is the number of resolvable paths, $\tau_l^{(j,i)}$ is the time delay in samples of the l th path, and $\mu_l^{(j,i)} (t)$ is the complex Gaussian fading gain of the l th path. All paths $\mu_l^{(j,i)} (t)$, for $l = 1, \dots, L^{(j,i)}$, are assumed to be independent of each other. Thus, with proper cyclic extension, the channel frequency response can be expressed as

$$H^{(j,i)} [t, k] = \sum_{l=1}^{L^{(j,i)}} \mu_l^{(j,i)} (t) e^{-j \frac{2\pi k \tau_l^{(j,i)}}{K}} \quad (4.2)$$

where k is the subcarrier index.

4.2.3 Received Signals

We assume that both timing and carrier frequency synchronization are perfect, and that the length of channel impulse response is always smaller than the length of the CP. Another assumption here is that the channel is quasi-static over the duration of each time slot. The time index t in $H^{(j,i)} [t, k]$ is dropped hereafter. Hence, after the OFDM demodulator in Figure 4.1(a),

the received preamble at the j th receive antenna can be expressed as

$$\check{R}^{(j)} [k] = \sum_{i=1}^{N_T} H^{(j,i)} [k] P^{(i)} [k] + Z^{(j)} [k] \quad (4.3)$$

for $k \in \mathbf{Q} \cup \mathbf{J}$. Moreover, the T successively received OFDM data symbols at the j th receive antenna are given by

$$R^{(j)} [t, k] = \sum_{i=1}^{N_T} H^{(j,i)} [k] X^{(i)} [t, k] + Z^{(j)} [t, k] \quad (4.4)$$

for $t = 1, \dots, T$ and $k \in \mathbf{Q} \cup \mathbf{J}$. Note that $Z^{(j)} [k]$ in (4.3) or $Z^{(j)} [t, k]$ in (4.4) are uncorrelated AWGN with zero-mean and variance σ_Z^2 on the j th receive antenna.

4.3 Proposed Channel Estimation Method

In this section, we first present the MPIC-based decorrelation method for the initialization stage. Before introducing the refined DF DFT-based channel estimation method for the tracking stage, we briefly review the equivalence between the DF DFT-based channel estimation method and Newton's method.

4.3.1 Initialization Stage: The MPIC-Based Decorrelation Method

We all know that CIR can be estimated using the preamble placed at the beginning of each OFDM frame, while the difficulty is that for most wireless standards, the preamble does not have ideal auto-correlation due to the

Step1:

Set preassumed number of paths N_h & observation window \mathbf{W}_b

Calculate $C_{\bar{R}P}[\tau]$ & $C_{PP}[\tau]$ by

$$C_{\bar{R}P}[\tau] = \text{IDFT}\{\tilde{R}[k] \cdot P^*[k]\}, \text{ for } \tau = 0, \dots, K-1$$

$$C_{PP}[\tau] = \text{IDFT}\{P[k] \cdot P^*[k]\}, \text{ for } \tau = 0, \dots, K-1$$

Initialize $\rho = 0$ & $\kappa = 0$

Step2:

Estimate multipath delays and complex gains coarsely:

while $\rho < N_h$

$\rho \leftarrow \rho + 1$

$$\tilde{\tau}_\rho = \arg \max_{\tau \in \mathbf{W}_b} \{C_{\bar{R}P}[\tau]\}$$

 if $\tilde{\tau}_\rho > G \cdot K$

$$C_{\bar{R}P}[\tilde{\tau}_\rho] = 0$$

 else

$\kappa \leftarrow \kappa + 1$

 find a legal path with time delay $\hat{\tau}_\kappa = \tilde{\tau}_\rho$ & complex path gain $\hat{\mu}_\kappa = C_{\bar{R}P}[\tilde{\tau}_\rho]$

$$C_{\bar{R}P}[\tau] \leftarrow C_{\bar{R}P}[\tau] - \hat{\mu}_\kappa C_{PP}[\tau - \hat{\tau}_\kappa], \quad \tau \in \mathbf{W}_b \setminus \{\tilde{\tau}_1, \dots, \tilde{\tau}_{\rho-1}\}$$

 end

end

Figure 4.2: The MPIC-based decorrelation method in the initialization stage. (G is the ratio of the CP length to the useful OFDM symbol time, and $\text{IDFT}\{\cdot\}$ is a K -point IDFT operation.)

use of either guard band or non-equally spaced pilot tones. Figure 4.2 outlines the MPIC-based decorrelation method to estimate CIR path-by-path by canceling out already known multipath interference. Since the preambles transmitted from different antennas do not interfere with each other at the receiver side, channel estimation can be independently performed for each transceiver antenna pair, and therefore the antenna indices j and i are omitted in the following. In step 1, we first define two parameters N_h and \mathbf{W}_b which represent a preassumed number of paths in a mobile radio channel and a multipath observation window, respectively. Next, we calculate the cyclic cross-correlation $C_{RP}[\tau]$ between the received and the transmitted preamble as well as the normalized cyclic auto-correlation $C_{PP}[\tau]$ of the transmitted preamble. Both ρ and κ , which stand for a path counting variable and the number of legal paths found by the MPIC-based decorrelation method, respectively, are initialized to zero. In step 2, we start by increasing the value of the path counting variable, ρ , by one, and picking only one path whose time delay $\tilde{\tau}_\rho$ yields the largest value in $|C_{RP}[\tau]|$, for $\tau \in \mathbf{W}_b$. If the time delay $\tilde{\tau}_\rho$ is larger than the length of the CP, this path is treated as an illegal path, thereby discarded by setting $C_{RP}[\tilde{\tau}_\rho] = 0$. Otherwise, we increase the number of legal paths found, κ , by one, and then reserve this path as the κ th legal path with time delay $\hat{\tau}_\kappa = \tilde{\tau}_\rho$ and complex path gain $\hat{\mu}_\kappa = C_{RP}[\tilde{\tau}_\rho]$. The replica of the interference associated with this legal path is regenerated and subtracted from $C_{RP}[\tau]$ to obtain a refined cross-correlation function:

$$C_{RP}[\tau] \leftarrow C_{RP}[\tau] - \hat{\mu}_\kappa C_{PP}[|\tau - \hat{\tau}_\kappa|] \quad (4.5)$$

for $\tau \in \mathbf{W}_b \setminus \{\tilde{\tau}_1, \dots, \tilde{\tau}_{\rho-1}\}$, where " \leftarrow " is the assignment operation. We continue the iterative process of the step 2 until ρ reaches the preassumed value of N_h .

4.3.2 Equivalence between DF DFT-Based Method and Newton's Method

Through the initialization stage, we are able to obtain information on the number of paths $\kappa^{(j,i)} (\leq N_h)$, the multipath delays $\hat{\tau}_l^{(j,i)}$, the multipath complex gains $\hat{\mu}_l^{(j,i)}$, for $l = 1, \dots, \kappa^{(j,i)}$, and therefore the corresponding channel frequency response. With the assumption of the unchanged multipath delays over the duration of each OFDM frame, the DF DFT-based channel estimation method using the ML criterion (at the v th iteration) can be summarized as follows [36–38, 47]. First, the re-encoded STBC matrix is used to obtain the LS channel estimate [48]:

$$\boldsymbol{\delta}_v^{(j)} = \hat{\mathbf{C}}_v^{-1} \hat{\mathbf{X}}_v^H \mathbf{R}^{(j)} \quad (4.6)$$

where $\mathbf{R}^{(j)} = [\mathbf{R}^{(j)T} [\Theta_1], \dots, \mathbf{R}^{(j)T} [\Theta_{|\Theta|}]]^T$ is the received signal vector at the j th receive antenna in which we have $\mathbf{R}^{(j)} [k] = [R^{(j)} [1, k], \dots, R^{(j)} [T, k]]^T$ and Θ is a subset of \mathbf{Q} used to track channel variations, v is the iteration number from 1 to V , $\hat{\mathbf{X}}_v = \text{diag} \left\{ \hat{\mathbf{X}}_v [\Theta_1], \dots, \hat{\mathbf{X}}_v [\Theta_{|\Theta|}] \right\}$ consists of the re-encoded STBC matrix $\hat{\mathbf{X}}_v [k]$ with decision symbols $\hat{X}_v^{(i)} [t, k]$ which are obtained by applying the previously estimated channel frequency response to decode the received signal $\mathbf{R}^{(j)} [k]$ according to [12, 13, 40], and $\hat{\mathbf{C}}_v = \text{diag} \left\{ \hat{C}_v [\Theta_1] \mathbf{I}_{N_T}, \dots, \hat{C}_v [\Theta_{|\Theta|}] \mathbf{I}_{N_T} \right\}$ is the energy normalization factor in which $\hat{C}_v [k] = \hat{C}_v^{(i)} [k] = \sum_{t=1}^T \left| \hat{X}_v^{(i)} [t, k] \right|^2$. The notation $(\cdot)^T$ takes

the transpose of (\cdot) , the notation $(\cdot)^{-1}$ represents the matrix inversion of (\cdot) , the notation $diag\{\mathbf{X}_1, \dots, \mathbf{X}_N\}$ represents a block diagonal matrix, and the notation $|\Theta|$ denotes the dimension of the set Θ . Thus, it is clear to show that the $((m-1)N_T + i)$ th entry of $\delta_v^{(j)}$ represents the LS estimate for the Θ_m th subcarrier at the i th transmit and the j th receive antenna. Second, a truncated DFT matrix is applied to improve the LS estimate as follows:

$$\mathbf{q}_v^{(j)} = \mathbf{F}^{(j)H} \Delta_v^{(j)} \quad (4.7)$$

$$\mathbf{E}^{(j)} = \mathbf{F}^{(j)H} \mathbf{F}^{(j)} \quad (4.8)$$

$$\mathbf{M}_v^{(j)} = \mathbf{F}^{(j)} \mathbf{E}^{(j)^{-1}} \mathbf{q}_v^{(j)} \quad (4.9)$$

where $\Delta_v^{(j)} = \mathbf{\Pi} \delta_v^{(j)}$ and $\mathbf{\Pi}$ is an $N_T |\Theta| \times N_T |\Theta|$ permutation matrix obtained by making the $((i-1)|\Theta| + m)$ th row have all zeros except for a single one at the $((m-1)N_T + i)$ th column for $i = 1, \dots, N_T$ and $m = 1, \dots, |\Theta|$, $\mathbf{F}^{(j)} = diag\{\mathbf{F}^{(j,1)}, \dots, \mathbf{F}^{(j,N_T)}\}$, $\mathbf{F}^{(j,i)}$ is the $|\Theta| \times \kappa^{(j,i)}$ truncated DFT matrix whose (m, l) th entry is defined as $\exp\{-j2\pi\Theta_m \hat{\tau}_l^{(j,i)}/K\}$, and the $((i-1)|\Theta| + m)$ th entry of $\mathbf{M}_v^{(j)}$ represents the estimated channel frequency response for the Θ_m th subcarrier at the v th iteration, corresponding to the (j, i) th antenna pair. In order to initialize the channel estimator of (4.9), the CSI estimated in the previous time slot has to be taken as the initial value of the CSI for the current time slot.

It has been shown in [47] that the DF DFT-based channel estimation method can be equivalently expressed in the framework of Newton's method as [41]:

$$\bar{\delta}_v^{(j)} = \bar{\mathbf{M}}_{v-1}^{(j)} - \hat{\mathbf{C}}_v^{-1} \hat{\mathbf{X}}_v^H \mathbf{R}^{(j)} \quad (4.10)$$

$$\bar{\mathbf{q}}_v^{(j)} = \mathbf{F}^{(j)H} \bar{\Delta}_v^{(j)} \quad (4.11)$$

$$\mathbf{M}_v^{(j)} = \mathbf{M}_{v-1}^{(j)} - \mathbf{F}^{(j)} \mathbf{E}^{(j)-1} \bar{\mathbf{q}}_v^{(j)} \quad (4.12)$$

where $\bar{\mathbf{M}}_{v-1}^{(j)} = \mathbf{\Pi}^{-1} \mathbf{M}_{v-1}^{(j)}$ and $\bar{\Delta}_v^{(j)} = \mathbf{\Pi} \bar{\delta}_v^{(j)}$. According to [47], the purpose of calculating the difference between $\bar{\mathbf{M}}_{v-1}^{(j)}$ and $\hat{\mathbf{C}}_v^{-1} \hat{\mathbf{X}}_v^H \mathbf{R}^{(j)}$ in (4.10), followed by the IDFT matrix $\mathbf{F}^{(j)H}$ in (4.11), is to form the gradient vector $\bar{\mathbf{q}}_v^{(j)}$ in Newton's method, as observed in (4.12). Furthermore, it is also proved in [47] that the role of the weighting matrix $\mathbf{E}^{(j)-1}$ in (4.12) is in fact the inverse of the Hessian matrix in Newton's method.

4.3.3 Refined DF DFT-Based Channel Estimation

Through the equivalence relation described in Section 4.3.2, it is concluded that the classical DF DFT-based channel estimation method (called method I, hereafter) is limited by the local search capability of Newton's method and only applicable to very slowly time-varying channels. In the previous studies [36] [37], pilot tones as well as decision data symbols are simultaneously adopted to perform channel estimation at each iteration. This is, however, not a good solution in time-varying channels because decision data symbols easily induces the error propagation effect, whereas pilot tones are much more reliable than decision data symbols. From the viewpoint of optimization, the pilot tones inserted in each OFDM data symbol can play a more important role in providing a global search direction at the first iteration of the method I. Figure 4.3 shows the refined DF DFT-based channel estimation method in the tracking stage. With the help from a few pilot tones to form a gradient

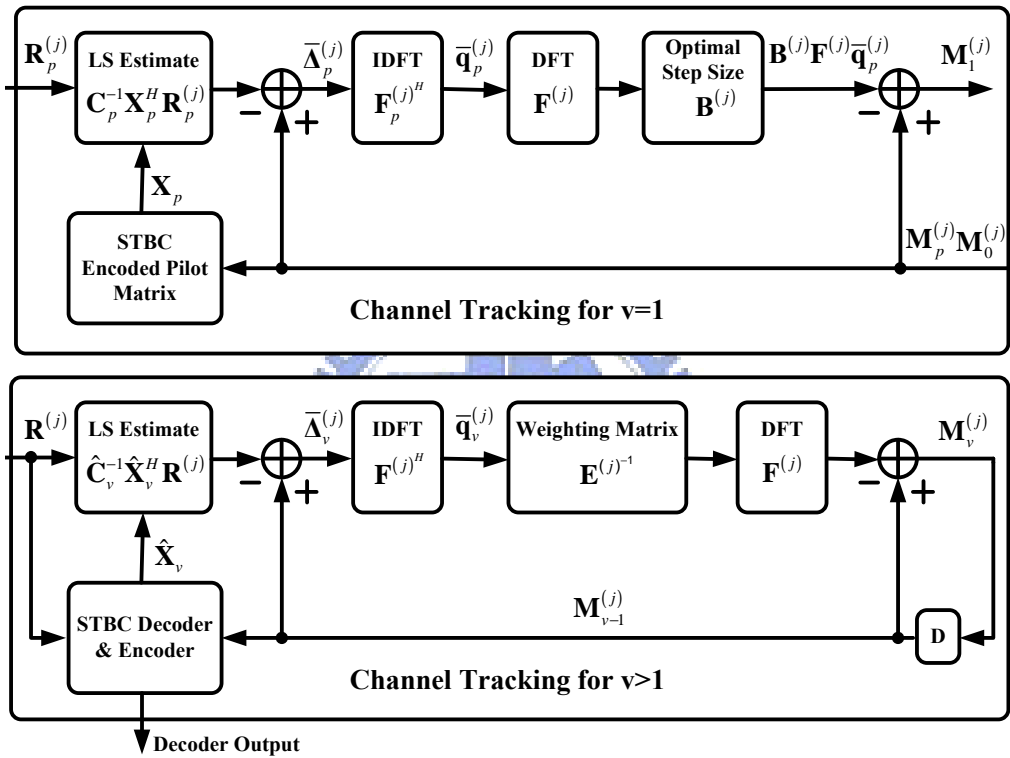


Figure 4.3: Block diagram of the refined DF DFT-based channel estimation method in the tracking stage. (The subscript "p" is to indicate that the calculation is only associated with the pilot subcarrier set.)

vector according to (4.10) and (4.11), the refined channel tracking method (called method II, hereafter) is proposed here by only modifying the first iteration ($v = 1$) of (4.12) as follows

$$\mathbf{M}_1^{(j)} = \mathbf{M}_0^{(j)} - \mathbf{B}^{(j)} \mathbf{F}^{(j)} \bar{\mathbf{q}}_p^{(j)} \quad (4.13)$$

where $\bar{\mathbf{q}}_p^{(j)}$ is the gradient vector calculated according to (4.10) and (4.11) by only utilizing the set \mathbf{J} , instead of Θ , $\mathbf{B}^{(j)}$ is a block diagonal matrix defined as $\text{diag} \{ \beta^{(j,1)} \mathbf{I}_{|\Theta|}, \dots, \beta^{(j,N_T)} \mathbf{I}_{|\Theta|} \}$, and $\beta^{(j,i)}$ is a real-valued step size which can be determined by minimizing the following ML cost function over the pilot subcarrier indices:

$$\begin{aligned} \boldsymbol{\beta}_{opt} &= \arg \min_{\boldsymbol{\beta}} \Omega(\boldsymbol{\beta}) \\ &= \arg \min_{\boldsymbol{\beta}} \sum_{j=1}^{N_R} \left| \mathbf{R}_p^{(j)} - \mathbf{X}_p \mathbf{\Pi}_p^{-1} (\mathbf{M}_p^{(j)} - \mathbf{B}_p^{(j)} \mathbf{F}_p^{(j)} \bar{\mathbf{q}}_p^{(j)}) \right|^2 \end{aligned} \quad (4.14)$$

where $\boldsymbol{\beta} = [\boldsymbol{\beta}^{(1)T}, \dots, \boldsymbol{\beta}^{(N_R)T}]^T$ in which we define $\boldsymbol{\beta}^{(j)} = [\beta^{(j,1)}, \dots, \beta^{(j,N_T)}]^T$, the notations $\mathbf{R}_p^{(j)}$, \mathbf{X}_p , $\mathbf{M}_p^{(j)}$, $\mathbf{B}_p^{(j)}$, $\mathbf{F}_p^{(j)}$ and $\mathbf{\Pi}_p$ are defined similar to $\mathbf{R}^{(j)}$, $\hat{\mathbf{X}}_v$, $\mathbf{M}_v^{(j)}$, $\mathbf{B}^{(j)}$, $\mathbf{F}^{(j)}$ and $\mathbf{\Pi}$, respectively, but here the set \mathbf{J} is used instead of Θ in the definition. Note that $\mathbf{M}_p^{(j)}$ is the estimated channel frequency obtained from the previous time slot. Let $M^{(j,i)}[k]$ and $\xi^{(j,i)}[k]$ for $k = J_m$ denote the $((i-1)|\mathbf{J}| + m)$ th entries of $\mathbf{M}_p^{(j)}$ and $\mathbf{F}_p^{(j)} \bar{\mathbf{q}}_p^{(j)}$, respectively. Furthermore, we define $\gamma^{(j)}[t, k] = R^{(j)}[t, k] - \sum_{i=1}^{N_T} M^{(j,i)}[k] X^{(i)}[t, k]$ and $\phi^{(j,i)}[t, k] = \xi^{(j,i)}[k] X^{(i)}[t, k]$. By expanding (4.14) and taking $\partial \Omega(\boldsymbol{\beta}) / \partial \boldsymbol{\beta} = \mathbf{0}$, after straightforward manipulations, the optimum value of $\boldsymbol{\beta}^{(j)}$ is given by

$$\boldsymbol{\beta}_{opt}^{(j)} = -\Phi^{(j)-1} \mathbf{\Gamma}^{(j)} \quad (4.15)$$

where the (m, u) th entry of the matrix $\mathbf{\Phi}^{(j)}$ is calculated by

$$(\mathbf{\Phi}^{(j)})_{m,u} = \sum_{t=1}^T \sum_{k \in \mathbf{J}} \Re \{ \phi^{(j,m)} [t, k] \phi^{(j,u)*} [t, k] \} \quad (4.16)$$

and the u th entry of the vector $\mathbf{\Gamma}^{(j)}$ is given by

$$(\mathbf{\Gamma}^{(j)})_u = \sum_{t=1}^T \sum_{k \in \mathbf{J}} \Re \{ \phi^{(j,u)} [t, k] \gamma^{(j)*} [t, k] \} \quad (4.17)$$

The notation $\Re(\cdot)$ takes the real part of (\cdot) . Note that, after the first iteration, we execute the channel tracking process of (4.10)–(4.12) for the second and subsequent iterations until a stopping criterion holds. The stopping criterion is to check whether the absolute value of each entry in $\mathbf{F}^{(j)} \mathbf{E}^{(j)-1} \mathbf{q}_v^{(j)}$ is less than a prespecified threshold ε or the iteration number v reaches the maximum value of V . The channel tracking process for the current time slot will be stopped when either of the above two conditions holds.

In order to reduce computational complexity in the method II, we further propose method III to avoid the matrix inverse of the weighting matrix, $\mathbf{E}^{(j)-1}$, by taking into account the strongly diagonal property of $\mathbf{E}^{(j)}$ which is originally proposed for reducing the complexity of multiuser detection in code division multiple access (CDMA) systems [49]. Define $\mathbf{E}^{(j)} = |\mathbf{\Theta}| (\mathbf{I}_{\kappa^{(j)}} + \mathbf{O}_{off})$, where $\kappa^{(j)} = \sum_{i=1}^{N_T} \kappa^{(j,i)}$ and \mathbf{O}_{off} is a zero-diagonal matrix. Then, it follows that if $|\mathbf{\Theta}|$ is large enough, an approximate weighting matrix of $\mathbf{E}^{(j)-1}$ takes the form:

$$\begin{aligned} \mathbf{E}^{(j)-1} &= \frac{1}{|\mathbf{\Theta}|} (\mathbf{I}_{\kappa^{(j)}} + \mathbf{O}_{off})^{-1} \\ &\approx \frac{1}{|\mathbf{\Theta}|} (\mathbf{I}_{\kappa^{(j)}} - \mathbf{O}_{off}) \\ &= \frac{1}{|\mathbf{\Theta}|^2} (2|\mathbf{\Theta}| \mathbf{I}_{\kappa^{(j)}} - \mathbf{E}^{(j)}) \end{aligned} \quad (4.18)$$

4.4 Computational Complexity

Now let us look at the computational complexity of the three methods, in terms of the number of real multiplications per transceiver antenna pair. In general, the operations of K -point IDFT and $K \times K$ matrix inversion need $4K \log_2 K$ and $4K^3$ real multiplications, respectively. Besides, the weighting matrix $\mathbf{E}^{(j)-1}$ only needs to be calculated once in each OFDM frame as it is only related to the multipath delays $\hat{\tau}_l^{(j,i)}$. Therefore, the complexity considered in the initialization stage is mainly due to the operations of $C_{RP}[\tau]$ and $\mathbf{E}^{(j)-1}$. The calculation and update of $C_{RP}[\tau]$ require at most $4(|\mathbf{Q}| + |\mathbf{J}|)/N_T + 4K \log_2 K + 4N_h |\mathbf{W}_b|$ real multiplications. Moreover, the calculation of $\mathbf{E}^{(j)-1}$ needs at most $4N_h^3$ real multiplications, but it needs at most $2N_h^2$ real multiplications if the approximate weighting matrix of $\mathbf{E}^{(j)-1}$ in (4.18) is used instead. In the tracking stage, the computation for each iteration of the method I (or each of the second and subsequent iterations of the method II and the method III) involves the calculation of $\mathbf{F}^{(j)} \mathbf{E}^{(j)-1} \bar{\mathbf{q}}_v^{(j)}$, in total requiring at most $|\Theta|(4T+2) + 2|\Theta|T/N_T + 8K \log_2 K + 4N_h^2$ real multiplications. For the method II and the method III, an optimal gradient vector is formed at the first iteration, in which the computation of $\mathbf{B}^{(j)} \mathbf{F}^{(j)} \bar{\mathbf{q}}_p^{(j)}$ and the optimum $\beta^{(j)}$ at most requires $|\mathbf{J}|(4T+2) + 8K \log_2 K + 2|\Theta| + 10|\mathbf{J}|T + N_T^2 + N_T(2|\mathbf{J}|T+1)$ real multiplications. The computational complexity for the system parameters given in Section 4.5 is listed in Table 4.1. The values of the parameters $|\mathbf{W}_b|$, N_h , T , N_T , K , $|\mathbf{Q}|$, and $|\mathbf{J}|$ are set as 109, 8, 2, 2,

256, 192, and 8, respectively, while the value of the parameter $|\Theta|$ can be 48, 96, or 192. For a fair performance comparison, both the data subcarrier set Θ and the pilot subcarrier set \mathbf{J} are used for tracking channel variations, except that only the pilot subcarrier set is adopted at the first iteration of the method II and the method III. Hence, we use $|\Theta| + |\mathbf{J}|$ to replace $|\Theta|$ in the calculation of the complexity in Table 4.1. As observed in Table 4.1, the complexity of the method II and the method III is a little bit lower than that of the method I. It can also be noticed that the complexity in the tracking stage is mostly due to the operations of the DFT and the IDFT which in total require $8K \log_2 K = 16384$ real multiplications. Some complexity gain can be achieved by using partial DFT processing such as in [37]. If $|\Theta|$ is larger than N_h , the computation of the partial DFT processing mainly depends on the size of Θ . As a result, the complexity of the tracking stage is basically dominated by $|\Theta|$ since $|\mathbf{J}|$, N_T and T are usually much smaller than $|\Theta|$.

4.5 Computer Simulation

We demonstrate the performance of the proposed channel estimation methods through computer simulation of an STBC-OFDM system with two transmit antennas and a single receive antenna. The parameters are set the same as those in the 802.16e OFDM standard [43] and summarized in Table 4.2. The system occupies a bandwidth of $5MHz$ and operates in the $2.3GHz$ frequency band. The entire bandwidth is divided into $K = 256$ subcarriers among which $|\mathbf{Q}| + |\mathbf{J}| = 200$ subcarriers are used to transmit data symbols and pilot tones, and the remaining $M = 56$ subcarriers are used as virtual

Table 4.1: Computational complexity for the system parameters given in Section 4.5.

Initialization Stage			
Method I and Method II	14128		
Method III	12208		
Tracking Stage ($ \mathbf{J} = 8$)			
$ \Theta $	48	96	192
The first iteration of the Method II and the Method III	16806	16902	17094
Each iteration of the Method I or each of the second and subsequent iterations of the Method II and the Method III	17312	17888	19040

Table 4.2: Simulation parameters.

Parameter	Value
Carrier frequency	2.3GHz
Bandwidth	5MHz
FFT size (K)	256
Length of CP	64
Number of data and pilot subcarriers	200
Number of virtual subcarriers (M)	56
Modulation scheme for data subcarriers	QPSK, 16QAM
Modulation scheme for pilot subcarriers	BPSK
Number of OFDM data symbols per frame (N_D)	40
Channel power profiles	Two-path channel ITU Veh-A channel [34] Jakes model [33]
Channel delay profiles	0 ~ 50 (Uniform distribution)
Multipath observation window (\mathbf{W}_b)	[0, 108]
Preassumed number of paths (N_h)	4, 8

subcarriers at the two edges and a DC subcarrier. In the simulation, the modulation schemes for the data symbols are QPSK and 16QAM, while the BPSK modulation scheme is adopted for the pilot tones. Each pilot subcarrier transmits the same power as each data subcarrier. Each OFDM frame is composed of one OFDM preamble and $N_D = 40$ OFDM data symbols. The length of the CP is 64 sample periods, i.e., one quarter of the useful symbol time. The preambles transmitted from the first and second antennas use even and odd subcarriers respectively with a $3dB$ power boost, and the values of those subcarriers are set according to [43]. Both a conventional two-path channel and an International Telecommunication Union (ITU) Veh-A channel are simulated with path delays uniformly distributed from 0 to 50 sample periods, where the relative path power profiles are set as 0, 0 (dB) for the two-path channel and 0, -1 , -9 , -10 , -15 , -20 (dB) for the ITU Veh-A channel [34]. The vehicle speed v_e of $240km/hr$ is used to simulate mobile radio environments, for which Rayleigh fading is generated by Jakes model [33]. Moreover, the multipath observation window \mathbf{W}_b is set as $[0, 108]$. The preassumed number of paths N_h is set as 4 and 8 in the two-path channel and the ITU Veh-A channel, respectively. Both the data subcarrier set Θ and the pilot subcarrier set \mathbf{J} are used in the tracking stage. The subcarrier indices of \mathbf{J} are uniformly assigned within the available subcarriers. The set Θ is uniformly selected from \mathbf{Q} , and the parameter $|\Theta|$ could be 192, 96, 48, 24, 12, 6 or 3. The values of the maximum iteration number V are set as 5 and 7 for QPSK and 16QAM modulation, respectively. For the stopping criterion, the prespecified threshold ε is set as 10^{-4} . The entire simulations

are conducted in the equivalent baseband, and we assume both symbol synchronization and carrier synchronization are perfect. Finally, throughout the simulation, the parameter E_b/N_o is defined as a ratio of received bit energy to the power spectral density of noise.

For comparison purpose, the performance curve with ideal channel estimation, denoted as perfect CSI, is provided for reference and served as a performance lower bound. We also compare the proposed methods with both the STBC-based MMSE method [48] and the Kalman filtering method [50] [51] where the decision-feedback methodology is employed under the assumption of ideal channel estimation in the initialization stage. Some statistical information such as Doppler frequency, auto-covariance of channels, and noise power is assumed to be known for these two existing methods. It is noted that the Kalman filtering method is mainly based on [50] and the received signals within a time slot are used to perform channel estimation according to the decision-feedback steps in [51].

4.5.1 NSE Performance of MPIC-Based Decorrelation Method

Figure 4.4 shows the NSE performance of the MPIC-based decorrelation method in the initialization stage. The performance measure is the averaged NSE between the actual channel frequency response and the estimated channel frequency response. We can see from this figure that the NSE decreases monotonically as the E_b/N_o increases, but the channel estimation performance is eventually limited by the residual multipath interference at

high E_b/N_o . It is also observed that the MPIC-based decorrelation method in the two-path channel has better NSE performance than that in the ITU Veh-A channel.

4.5.2 BER Performance

Figure 4.5 shows the BER performance for QPSK modulation in the two-path channel at $v_e = 240km/hr$ with $|\mathbf{J}| = 8$ and $|\Theta| = 192$. As can be observed in Figure 4.5, at $BER=10^{-2}$, the required E_b/N_o for both the method II and the method III is about $3dB$ less than that for the method I, and at most $0.2dB$ more than that for the perfect CSI case. Compared with the STBC-based MMSE method and the Kalman filtering method, the method II and the method III also yield much better BER performance. Figure 4.6 and Figure 4.7 give the BER performance for 16QAM modulation in the two-path channel and the ITU Veh-A channel, respectively, at $v_e = 240km/hr$ with $|\mathbf{J}| = 8$ and $|\Theta| = 192$. As shown in Figure 4.6, it is clear that with the help from the pilot gradient vector at the first iteration, both the method II and the method III provide a substantial gain in E_b/N_o , and there is only $0.8dB$ degradation in the required E_b/N_o at $BER=10^{-2}$ compared with the performance curve of the perfect CSI case. However, for the method I, a higher error floor is clearly visible at a BER of 4×10^{-2} . Obviously, the two existing methods also suffer from severe degradation in the BER performance due to the error propagation effect, even though they are a little better than the method I. Similarly to the above results, Figure 4.7 shows that both the method II and the method III significantly outperform the method I, whereas

they still have $3.6dB$ gap in E_b/N_o compared with the perfect CSI case. The two existing methods still have a significant E_b/N_o gap compared with the method II and the method III.

4.5.3 Effect of Normalized Maximum Doppler Frequency

Figure 4.8 and Figure 4.9 show BER versus f_d in the ITU Veh-A channel for QPSK modulation and 16QAM modulation, respectively, where f_d is the maximum Doppler frequency, normalized by the subcarrier spacing. Clearly, the BER performance gap between the method I and the method II (or the method III) becomes larger as f_d increases up to 0.07 and 0.05 in the cases of QPSK and 16QAM modulation, respectively. It is worthy to mention that the method III is comparable to the method II at different maximum Doppler frequencies, even though an approximate weighting matrix is adopted in the method III. For calibration purpose, the BER performance for the three proposed methods without error propagation is also simulated in these two figures. We observe that the adoption of the optimal gradient vector effectively reduces the error propagation effect and even the BER performance without error propagation is attainable for f_d up to 0.05 and 0.03 in the cases of QPSK and 16QAM modulation, respectively.

4.5.4 Effect of Number of Pilot Tones

Figure 4.10 shows BER versus $|\mathbf{J}|$ in the ITU Veh-A channel for 16QAM modulation at $v_e = 240km/hr$ with $|\Theta| = |\mathbf{Q}|$ and $E_b/N_o = 20dB$. As can be observed from this figure, the BER performance of the method II (or

the method III) converges faster than the method I, while it is improved continually until $|\mathbf{J}|$ increases to 8. Moreover, both the method II and the method III can still give better BER performance even when the number of pilot tones is small.

4.5.5 Effect of Number of Data Subcarriers Used

Figure 4.11 shows BER versus $|\Theta|$ in the ITU Veh-A channel for 16QAM modulation at $v_e = 240km/hr$ with $|\mathbf{J}| = 8$ and $E_b/N_o=20dB$. This figure indicates that there is a trade-off between BER and $|\Theta|$, and that one can reduce computational complexity by using a smaller $|\Theta|$ at the cost of slight performance degradation whenever $|\Theta|$ is larger than 48.

4.5.6 Average Number of Iterations

Figure 4.12 compares the average number of iterations in the ITU Veh-A channel for 16QAM modulation at $v_e = 240km/hr$ with $|\mathbf{J}| = 8$ and $E_b/N_o=20dB$. Due to the use of the approximate weighting matrix, when $|\Theta|$ is larger than 48, the method III requires one more iteration compared with the method II. However, the method III is still an attractive approach by decreasing V from 7 to 6 at the price of slight BER performance degradation when the computational complexity of matrix inverse is an issue.

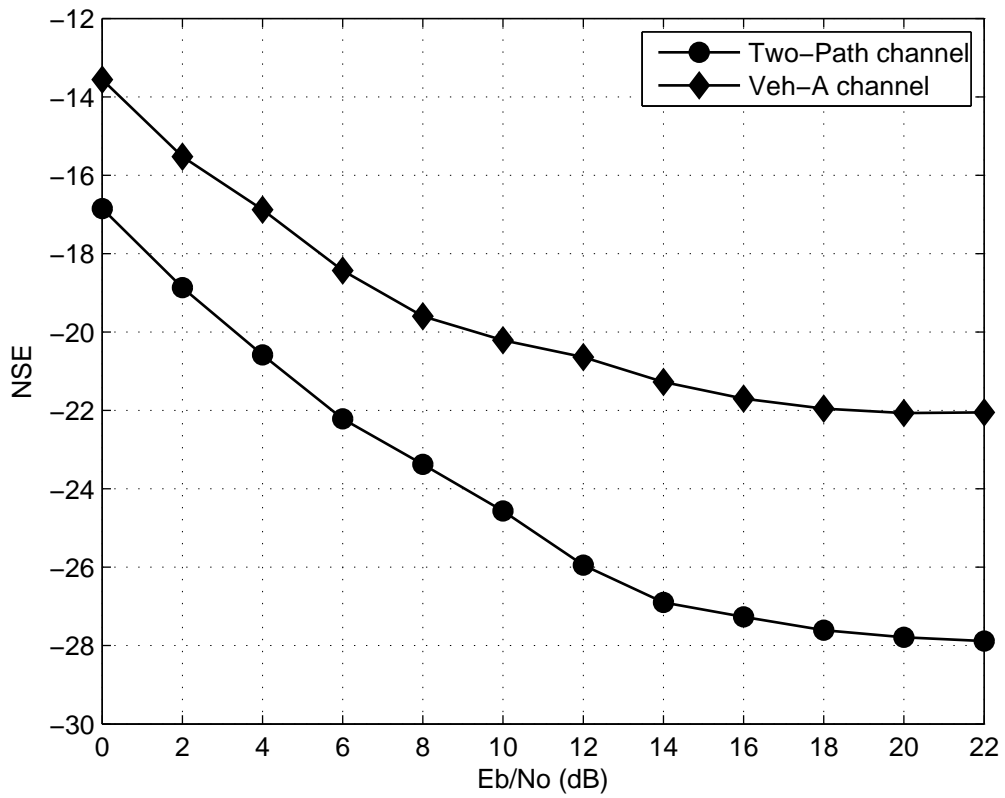


Figure 4.4: NSE performance of the MPIC-based decorrelation method in the initialization stage ($v_e = 240km/hr$).

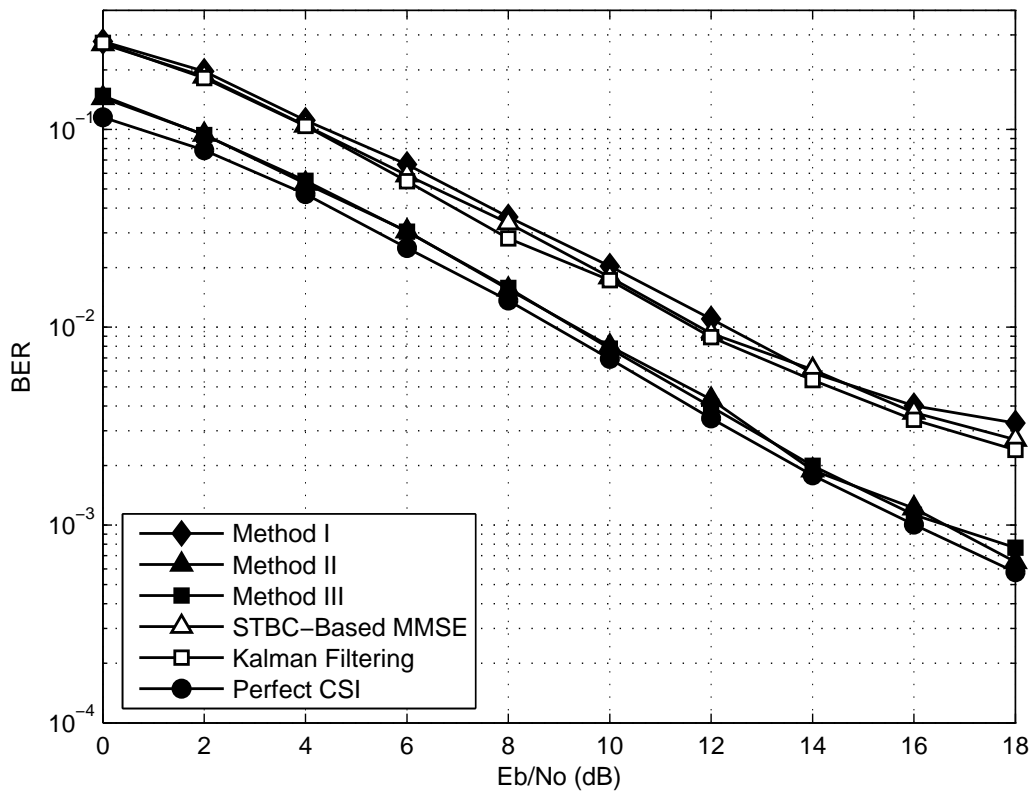


Figure 4.5: BER performance for QPSK modulation in the two-path channel at $v_e = 240\text{km/hr}$ ($|\mathbf{J}| = 8$ and $|\Theta| = |\mathbf{Q}| = 192$).

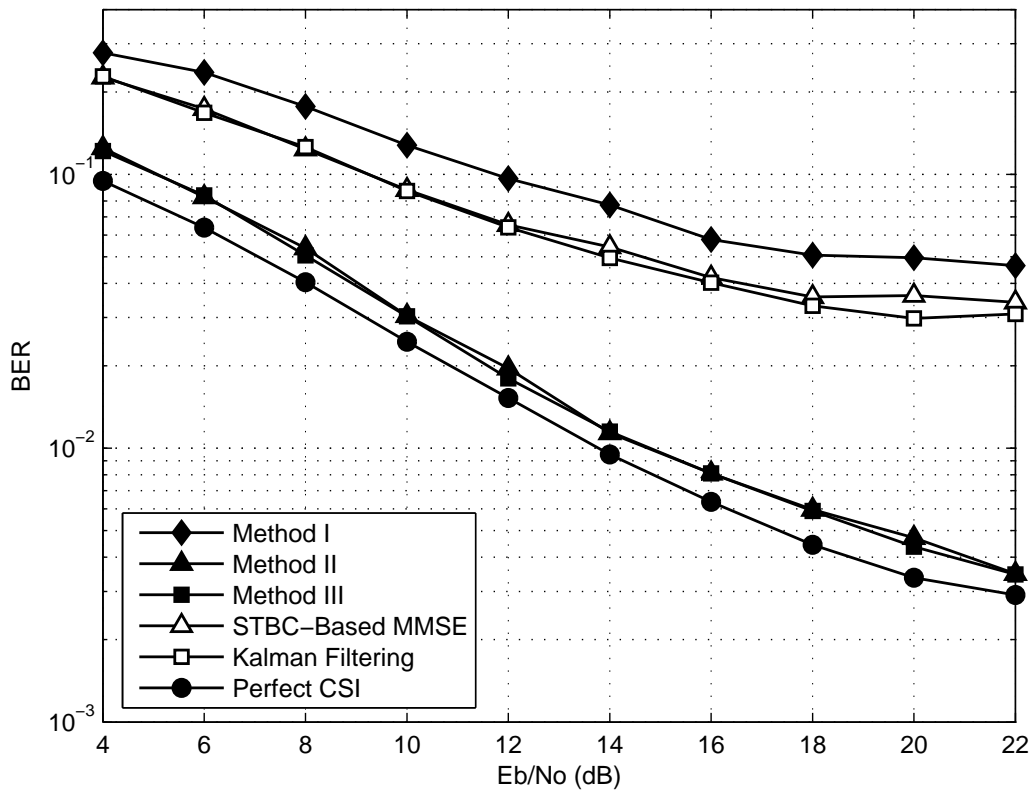


Figure 4.6: BER performance for 16QAM modulation in the two-path channel at $v_e = 240\text{km/hr}$ ($|\mathbf{J}| = 8$ and $|\Theta| = |\mathbf{Q}| = 192$).

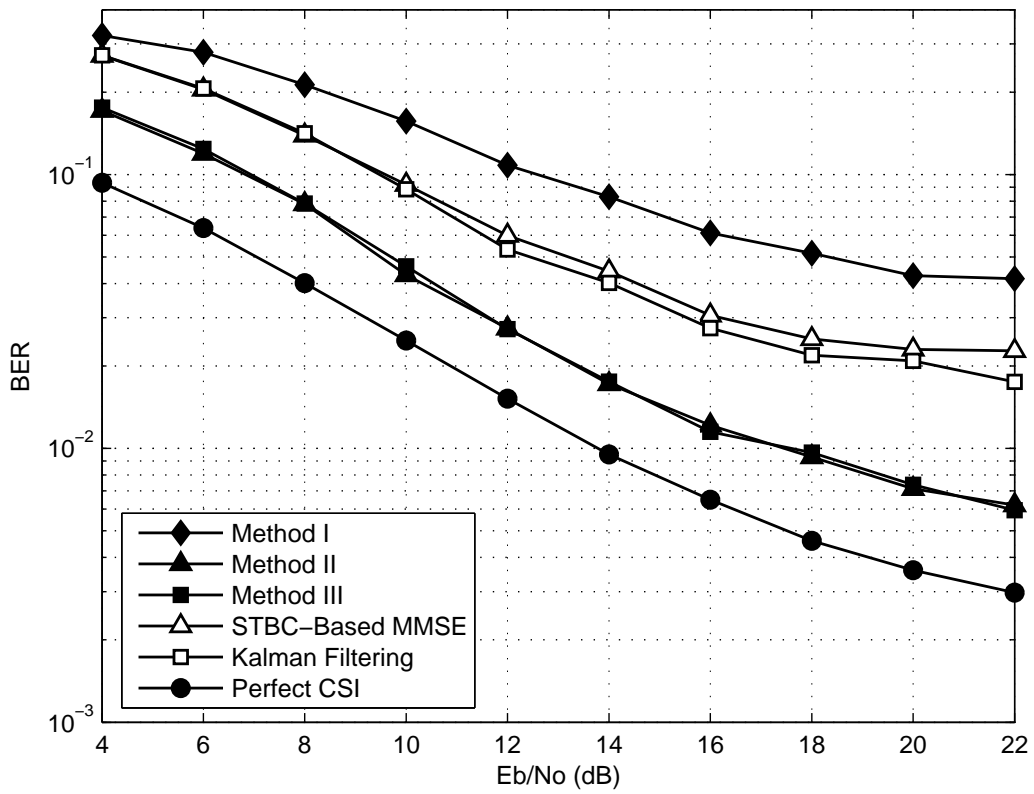


Figure 4.7: BER performance for 16QAM modulation in the ITU Veh-A channel at $v_e = 240\text{km/hr}$ ($|\mathbf{J}| = 8$ and $|\Theta| = |\mathbf{Q}| = 192$).

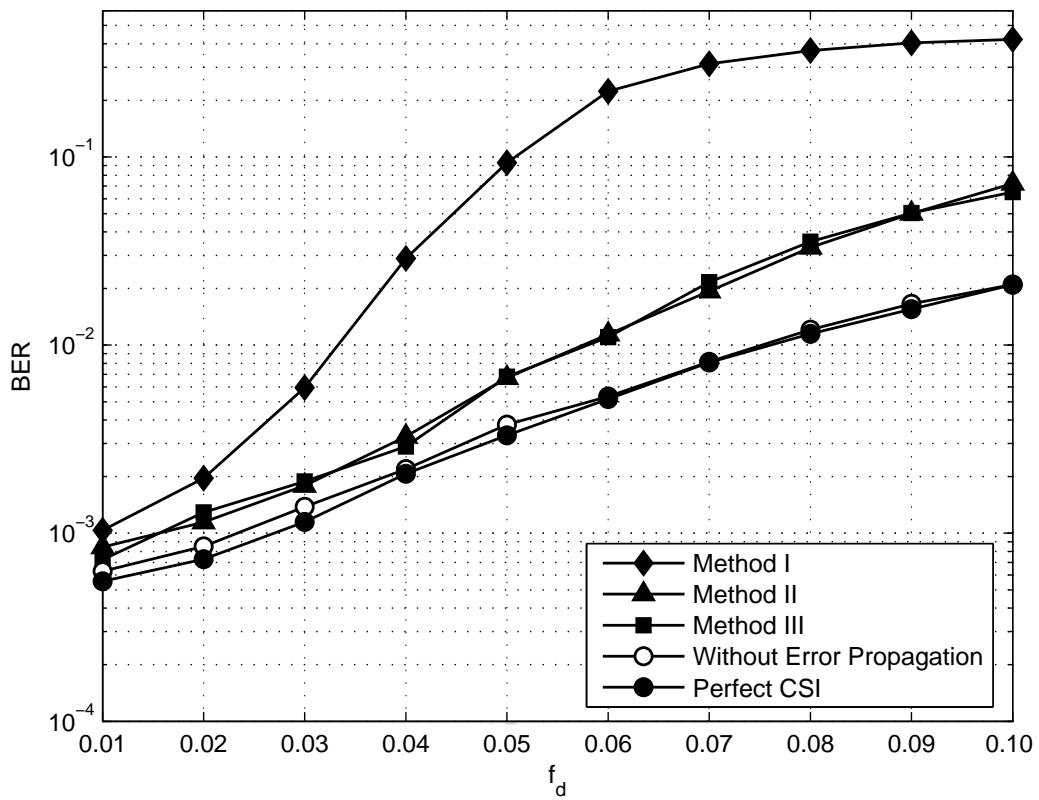


Figure 4.8: BER versus normalized maximum Doppler frequency in the ITU Veh-A channel for QPSK modulation ($|\mathbf{J}| = 8$, $|\Theta| = |\mathbf{Q}| = 192$, and $E_b/N_o=16dB$).

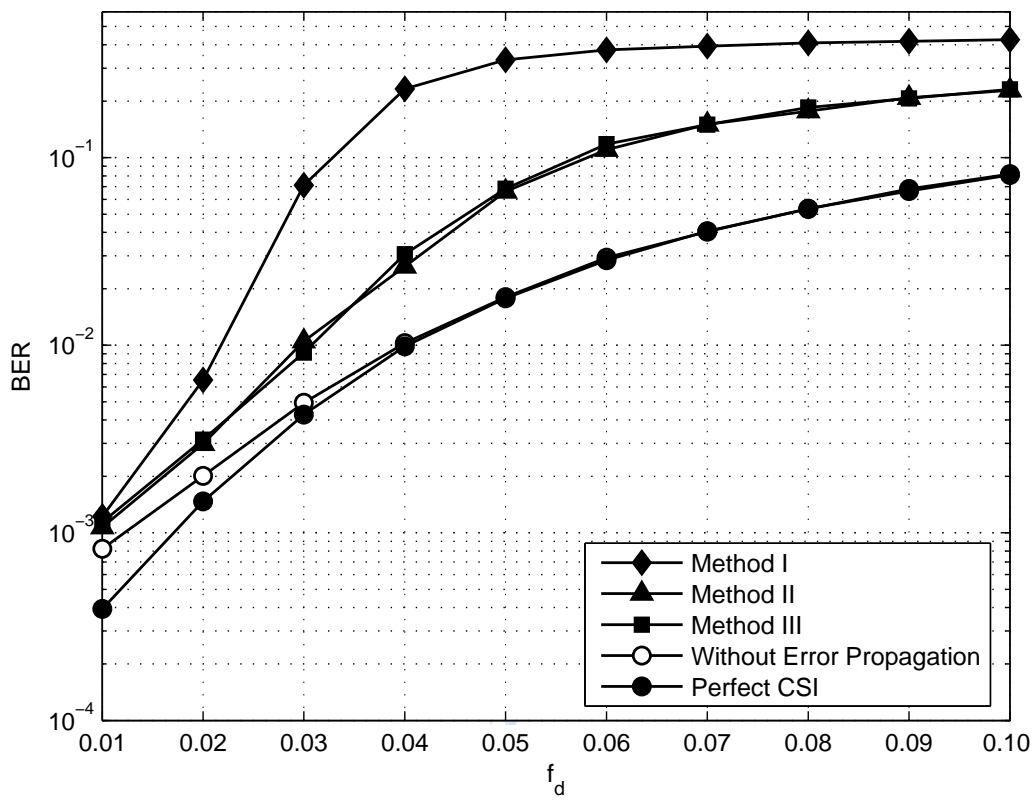


Figure 4.9: BER versus normalized maximum Doppler frequency in the ITU Veh-A channel for 16QAM modulation ($|\mathbf{J}| = 8$, $|\Theta| = |\mathbf{Q}| = 192$, and $E_b/N_o=22dB$).

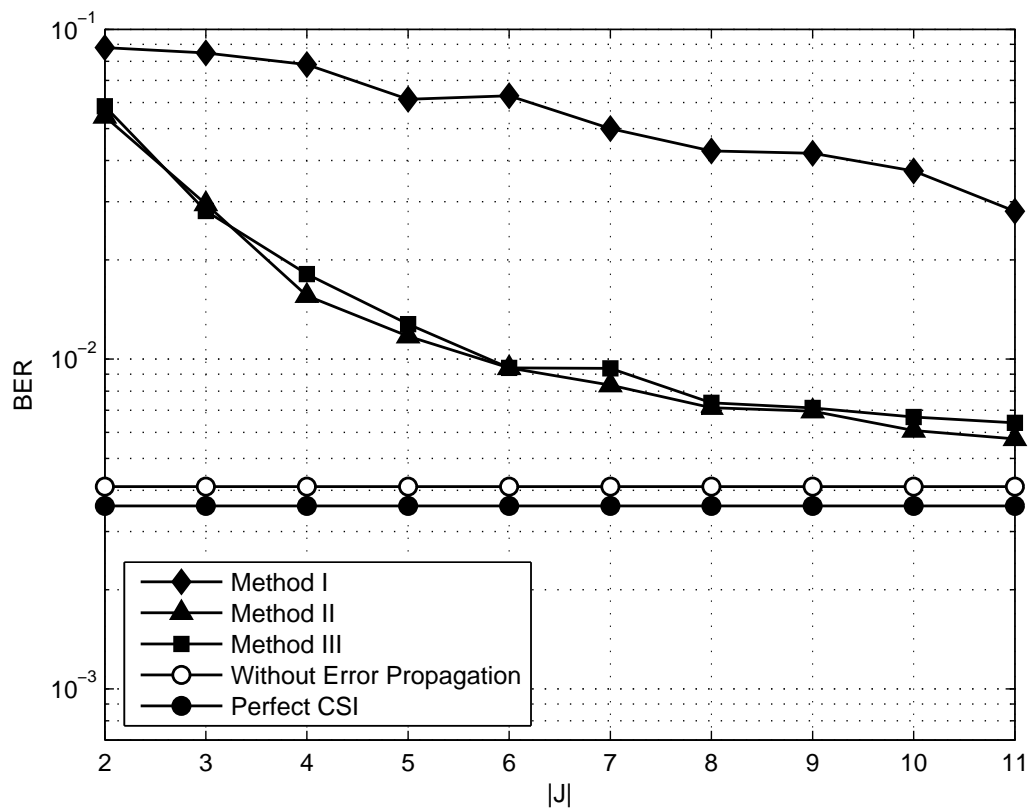


Figure 4.10: BER versus number of pilot tones used in the ITU Veh-A channel for 16QAM modulation ($v_e = 240km/hr$, $|\Theta| = |\mathbf{Q}|$, and $E_b/N_o=20dB$).

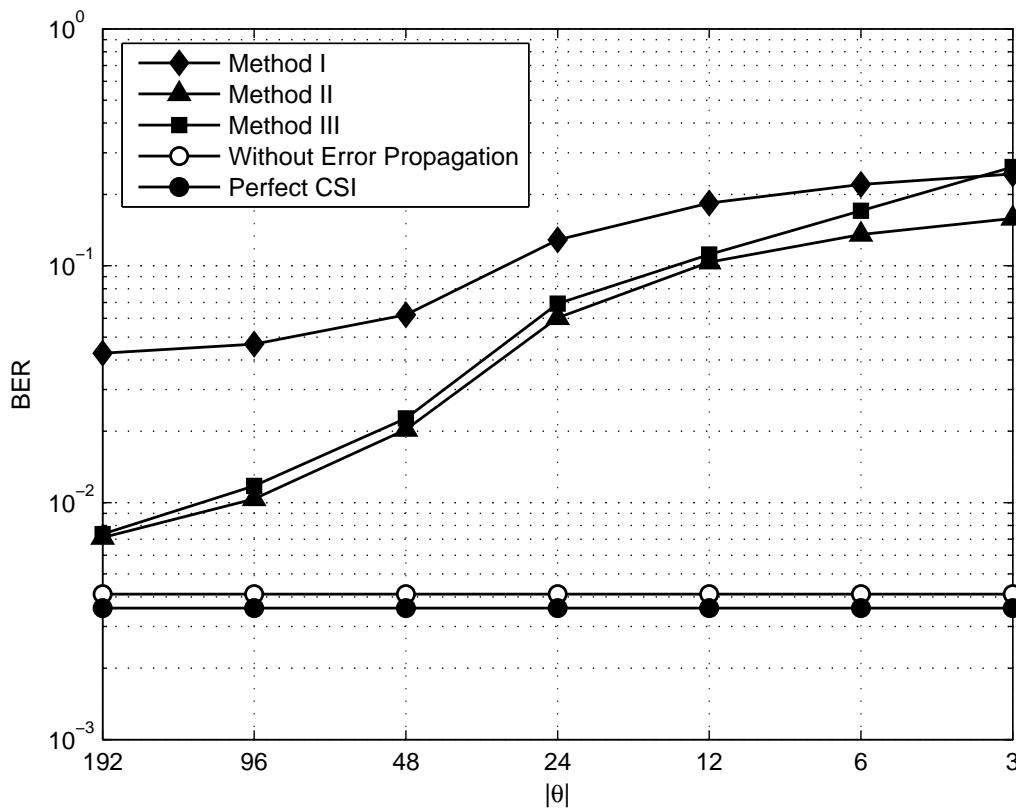


Figure 4.11: BER versus number of data subcarriers used in the ITU Veh-A channel for 16QAM modulation ($v_e = 240km/hr$, $|\mathbf{J}| = 8$, and $E_b/N_o=20dB$).

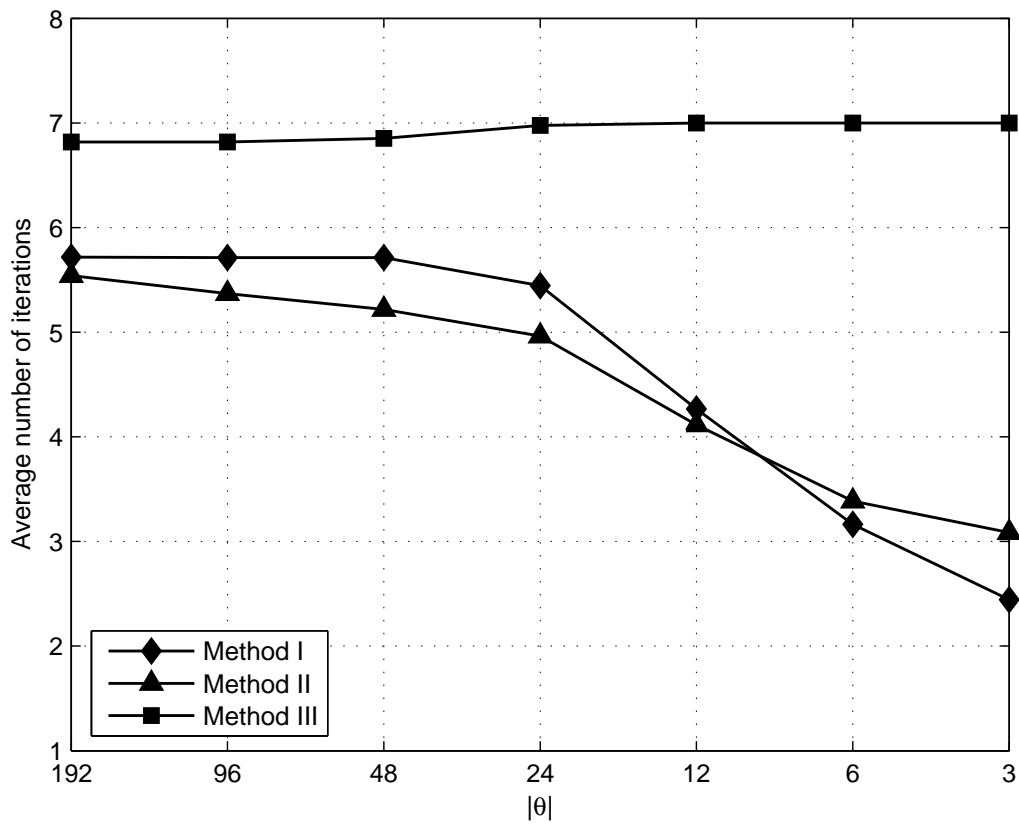


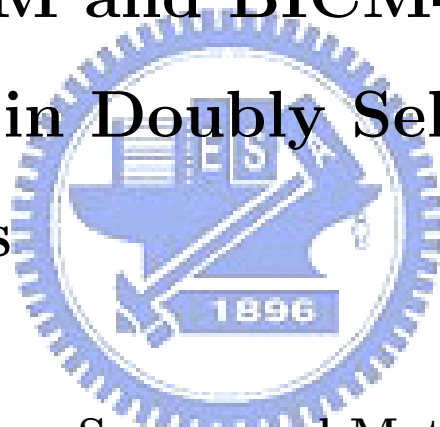
Figure 4.12: Average number of iterations versus $|\Theta|$ in the ITU Veh-A channel for 16QAM modulation ($v_e = 240km/hr$, $|\mathbf{J}| = 8$, and $E_b/N_o=20dB$).

4.6 Summary

In this chapter, we present a two-stage channel estimation method for STBC-OFDM systems in mobile wireless channels. In the initialization stage, an MPIC-based decorrelation method is used to estimate multipath delays and multipath complex gains. In the tracking stage, two refined DF DFT-based channel estimation methods are proposed by using a few pilot tones to form an optimal gradient vector at the first iteration, and the optimal step size is directly calculated from the received signals. Further, in order to reduce computational complexity of matrix inverse in the method II, an approximate weighting matrix is proposed and used in the method III. The simulation results show that both the method II and the method III can effectively alleviate the error propagation effect and thus significantly improve the performance of the method I (i.e., the classical DF DFT-based channel estimation method). The two refined methods also outperform the STBC-based MMSE method and the Kalman filtering method, especially when a high-level modulation scheme, e.g. 16QAM, is adopted in mobile environments.

Chapter 5

EM-based Iterative Receivers for OFDM and BICM-OFDM Systems in Doubly Selective Channels



5.1 Literature Survey and Motivation

OFDM is a promising technique to realize high data rate transmission over multipath fading channels. Due to the use of a GI, it allows for a simple one-tap equalizer [52]. In addition, BICM combined with OFDM, known as BICM-OFDM, is introduced as a way to offer superior performance by exploiting frequency diversity [53]. Over the past decade, OFDM has found widespread application in several standards such as 802.16e WMAN [54]. However, in mobile radio environments, multipath channels are usually time-variant. The channel time variation destroys the orthogonality among sub-

carriers, and thereby yields ICI. The effect of ICI on the BER performance has been intensively studied in [55, 56]. As the maximum Doppler frequency increases, the one-tap equalizer is no longer sufficient to conquer this channel distortion. It is shown in [56] that if the maximum Doppler frequency is larger than 8% of the subcarrier spacing, the signal-to-ICI plus noise ratio is less than $20dB$. Hence, in order to obtain reliable reception, there is a need for efficient algorithms to combat the ICI effect in a mobile OFDM receiver.

A wide variety of schemes for ICI mitigation have been proposed, mainly consisting of ICI self-cancellation, blind equalization, and ICI cancellation-based equalization [56–69]. At the expense of reduced bandwidth efficiency, the ICI self-cancellation scheme is simple and effective to provide good BER performance [57, 58]. The scheme, however, is not suitable for existing standards as modification to transmit formats is required. In contrast, the blind equalization scheme is efficient in saving bandwidth but it involves high computation complexity [59]. Among the three ICI mitigation schemes, the ICI cancellation-based equalization scheme is the most common [60–69]. Based on zero-forcing or MMSE criterion, two optimal frequency-domain equalizers are derived in [60–62]. To enhance the performance, successive interference cancellation with optimal ordering can be incorporated with the MMSE equalizer [63]. Several works, like [56] and [64–66], are targeted toward reducing the complexity of frequency-domain equalizers. By ignoring small ICI terms, a partial MMSE equalizer is proposed in [64] to avoid the inversion of a large-size matrix, while a recursive algorithm is developed in [56] for calculation of equalizer coefficients. Moreover, [65] incorporates a partial MMSE

equalizer with successive interference cancellation, and [66] combines the partial MMSE equalizer with BICM. Both methods benefit greatly from time diversity gains induced by mobility. We also find two DF equalizers in [67,68], which make use of power series expansion on time-variant frequency response. Apart from using frequency-domain equalizers, [60] and [69] consider time-domain equalizers which first achieve ICI shortening, followed by MMSE detection and parallel interference cancellation, respectively, to remove the residual ICI.

For successful implementation of the ICI cancellation-based equalization, it is essential to obtain an accurate estimate of channel variation or the equivalent ICI channel matrix. In general, this can be accomplished through the use of embedded reference signals such as pilot symbols or pilot tones. In [63], an MMSE estimator, which demands frequent pilot symbols inserted among OFDM data symbols, is proposed to estimate time-variant CIR. As complexity is concerned, most studies model the time variation of each channel tap as a polynomial function. By assuming CIR varies in a linear fashion within an OFDM symbol, [64] and [67] exploit pilot symbols for parameter estimation, whereas [68] and [70] belong to the category which uses pilot tones. It is concluded that a first-order polynomial is adequate to capture channel dynamics with the normalized maximum Doppler frequency up to 0.1. When normalized maximum Doppler frequency is larger than 0.1, a 2-D polynomial surface function is suggested in [61] to model time-varying channel frequency response and to gain better performance.

The EM algorithm can facilitate solving the ML estimation problem in

an iterative manner which alternates between an E-step, calculating an expected complete log-likelihood (ECLL) function, and an M-step, maximizing the ECLL function with respect to some unknown parameters [71, 72]. Recently, a few EM-based methods have been proposed for channel estimation and data detection in OFDM systems [73–75]. The major difference among these methods lies whether they formulate the original ML problem into a data sequence detection problem or a channel variable estimation problem. Yet, the wireless channel is assumed to be quasi-static in all these works, i.e., channel gain remains constant over the duration of one OFDM symbol. In this chapter, we investigate two EM-based iterative receivers for OFDM and BICM-OFDM systems in doubly selective fading channels. By assuming channel varies in a linear fashion, we first analyze the ICI effect in frequency domain and derive a data detection method based on the EM algorithm using the ML criterion. In an effort to reduce complexity, groupwise processing is adopted for the two EM-based receivers. For OFDM systems, we implement an ML-EM receiver which iterates between a groupwise ICI canceller and an EM detector. Based on this receiver structure, a TURBO-EM receiver for BICM-OFDM systems is then proposed to successively improve the performance by applying the turbo principle. Finally, for the initial setting of the two receivers, MMSE-based channel estimation is first performed by using a few pilot tones and it is later improved via the DF methodology.

The rest of this chapter is organized as follows. In Section 5.2, we describe the OFDM and BICM-OFDM systems, followed by the analysis of ICI in frequency domain. According to the frequency domain ICI model, an

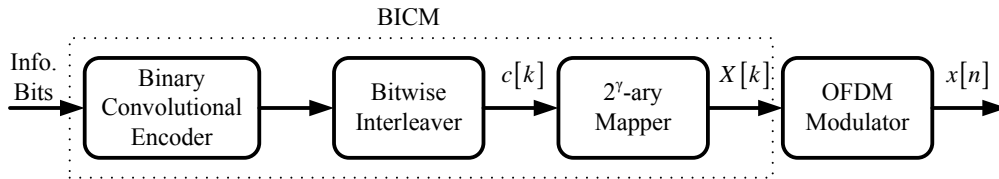


Figure 5.1: BICM-OFDM systems.

EM-based data detection method is developed in Section 5.3. In Section 5.4, an ML-EM receiver and a TURBO-EM receiver are proposed. Afterwards, we describe the initialization procedure of the two receivers and discuss their computational complexity. In Section 5.5, we present our computer simulation results. Finally, some conclusions are drawn in Section 5.6.

5.2 System Model

5.2.1 Transmitted and Received Signals

Figure 5.1 shows a BICM-OFDM system, where information bits are modulated by BICM along with an OFDM modulator [53]. Data symbols are generated by concatenating a binary convolutional encoder with a 2^γ -ary mapper through a bitwise interleaver (denoted as Π). Throughout this chapter, we only consider BPSK modulation ($\gamma = 1$); therefore, data symbols are one-to-one mapped from coded bits. Subsequently, these data symbols are transmitted over N_F consecutive OFDM symbols. Let $X[k]$ be the data symbols to be transmitted over the k th subcarrier for an OFDM symbol. After modulated by an N -point IDFT and appended with GI of length N_G , time

domain samples of an OFDM symbol are given by

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k] e^{j\frac{2\pi kn}{N}} \quad (5.1)$$

for $n = -N_G, \dots, N-1$, where we assume that $X[k]$ is mapped from the coded bit stream $c[k]$.

At the receiver, by removing the GI and taking the DFT, the demodulated signal in frequency domain is given by [19]:

$$Y[k] = H[k, k] X[k] + \underbrace{\sum_{m=0, m \neq k}^{N-1} H[k, m] X[m]}_{\text{ICI term}} + Z[k] \quad (5.2)$$

for $k = 0, \dots, N-1$, where $H[k, m]$ represents the leakage term of ICI from the m th subcarrier to the k th subcarrier, $\alpha[k, m, l]$ is the DFT of a time series $h[l, n]$ corresponding to the l th channel tap at time delay l , for $l = 0, \dots, L-1$ and $n = 0, \dots, N-1$, $((\cdot))_N$ denotes the modulo- N operation, and $Z[k]$ is AWGN with zero-mean and variance σ_Z^2 . Therefore, we have

$$H[k, m] = \sum_{l=0}^{L-1} \alpha[k, m, l] e^{-j\frac{2\pi ml}{N}} \quad (5.3)$$

and

$$\alpha[k, m, l] = \frac{1}{N} \sum_{n=0}^{N-1} h[l, n] e^{-j\frac{2\pi n((k-m))_N}{N}} \quad (5.4)$$

Moreover, we assume that the channel tap $h[l, n]$ for different l is an independent and identically distributed (i.i.d.) complex Gaussian random variable with zero mean and variance Ξ_l . From (5.2), we can observe that a demodulated subcarrier is affected by the ICI contributed from all the other subcarriers, and this effect severely degrades the system performance if a conventional one-tap equalizer is employed [56].

5.2.2 Modeling of ICI in Frequency Domain

We adopt a linear function to model the temporal variation of each channel tap over an OFDM symbol, as follows:

$$h[l, n] = a[l, 1]n + a[l, 0] \quad (5.5)$$

for $l = 0, \dots, L - 1$ and $n = 0, \dots, N - 1$, where $a[l, p]$ is the complex coefficient of the p th order for the l th tap. Substituting (5.5) into $\alpha[k, m, l]$ of (5.4), we can obtain

$$\alpha[k, m, l] = \begin{cases} \frac{N-1}{2}a[l, 1] + a[l, 0], & \text{for } k = m \\ \Phi[k, m]a[l, 1], & \text{otherwise} \end{cases} \quad (5.6)$$

where $\Phi[k, m]$ can be derived as

$$\Phi[k, m] = \frac{1}{N} \sum_{n=0}^{N-1} n e^{-j \frac{2\pi n(k-m)}{N}} = -\frac{1}{2} + j \frac{1}{2 \tan\left(\frac{\pi((k-m))_N}{N}\right)} \quad (5.7)$$

According to the fact of $1 \leq ((k-m))_N \leq N-1$, we observe that the value of $\pi((k-m))_N/N$ ranges from π/N to $\pi(N-1)/N$. Apply the Maclaurin series of $\tan(x) \approx x$, for $|x| < \pi/2$, and after some straightforward derivation, we can represent $\Phi[k, m]$ as

$$\Phi[k, m] \approx \begin{cases} -\frac{1}{2}, & \text{for } ((k-m))_N = \frac{N}{2} \\ -\frac{1}{2} + j \frac{N}{2\pi((k-m))_N}, & \text{for } 1 \leq ((k-m))_N < \frac{N}{2} \\ -\frac{1}{2} + j \frac{N}{2\pi[(k-m))_{N-N}]}, & \text{for } \frac{N}{2} < ((k-m))_N \leq N-1 \end{cases} \quad (5.8)$$

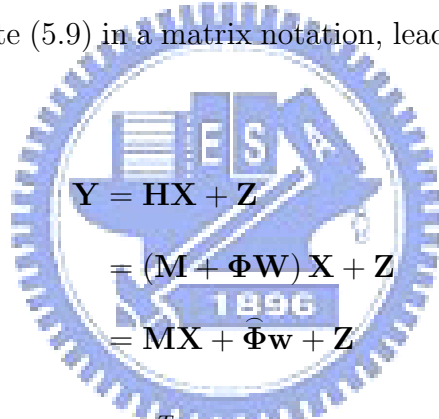
From (5.8), it follows that $\Phi[k, m]$ is a fixed value, which only depends on $(k-m)$ modulo N and it can be calculated in advance. By using (5.6) and (5.8), (5.2) can be rewritten as

$$Y[k] = H[k, k]X[k] + \underbrace{\sum_{m=0, m \neq k}^{N-1} \Phi[k, m]w[m]X[m]}_{\text{ICI term}} + Z[k] \quad (5.9)$$

where $w[m]$ is a new channel variable in frequency domain and can be defined as

$$w[m] = \sum_{l=0}^{L-1} a[l, 1] e^{-j\frac{2\pi ml}{N}} \quad (5.10)$$

It is worthy to mention that in orthogonal frequency division multiple access (OFDMA) systems, each user merely detects a set of nearby subcarriers of interest (e.g. zones or clusters in 802.16e), instead of all the N subcarriers [54]. With the formulation of (5.9), one can deal only with a small number of channel variables even when the number of channel taps is large. Finally, we can rewrite (5.9) in a matrix notation, leading to a more compact representation:



$$\begin{aligned} \mathbf{Y} &= \mathbf{H}\mathbf{X} + \mathbf{Z} \\ &= (\mathbf{M} + \mathbf{\Phi}\mathbf{W})\mathbf{X} + \mathbf{Z} \\ &= \mathbf{M}\mathbf{X} + \widehat{\mathbf{\Phi}}\mathbf{w} + \mathbf{Z} \end{aligned} \quad (5.11)$$

where $\mathbf{Y} = [Y[0], \dots, Y[N-1]]^T$, $\mathbf{X} = [X[0], \dots, X[N-1]]^T$, the (k, m) th entry of \mathbf{H} is $H[k, m]$, $\mathbf{Z} = [Z[0], \dots, Z[N-1]]^T$, the (k, m) th entry of $\mathbf{\Phi}$ is just $\Phi[k, m]$, $\mathbf{w} = [w[0], \dots, w[N-1]]^T$, $\mathbf{W} = \text{diag}\{\mathbf{w}\}$, the (k, m) th entry of $\widehat{\mathbf{\Phi}}$ is given by $\Phi[k, m]X[m]$, $\mathbf{M} = \text{diag}\{[H[0, 0], \dots, H[N-1, N-1]]^T\}$, and the superscript $(\cdot)^T$ stands for the transpose operation. Moreover, we have $\mathbf{w} = \mathbf{F}\mathbf{s}$, where $\mathbf{s} = [a[0, 1], \dots, a[L-1, 1]]^T$ and \mathbf{F} is a DFT matrix of size $N \times L$, with the (m, l) th entry given by $\exp\{-j2\pi ml/N\}$.

5.3 EM-based Data Detection Method

The EM algorithm was originally introduced by Dempster et al in 1977 for iteratively computing the ML estimate in a missing data model. [71]. The algorithm was given its name since the iteration procedure consists of an expectation step (E-step), followed by a maximization step (M-step). Dempster et al pointed out that the EM algorithm had been used in a wide range of applications, but they generalized the algorithm and more importantly, developed the theory behind it. Since 1977, due to this development, the EM algorithm has been applied in many research works and fields, such as image signal processing, communication signal processing, and even economic and social science. From (5.11), the optimum ML data detection problem can be formulated as follows:

$$\begin{aligned} \mathbf{X}_{ML} &= \arg \max_{\mathbf{X} \in \{1, -1\}^N} L(\mathbf{Y} | \mathbf{X}) \\ &= \arg \max_{\mathbf{X} \in \{1, -1\}^N} \int L(\mathbf{Y} | \mathbf{w}, \mathbf{X}) P(\mathbf{w}) d\mathbf{w} \end{aligned} \quad (5.12)$$

where $L(\cdot)$ is a log-likelihood function, obtained by taking logarithm of the corresponding probability density function $P(\cdot)$. Direct calculation using (5.12), however, involves multidimensional integration over the hidden variable \mathbf{w} . With the ability to tackle missing data models, the EM algorithm is considered as a good alternative to solve (5.12), and the core idea behind this algorithm is to iterate between the E-step and the M-step such that monotonic increase in $L(\mathbf{Y} | \mathbf{X})$ is obtained. The theorem of the EM algorithm is reviewed and proved in Appendix E. More details of the algorithm and its application can be found in [71, 72].

The E-step and the M-step associated with the optimization problem of (5.12) are expressed respectively as

$$\Omega \left(\mathbf{X} | \mathbf{Y}, \hat{\mathbf{X}}^{(m-1)} \right) = \mathbb{E}_{\mathbf{w} | \mathbf{Y}, \hat{\mathbf{X}}^{(m-1)}} [L(\mathbf{Y}, \mathbf{w} | \mathbf{X})] \quad (5.13)$$

$$\hat{\mathbf{X}}^{(m)} = \arg \max_{\mathbf{x} \in \{1, -1\}^N} \Omega \left(\mathbf{X} | \mathbf{Y}, \hat{\mathbf{X}}^{(m-1)} \right) \quad (5.14)$$

where $\hat{\mathbf{X}}^{(m)}$ denotes the hard decision of \mathbf{X} at the m th EM iteration, and $\Omega \left(\mathbf{X} | \mathbf{Y}, \hat{\mathbf{X}}^{(m-1)} \right)$ is known as the ECLL function, to be maximized in the M-step of (5.14). By using the fact that $L(\mathbf{Y}, \mathbf{w} | \mathbf{X}) = L(\mathbf{Y} | \mathbf{w}, \mathbf{X}) + L(\mathbf{w})$ and from (5.11), we can further simplify (5.13) as

$$\begin{aligned} \Omega \left(\mathbf{X} | \mathbf{Y}, \hat{\mathbf{X}}^{(m-1)} \right) &= \mathbb{E}_{\mathbf{w} | \mathbf{Y}, \hat{\mathbf{X}}^{(m-1)}} [L(\mathbf{Y} | \mathbf{w}, \mathbf{X})] + \text{const.} \\ &= \mathbb{E}_{\mathbf{w} | \mathbf{Y}, \hat{\mathbf{X}}^{(m-1)}} \left[\frac{-1}{\sigma_Z^2} \|\mathbf{Y} - \mathbf{H}\mathbf{X}\|^2 \right] + \text{const.} \\ &= \frac{-1}{\sigma_Z^2} \left(\mathbf{Y}^H \mathbf{Y} - \mathbf{Y}^H \tilde{\mathbf{H}} \mathbf{X} \right. \\ &\quad \left. - \mathbf{X}^H \tilde{\mathbf{H}}^H \mathbf{Y} + \mathbf{X}^H \tilde{\Sigma} \mathbf{X} \right) + \text{const.} \end{aligned} \quad (5.15)$$

where $\tilde{\mathbf{H}}$ and $\tilde{\Sigma}$ denote $\mathbb{E}_{\mathbf{w} | \mathbf{Y}, \hat{\mathbf{X}}^{(m-1)}} [\mathbf{H}]$ and $\mathbb{E}_{\mathbf{w} | \mathbf{Y}, \hat{\mathbf{X}}^{(m-1)}} [\mathbf{H}^H \mathbf{H}]$, respectively, and the superscript $(\cdot)^H$ represents the Hermitian operation. The constant term in (5.15) can be dropped for simplicity. Without loss of generality, the CSI \mathbf{M} can be estimated through pilot tones embedded in each OFDM symbol, and we denote the estimate as $\hat{\mathbf{M}}$. By inserting $\mathbf{H} = \hat{\mathbf{M}} + \Phi \mathbf{W}$ into $\tilde{\mathbf{H}}$ and $\tilde{\Sigma}$, it is straightforward to calculate the two terms as

$$\tilde{\mathbf{H}} = \mathbb{E}_{\mathbf{w} | \mathbf{Y}, \hat{\mathbf{X}}^{(m-1)}} \left[\hat{\mathbf{M}} + \Phi \mathbf{W} \right] = \hat{\mathbf{M}} + \Phi \tilde{\mathbf{W}} \quad (5.16)$$

$$\begin{aligned}
\tilde{\Sigma} &= \mathbb{E}_{\mathbf{w}|\mathbf{Y}, \hat{\mathbf{x}}^{(m-1)}} \left[\left(\hat{\mathbf{M}} + \Phi \mathbf{W} \right)^H \left(\hat{\mathbf{M}} + \Phi \mathbf{W} \right) \right] \\
&= \mathbb{E}_{\mathbf{w}|\mathbf{Y}, \hat{\mathbf{x}}^{(m-1)}} \left[\hat{\mathbf{M}}^H \hat{\mathbf{M}} + \hat{\mathbf{M}}^H \Phi \mathbf{W} + \mathbf{W}^H \Phi^H \hat{\mathbf{M}} \right. \\
&\quad \left. + \mathbf{W}^H \Phi^H \Phi \mathbf{W} \right] \\
&= \hat{\mathbf{M}}^H \hat{\mathbf{M}} + \hat{\mathbf{M}}^H \Phi \tilde{\mathbf{W}} + \tilde{\mathbf{W}}^H \Phi^H \hat{\mathbf{M}} \\
&\quad + \mathbb{E}_{\mathbf{w}|\mathbf{Y}, \hat{\mathbf{x}}^{(m-1)}} \left[(\mathbf{w} \mathbf{w}^H)^T \odot (\Phi^H \Phi) \right] \\
&= \hat{\mathbf{M}}^H \hat{\mathbf{M}} + \hat{\mathbf{M}}^H \Phi \tilde{\mathbf{W}} + \tilde{\mathbf{W}}^H \Phi^H \hat{\mathbf{M}} \\
&\quad + \mathbb{E}_{\mathbf{w}|\mathbf{Y}, \hat{\mathbf{x}}^{(m-1)}} \left[(\mathbf{w} \mathbf{w}^H)^T \right] \odot (\Phi^H \Phi) \\
&= \hat{\mathbf{M}}^H \hat{\mathbf{M}} + \hat{\mathbf{M}}^H \Phi \tilde{\mathbf{W}} + \tilde{\mathbf{W}}^H \Phi^H \hat{\mathbf{M}} \\
&\quad + \left(\mathbb{E}_{\mathbf{w}|\mathbf{Y}, \hat{\mathbf{x}}^{(m-1)}} [\mathbf{w} \mathbf{w}^H] \right)^T \odot (\Phi^H \Phi) \tag{5.17}
\end{aligned}$$

where we have $\tilde{\mathbf{W}} \triangleq \mathbb{E}_{\mathbf{w}|\mathbf{Y}, \hat{\mathbf{x}}^{(m-1)}} [\mathbf{W}]$ and the notation \odot denotes as the Hadamard product. Let

$$\tilde{\mathbf{W}} \triangleq \text{diag} \{ \tilde{\mathbf{w}} \} \tag{5.18}$$

$$\tilde{\mathbf{w}} \triangleq \mathbb{E}_{\mathbf{w}|\mathbf{Y}, \hat{\mathbf{x}}^{(m-1)}} [\mathbf{w}] \tag{5.19}$$

and

$$\begin{aligned}
\tilde{\Sigma}_{\mathbf{w}} &\triangleq \mathbb{E}_{\mathbf{w}|\mathbf{Y}, \hat{\mathbf{x}}^{(m-1)}} \left[(\mathbf{w} - \tilde{\mathbf{w}}) (\mathbf{w} - \tilde{\mathbf{w}})^H \right] \\
&= \mathbb{E}_{\mathbf{w}|\mathbf{Y}, \hat{\mathbf{x}}^{(m-1)}} [\mathbf{w} \mathbf{w}^H] - \tilde{\mathbf{w}} \tilde{\mathbf{w}}^H \tag{5.20}
\end{aligned}$$

We can rewrite (5.17) as

$$\begin{aligned}
\tilde{\Sigma} &= \hat{\mathbf{M}}^H \hat{\mathbf{M}} + \hat{\mathbf{M}}^H \Phi \tilde{\mathbf{W}} + \tilde{\mathbf{W}}^H \Phi^H \hat{\mathbf{M}} \\
&\quad + \left(\tilde{\Sigma}_{\mathbf{w}} + \tilde{\mathbf{w}} \tilde{\mathbf{w}}^H \right)^T \odot (\Phi^H \Phi) \tag{5.21}
\end{aligned}$$

Also, from (5.11), it is observed that the conditional probability density function $P(\mathbf{w} | \mathbf{Y}, \hat{\mathbf{X}}^{(m-1)})$ is a Gaussian distribution, with mean and covariance given by [76]

$$\tilde{\mathbf{w}} = \boldsymbol{\mu}_{\mathbf{w}} + \mathbf{C}_{\mathbf{wY}} \mathbf{C}_{\mathbf{YY}}^{-1} (\mathbf{Y} - \boldsymbol{\mu}_{\mathbf{Y}}) \quad (5.22)$$

$$\tilde{\boldsymbol{\Sigma}}_{\mathbf{w}} = \mathbf{C}_{\mathbf{ww}} - \mathbf{C}_{\mathbf{wY}} \mathbf{C}_{\mathbf{YY}}^{-1} \mathbf{C}_{\mathbf{Yw}} \quad (5.23)$$

where the relevant terms are defined and statistics are calculated in the following way. We first apply a first-order autoregressive (AR) channel model to compute $\boldsymbol{\mu}_{\mathbf{w}}$ and $\mathbf{C}_{\mathbf{ww}}$. Details are provided in Appendix F, and the two terms can be derived as

$$\boldsymbol{\mu}_{\mathbf{w}} = \mathbb{E}[\mathbf{w}] = \mathbf{0} \quad (5.24)$$

$$\mathbf{C}_{\mathbf{ww}} = \mathbb{E}[(\mathbf{w} - \boldsymbol{\mu}_{\mathbf{w}})(\mathbf{w} - \boldsymbol{\mu}_{\mathbf{w}})^H] = \mathbf{F} \mathbf{C}_{\text{ss}} \mathbf{F}^H \quad (5.25)$$

where \mathbf{C}_{ss} in (5.25) is a diagonal matrix with the l th diagonal entry equal to $2(1 - \alpha) \Xi_l / (N - 1)^2$, and α is the channel tap autocorrelation as defined in (F.2). Moreover, we can get

$$\boldsymbol{\mu}_{\mathbf{Y}} = \mathbb{E}[\mathbf{Y}] = \hat{\mathbf{M}} \hat{\mathbf{X}}^{(m-1)} \quad (5.26)$$

$$\begin{aligned} \mathbf{C}_{\mathbf{YY}} &= \mathbb{E}[(\mathbf{Y} - \boldsymbol{\mu}_{\mathbf{Y}})(\mathbf{Y} - \boldsymbol{\mu}_{\mathbf{Y}})^H] \\ &= \hat{\boldsymbol{\Phi}}^{(m-1)} \mathbf{C}_{\mathbf{ww}} \hat{\boldsymbol{\Phi}}^{(m-1)H} + \sigma_Z^2 \mathbf{I} \end{aligned} \quad (5.27)$$

and

$$\mathbf{C}_{\mathbf{w}\mathbf{Y}} = \mathbf{C}_{\mathbf{Y}\mathbf{w}}^H = \mathbb{E} \left[(\mathbf{w} - \boldsymbol{\mu}_{\mathbf{w}}) (\mathbf{Y} - \boldsymbol{\mu}_{\mathbf{Y}})^H \right] = \mathbf{C}_{\mathbf{w}\mathbf{w}} \widehat{\boldsymbol{\Phi}}^{(m-1)H} \quad (5.28)$$

where $\widehat{\boldsymbol{\Phi}}^{(m-1)}$ is obtained by substituting the hard decision $\widehat{\mathbf{X}}^{(m-1)}$ into $\widehat{\boldsymbol{\Phi}}$. Using (5.16)-(5.28), we can calculate the ECLL function of (5.15). The EM algorithm for data detection is then summarized in Figure 5.2, and it is repeated until a stopping criterion holds. The stopping criterion is to check whether $\widehat{\mathbf{X}}^{(m)} = \widehat{\mathbf{X}}^{(m-1)}$ or the iteration number reaches a predefined limit N_{EM} .

5.4 Implementation: EM-based Iterative Receivers

Because the data detection method in (5.14) has high computation complexity, in this section, we investigate two EM-based iterative receivers for practical implementation.

5.4.1 ML-EM Receiver for OFDM Systems

As depicted in Figure 5.3, we consider an ML-EM receiver with N subcarriers partitioned into R groups, and each group consists of G subcarriers. Denote the j th group of subcarriers as $\mathbf{G}_j = \{jG, \dots, (j+1)G - 1\}$, for $j = 0, \dots, R - 1$. Next, we define the j th data group and observation group as

$$\mathbf{X}_j = [X[jG], \dots, X[(j+1)G - 1]]^T \quad (5.29)$$

Initialization:

Calculate $\hat{\mathbf{M}}$, choose $\hat{\mathbf{X}}^{(0)}$, and set $m = 0$

Execution of EM algorithm:

do {

$m = m + 1$

E-step:

Compute statistics

$$\mathbf{C}_{\mathbf{w}\mathbf{w}} = \mathbf{F}\mathbf{C}_{\mathbf{s}\mathbf{s}}\mathbf{F}^H$$

$$\tilde{\mathbf{w}} = \mathbf{C}_{\mathbf{w}\mathbf{w}} \hat{\Phi}^{(m-1)H} \left(\hat{\Phi}^{(m-1)} \mathbf{C}_{\mathbf{w}\mathbf{w}} \hat{\Phi}^{(m-1)H} + \sigma_Z^2 \mathbf{I} \right)^{-1} \left(\mathbf{Y} - \hat{\mathbf{M}} \hat{\mathbf{X}}^{(m-1)} \right)$$

$$\tilde{\Sigma}_{\mathbf{w}} = \mathbf{C}_{\mathbf{w}\mathbf{w}} - \mathbf{C}_{\mathbf{w}\mathbf{w}} \hat{\Phi}^{(m-1)H} \left(\hat{\Phi}^{(m-1)} \mathbf{C}_{\mathbf{w}\mathbf{w}} \hat{\Phi}^{(m-1)H} + \sigma_Z^2 \mathbf{I} \right)^{-1} \hat{\Phi}^{(m-1)} \mathbf{C}_{\mathbf{w}\mathbf{w}}$$

$$\tilde{\mathbf{H}} = \hat{\mathbf{M}} + \Phi \tilde{\mathbf{w}}$$

$$\tilde{\Sigma} = \hat{\mathbf{M}}^H \hat{\mathbf{M}} + \hat{\mathbf{M}}^H \Phi \tilde{\mathbf{w}} + \tilde{\mathbf{w}}^H \Phi^H \hat{\mathbf{M}} + \left(\tilde{\Sigma}_{\mathbf{w}} + \tilde{\mathbf{w}} \tilde{\mathbf{w}}^H \right)^T \odot (\Phi^H \Phi)$$

For all $\mathbf{X} \in \{1, -1\}^N$, calculate

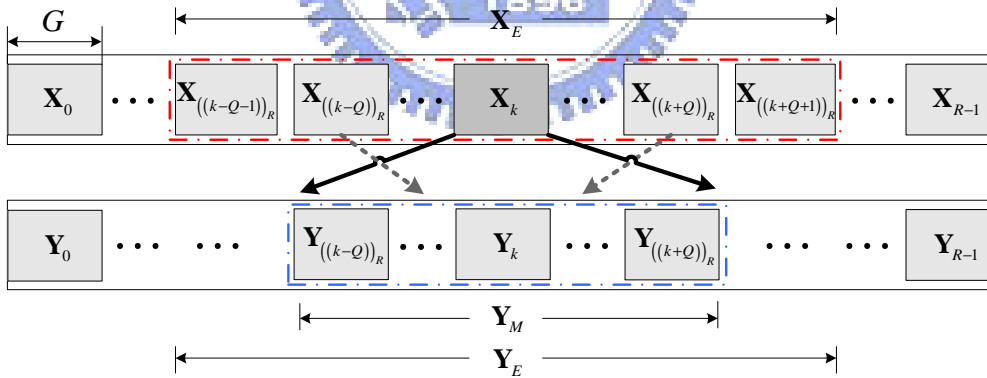
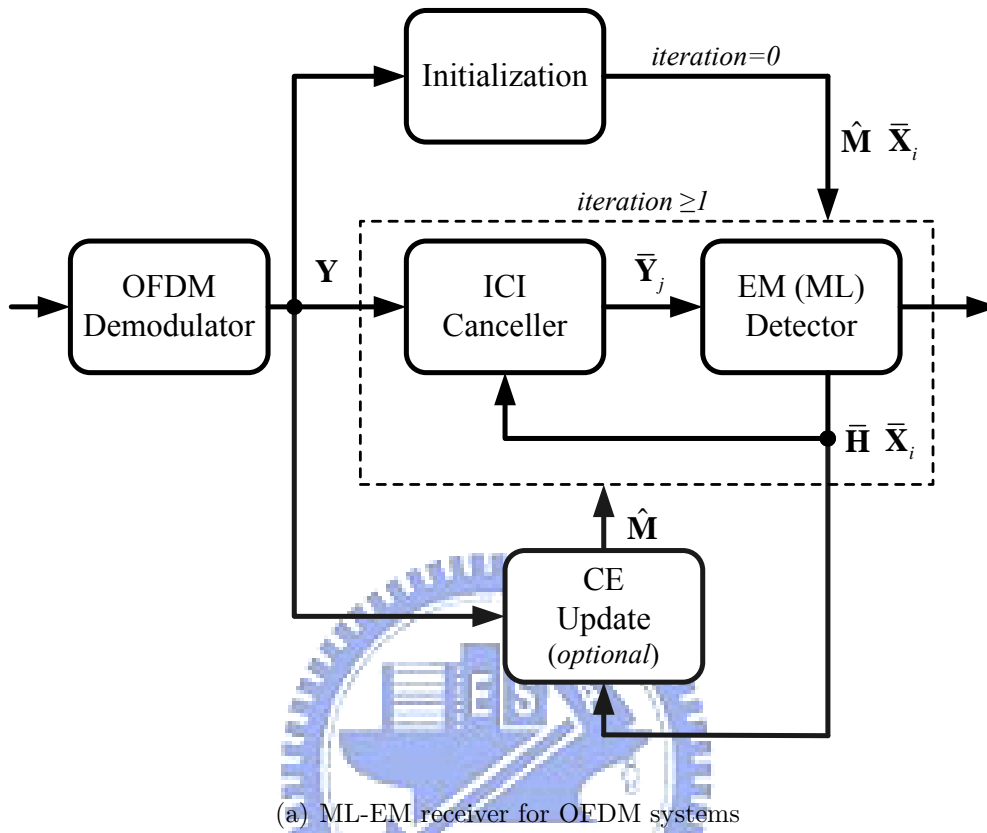
$$\Omega(\mathbf{X} | \mathbf{Y}, \hat{\mathbf{X}}^{(m-1)}) = -\frac{1}{\sigma_Z^2} \left\{ \mathbf{Y}^H \mathbf{Y} - \mathbf{Y}^H \tilde{\mathbf{H}} \mathbf{X} - \mathbf{X}^H \tilde{\mathbf{H}}^H \mathbf{Y} + \mathbf{X}^H \tilde{\Sigma} \mathbf{X} \right\}$$

M-step:

$$\hat{\mathbf{X}}^{(m)} = \arg \max_{\mathbf{X} \in \{1, -1\}^N} \Omega(\mathbf{X} | \mathbf{Y}, \hat{\mathbf{X}}^{(m-1)})$$

} while $(\hat{\mathbf{X}}^{(m)} \neq \hat{\mathbf{X}}^{(m-1)})$ and $m \leq N_{EM}$

Figure 5.2: EM-based data detection method.



1. Calculate statistics of \mathbf{w} , by using \mathbf{X}_E and \mathbf{Y}_E .
2. Calculate ECLL function, by using \mathbf{Y}_M .

(b) An illustration for group detection

Figure 5.3: (a) ML-EM receiver for OFDM systems (b) An illustration for group detection.

and

$$\mathbf{Y}_j = [Y [jG], \dots, Y [(j+1)G - 1]]^T \quad (5.30)$$

Further, we use \mathbf{B}_j to denote the set $\{((j-Q))_R, \dots, ((j+Q))_R\}$. Without loss of generality, we focus on detecting the k th data group. Assume that due to the ICI effect, the energy of \mathbf{X}_k is spread over $2Q + 1$ observation groups of \mathbf{Y}_j , for $j \in \mathbf{B}_k$, which also contains interfering energy caused by other adjacent data groups \mathbf{X}_j , for $j \in \mathbf{B}_j \setminus \{k\}$. As observed in Figure 5.3(a), there is an additional iteration loop outside the EM detector, called ML iteration. Within an ML iteration, the ICI is first reconstructed and subtracted from the observation group, yielding a signal:

$$\bar{\mathbf{Y}}_j = \mathbf{Y}_j - \sum_{i \in \mathbf{B}_k \setminus \{k\}} \bar{\mathbf{H}}_{j,i} \bar{\mathbf{X}}_i \quad (5.31)$$

for $j \in \mathbf{B}_k$, where $\bar{\mathbf{X}}_i$ is the tentative decision of \mathbf{X}_i , and the (p, q) th entry of $\bar{\mathbf{H}}_{j,i}$ is given by the $(jG + p, iG + q)$ th entry of $\bar{\mathbf{H}}$, the estimate of \mathbf{H} , for $p, q = 0, \dots, G - 1$. Both $\bar{\mathbf{X}}_i$ and $\bar{\mathbf{H}}$ are obtained from the output of the EM detector at the previous ML iteration. After ICI cancellation, the EM detector is executed by applying the EM-based data detection method in Section 5.3. Define $\mathbf{X}_E = [\bar{\mathbf{X}}_{((k-Q-1))_R}^T, \dots, \hat{\mathbf{X}}_k^{(m-1)T}, \dots, \bar{\mathbf{X}}_{((k+Q+1))_R}^T]^T$ and $\mathbf{Y}_E = [\mathbf{Y}_{((k-Q-1))_R}^T, \dots, \mathbf{Y}_{((k+Q+1))_R}^T]^T$, where $\hat{\mathbf{X}}_k^{(m-1)}$ is the hard decision of \mathbf{X}_k at the $(m-1)$ th EM iteration within the EM detector. Particularly, for $m = 1$, we initialize $\hat{\mathbf{X}}_k^{(0)}$ as $\bar{\mathbf{X}}_k$. In the E-step, at the m th EM iteration, we replace \mathbf{X} and \mathbf{Y} (in Figure 5.2) with \mathbf{X}_E and \mathbf{Y}_E to calculate the statistics $\tilde{\mathbf{W}}$, $\tilde{\mathbf{H}}$ and $\tilde{\Sigma}$. The size of these three matrices now becomes $(2Q + 3)G \times (2Q + 3)G$. After that, the interference-reduced signal

$\mathbf{Y}_M = \left[\bar{\mathbf{Y}}_{((k-Q))_R}^T, \dots, \bar{\mathbf{Y}}_{((k+Q))_R}^T \right]^T$ is taken to compute the ECLL function, for each combination of $\mathbf{X}_k \in \{1, -1\}^G$, as follows:

$$\Omega(\mathbf{X}_k | \mathbf{Y}_M, \mathbf{Y}_E, \mathbf{X}_E) = -\frac{1}{\sigma_Z^2} \left(\mathbf{Y}_M^H \mathbf{Y}_M - \mathbf{Y}_M^H \tilde{\mathbf{H}}_k \mathbf{X}_k - \mathbf{X}_k^H \tilde{\mathbf{H}}_k^H \mathbf{Y}_M + \mathbf{X}_k^H \tilde{\Sigma}_k \mathbf{X}_k \right) \quad (5.32)$$

where the matrices $\tilde{\mathbf{H}}_k$ and $\tilde{\Sigma}_k$ are of size $(2Q+1)G \times G$ and $G \times G$, with the (p, q) th entry given by the $(G+p, (Q+1)G+q)$ th entry of $\tilde{\mathbf{H}}$ and the $((Q+1)G+p, (Q+1)G+q)$ th entry of $\tilde{\Sigma}$, respectively. Finally, the decision of \mathbf{X}_k is calculated in the M-step according to:

$$\hat{\mathbf{X}}_k^{(m)} = \arg \max_{\mathbf{X}_k \in \{1, -1\}^G} \Omega(\mathbf{X}_k | \mathbf{Y}_M, \mathbf{Y}_E, \mathbf{X}_E) \quad (5.33)$$

Within the EM detector, the above procedure is conducted to detect R groups simultaneously, i.e., we use parallel processing for group detection. Once the stopping criterion is met, the receiver proceeds to the next ML iteration until a good performance is achieved, and $\bar{\mathbf{X}}_k$ and $\bar{\mathbf{H}}$ are updated. In other words, at the end of the k th parallel processing, $\bar{\mathbf{X}}_k$ is replaced by $\hat{\mathbf{X}}_k^{(m)}$, the $(kG+j)$ th diagonal entry of $\bar{\mathbf{W}}$ is renewed by the $((Q+1)G+j)$ th diagonal entry of $\tilde{\mathbf{W}}$, for $k = 0, \dots, R-1$ and $j = 0, \dots, G-1$, and $\bar{\mathbf{H}}$ is calculated as $\hat{\mathbf{M}} + \Phi \bar{\mathbf{W}}$.

The intuition behind the group detection is explained as follows. While computing and maximizing the ECLL function, we can acquire the diversity gains through examining the interference-reduced signals \mathbf{Y}_M , of which the energy is contributed mainly by the data group \mathbf{X}_k . Therefore, it is reasonable to expect that full diversity gain is achievable when the value of Q is sufficiently large and the ICI is perfectly cancelled out. The diversity gain we

5.4.2 TURBO-EM Receiver for BICM-OFDM Systems

Figure 5.4 shows the TURBO-EM receiver for BICM-OFDM systems. The receiver implements the turbo iterations by exchanging the extrinsic information between the EM detector (after the ICI canceller) and the soft-output Viterbi algorithm (SOVA) decoder. Within each TURBO iteration, the ICI is first reconstructed and subtracted from the observation group to obtain the interference-reduced signal $\bar{\mathbf{Y}}_j$ by using (5.31), but with soft decision $\hat{\mathbf{X}}_i$ replacing the hard decision $\bar{\mathbf{X}}_i$. In this way, we can mitigate the error propagation effect, and the soft decision for the BPSK case is given by [77]

$$\hat{\mathbf{X}}_i = \mathbb{E}[\mathbf{X}_i] = \tanh\left(\frac{\boldsymbol{\lambda}_i^{C,post}}{2}\right) \quad (5.34)$$

where $\boldsymbol{\lambda}_i^{C,post} = [\lambda^{C,post}[iG], \dots, \lambda^{C,post}[(i+1)G-1]]^T$ is the *a posteriori* log-likelihood ratio (LLR), associated with \mathbf{X}_i , from the SOVA decoder at the previous TURBO iteration, and the LLR of a data symbol ϑ is defined as the logarithm of the ratio of $P(\vartheta = +1)$ to $P(\vartheta = -1)$. We then apply the maximum *a posteriori* (MAP) EM algorithm to the EM detector, which further takes account of the *a priori* information to compute the ECLL function as follows:

$$\Omega(\mathbf{X}_k | \mathbf{Y}_M, \mathbf{Y}_E, \mathbf{X}_E) = \frac{-1}{\sigma_Z^2} \left(\mathbf{Y}_M^H \mathbf{Y}_M - \mathbf{Y}_M^H \tilde{\mathbf{H}}_k \mathbf{X}_k - \mathbf{X}_k^H \tilde{\mathbf{H}}_k^H \mathbf{Y}_M + \mathbf{X}_k^H \tilde{\boldsymbol{\Sigma}}_k \mathbf{X}_k \right) + L(\mathbf{X}_k) \quad (5.35)$$

where $L(\mathbf{X}_k) = \ln P(\mathbf{X}_k)$ is calculated from the interleaved extrinsic information $\boldsymbol{\lambda}_k^{C,ext} = [\lambda^{C,ext}[kG], \dots, \lambda^{C,ext}[(k+1)G-1]]^T$ with respect to \mathbf{X}_k which is generated by the SOVA decoder. Under the assumption of an ideal

interleaver, data symbols are independent of each other, and we obtain

$$L(\mathbf{X}_k) = \sum_{j=0}^{G-1} L(X[kG+j] = q_j) \quad (5.36)$$

where q_j denotes the value of $X[kG+j]$, and $L(X[kG+j] = q_j)$ is calculated by using (G.4) in Appendix G. At the final EM iteration, the extrinsic *a posteriori* LLR $\boldsymbol{\lambda}_k^{D,ext} = [\lambda^{D,ext}[kG], \dots, \lambda^{D,ext}[(k+1)G-1]]^T$ is generated at the output of the EM detector. From [78] and (G.3), we get

$$\begin{aligned} & \lambda^{D,ext}[kG+j] \\ &= \ln \frac{P(X[kG+j] = +1 | \mathbf{Y}_M)}{P(X[kG+j] = -1 | \mathbf{Y}_M)} - \lambda^{C,ext}[kG+j] \\ &\approx \max_{\mathbf{x}_k \in \Omega_j^+} \left(\frac{-1}{\sigma_Z^2} \|\mathbf{Y}_M - \tilde{\mathbf{H}}_k \mathbf{X}_k\|^2 + \frac{1}{2} \mathbf{X}_{k \setminus \{j\}}^T \boldsymbol{\lambda}_{k \setminus \{j\}}^{C,ext} \right) \\ &\quad - \max_{\mathbf{x}_k \in \Omega_j^-} \left(\frac{-1}{\sigma_Z^2} \|\mathbf{Y}_M - \tilde{\mathbf{H}}_k \mathbf{X}_k\|^2 + \frac{1}{2} \mathbf{X}_{k \setminus \{j\}}^T \boldsymbol{\lambda}_{k \setminus \{j\}}^{C,ext} \right) \end{aligned} \quad (5.37)$$

where Ω_j^+ denotes the set for which the j th entry of \mathbf{X}_k is "+1"; Ω_j^- is defined similarly, and the vectors $\mathbf{X}_{k \setminus \{j\}}$ and $\boldsymbol{\lambda}_{k \setminus \{j\}}^{C,ext}$ are obtained by omitting the j th entry of \mathbf{X}_k and $\boldsymbol{\lambda}_k^{C,ext}$, respectively. The extrinsic LLR $\boldsymbol{\lambda}_k^{D,ext}$ is then converted into soft bits using (5.34), modeled as the output of an AWGN channel with unit gain and variance σ_C^2 , deinterleaved through $\mathbf{\Pi}^{-1}$, and passed to the SOVA decoder. The variance σ_C^2 is estimated as [79]

$$\sigma_C^2 = \frac{1}{N_I} \sum_{i=1}^{N_I} (|\mu[i]| - 1)^2 \quad (5.38)$$

where $\mu[i]$ indicates the soft value of the coded bits ranging between -1 and $+1$, and N_I represents the interleaver size. It is mentioned in [79] that the Gaussian assumption is not satisfied at the beginning of TURBO iterations, but it becomes a good approximation as the number of iterations increases.

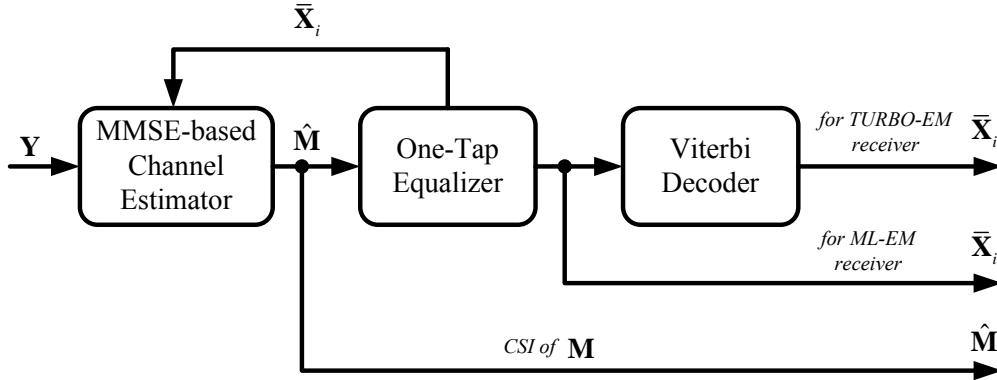


Figure 5.5: Initialization procedure for ML-EM and TURBO-EM receivers.

After the SOVA decoder produces the soft information by considering the ML path and its strongest competitor in the trellis diagram, the receiver progresses toward the next TURBO iteration until a preset maximum number of iteration, N_{TB} , is reached.

5.4.3 Initial Setting and Channel Estimation Update

Figure 5.5 depicts the block diagram for initialization of the two receivers. The initial channel estimation is performed through the use of pilot tones and improved via the decided data symbols. Let \mathbf{X}_P be a diagonal matrix whose diagonal elements are obtained from the stacked vector of J pilot tones on subcarriers $\{P_0, \dots, P_{J-1}\}$ within the OFDM symbol. Applying the MMSE-based channel estimation method, we obtain [39]

$$\hat{\mathbf{M}} = \mathbf{F} (\mathbf{F}_P^H \mathbf{X}_P^H \mathbf{X}_P \mathbf{F}_P + (\sigma_Z^2 + \sigma_{ICI}^2) \mathbf{I})^{-1} \mathbf{F}_P^H \mathbf{X}_P^H \mathbf{Y}_P \quad (5.39)$$

where \mathbf{Y}_P and \mathbf{F}_P are defined similar to \mathbf{Y} and \mathbf{F} , respectively, but here related to subcarriers $\{P_0, \dots, P_{J-1}\}$ only. By invoking central limit theorem,

the ICI energy σ_{ICI}^2 is approximated as $(2\pi f_d)^2/12$ [55]. Subsequently, an one-tap equalizer is used for data detection, and a DF approach is carried out to initialize the two receivers. For the ML-EM receiver, decided data symbols together with the pilot tones are used to generate a new channel estimate $\hat{\mathbf{M}}$ by using (5.39) and then produce an updated decision symbol $\bar{\mathbf{X}}_i$, while for the TURBO-EM receiver, much more reliable decision symbols are generated by the Viterbi decoder. At the first TURBO iteration of the TURBO-EM receiver, $\boldsymbol{\lambda}_k^{C,ext}$ is set to $\mathbf{0}$, and $\bar{\mathbf{X}}_i$ is used to replace $\hat{\mathbf{X}}_i$ in (5.34). Moreover, we initialize $\bar{\mathbf{H}}$ as $\hat{\mathbf{M}}$, i.e., set $\bar{\mathbf{W}} = \mathbf{O}$, since no information on $\bar{\mathbf{W}}$ is available at the first iteration of the two receivers.

Due to ICI, the initial estimate of \mathbf{M} becomes inaccurate as f_d increases. Hence, Figure 5.3(a) and Figure 5.4 offer an option for channel estimation update, wherein at the second and subsequent iterations (outer loop), the ICI (reconstructed from N_U adjacent subcarriers in a hard or soft manner) to the subcarriers is canceled out in the received signal \mathbf{Y} , and the MMSE-based channel estimation is again used to refine the estimate $\hat{\mathbf{M}}$ by setting $\sigma_{ICI}^2 = 0$.

5.4.4 Computational Complexity

Now let us look at the number of complex multiplications required for the two proposed receivers. Assume that the operation of $K \times K$ matrix inversion needs K^3 complex multiplications. In Table 5.1, the first and second rows indicate the complexity of the ICI canceller and the EM detector, respectively. The third row gives the complexity for precomputing $\mathbf{C}_{\mathbf{w}\mathbf{w}}$ and $\boldsymbol{\Phi}^H \boldsymbol{\Phi}$ in the

EM detector. Note that the calculation of (5.36) in the MAP EM detector does not require any multiplications, and the number of multiplications required to calculate (5.37) is presented in the fourth row of the Table 5.1. For example, in the case of $G = 4$, $Q = 4$, $\gamma = 1$ and $L = 6$, the complexity is calculated in the third column as well. Some complexity reduction can be achieved by applying the SAGE algorithm and the Viterbi algorithm, as proposed in [80] and [77] respectively, to simplify (5.33) and (5.37) when the values of G and γ are relatively large to dominate the overall computation complexity. Moreover, the complexity of the SOVA decoder is, in general, upper-bounded by two times that of the Viterbi decoder. Finally, the computational complexity of the MMSE-based channel estimation method can be referred from [39] for details.

5.5 Computer Simulation

Our simulation demonstrates the performance of the two EM-based receivers. The simulation parameters are defined according to the 802.16e OFDM standard [54] and listed in Table 5.2. The entire bandwidth, $5MHz$, is divided into $N = 256$ subcarriers among which 192 subcarriers carry data symbols, $J = 8$ subcarriers transmit pilot tones, and the remaining 56 subcarriers are virtual subcarriers. The BPSK modulation scheme is adopted for the pilot tones, and a pilot subcarrier transmits at the same power level as a data subcarrier. Each OFDM data frame is composed of $N_F = 40$ OFDM data symbols, and the length of GI is set to $N_G = 64$. For the BICM scheme, we employ a rate-1/2 convolutional code with generator polynomial

Table 5.1: Computational Complexity (Ex: $G = 4$, $Q = 4$, $\gamma = 1$, and $L = 6$).

Operation	Number of Complex Multiplications	Example	Unit
ICI canceller	$(2Q) G^2 (2Q + 1)$	1152	/ML (or TURBO) iteration /group
EM detector	$5(2Q + 3)^3 G^3$ + $[36Q^2 + (108 + 2^{\gamma G+1}) Q$ + $(81 + 2^{\gamma G+1})] G^2$ + $[(2^{\gamma G+2} + 4) Q + 3 \times 2^{\gamma G}$ + $6] G$	447208	/ML (or TURBO) iteration /EM iteration /group
Precalculation of $\mathbf{C}_{\mathbf{w}\mathbf{w}}$ & $\mathbf{\Phi}^H \mathbf{\Phi}$	$(2Q + 3)^2 (L + 1) G^2$ + $(2Q + 3) LG$	13816	/group
Eq. (5.37)	$2^{\gamma(G-1)} [(2Q + 1) G^2$ + $(2Q + 2) G + 1]$	1480	/TURBO iteration /group

(133, 171) represented in octal and a block interleaver with 96 rows and 80 columns. Both a conventional two-path channel and an ITU Veh-A channel are used in our simulation with path delays uniformly distributed from 0 to 50 sample periods, where the relative path power profiles are set as 0, 0 (dB) for the two-path channel and 0, -1, -9, -10, -15, -20 (dB) for the ITU Veh-A channel [34]. The fading channel is generated with Jakes model by setting $f_d = 0.1$ [33]. The user-defined parameters are chosen as $N_U = 10$ and $N_{EM} = 5$. Some statistical information such as power delay profiles, Doppler frequency, and noise power is assumed to be known to the receivers. Throughout the simulation, the parameter E_b/N_o is defined as a ratio of averaged receive bit energy to the power spectral density of noise. As a benchmark, the performance curve with ideal initialization, labeled as CSI and data known, serves as a performance lower bound, and the results obtained with ideal CSI, denoted as CSI known, is provided for reference. Furthermore, we also include the performance curve of the one-tap equalizer in quasi-static channels, under the assumption of ideal CSI.

5.5.1 BER Performance of ML-EM Receiver

Figure 5.6 and Figure 5.7 show the BER performance of the ML-EM receiver in the two-path and the ITU Veh-A channel, respectively. The parameter of $[G, Q]$ is set to $[4, 4]$. It is seen from Figure 5.6 that after three iterations, the ML-EM receivers with or without channel estimation update have much better performance than the same receiver at the initialization stage. Since time-variant channels introduce diversity gains, the ML-EM receiver also per-

Table 5.2: Simulation parameters.

Parameter	Value
Carrier frequency	2.3GHz
Bandwidth	5MHz
FFT size (N)	256
Length of CP (N_G)	64
Number of data subcarriers	192
Number of pilot subcarriers (J)	8
Number of virtual subcarriers	56
Modulation	BPSK
Number of OFDM data symbols per frame (N_F)	40
Generator polynomial of convolutional code	(133, 171) in octal
Block interleaver size (rows, columns)	(96, 80)
Channel power profiles	Two-path channel ITU Veh-A channel [34] Jakes model [33]
Channel delay profiles	0 ~ 50 (Uniform distribution)
Normalized maximum Doppler frequency (f_d)	0.1

forms better than the one-tap equalizer in quasi-static channels. In addition, its BER is comparable to that based on ideal CSI knowledge, whereas for the case without channel estimation update, an error floor occurs in the high E_b/N_o region at $\text{BER}=2 \times 10^{-4}$. When compared with the lower bound, there is still an E_b/N_o gap of 1.5dB and 4.5dB at $\text{BER}=10^{-3}$ and 10^{-4} , respectively, for the ML-EM receiver with channel estimation update. Clearly, this gap is due to the error propagation effect. It is worth noting that the performance lower bound in Figure 5.6 comes very close to the theoretical matched-filter bound analyzed in [56]. Similar performance trends are observed in Figure 5.7 for the ITU Veh-A channel. We can see that for the ML-EM receiver with channel estimation update, the required E_b/N_o at $\text{BER}=10^{-3}$ is almost the same as that for an one-tap equalizer in quasi-static channels, while a loss of 4.5dB and 3dB in E_b/N_o can be observed as compared with the lower bound and the case of ideal CSI, respectively.

5.5.2 Effect of Group Size

Figure 5.8 addresses the impact of group size on the BER performance, and the number of Q is selected to keep $(2Q + 1)G \approx 39$ for fair comparison. As expected, joint detection with more subcarriers will attain better performance, and the improvement eventually saturates as the size of G increases.

5.5.3 BER Performance of TURBO-EM Receiver

Figure 5.9 demonstrates the BER performance of the TURBO-EM receiver in the two-path channel. With the BICM, the effect of error propagation is

effectively suppressed. From Figure 5.9, we see that after three iterations, the receiver with channel estimation update achieves a performance close to the lower bound, and the gap between them is only about $0.5dB$ at $BER=10^{-5}$. We also notice that there is a significant performance improvement over the initialization case and it also outperforms the case without channel estimation update by $1dB$. Figure 5.10 depicts the BER performance in the ITU Veh-A channel. We see from Figure 5.10 that after four iterations, the receiver with channel estimation update achieves a performance gap with respect to the lower bound by less than $1dB$ at $BER=10^{-5}$, and the receiver exhibits a remarkable improvement as compared with the initialization stage. However, when CSI is not updated, the performance of the receiver deteriorates remarkably, although it is still better than that of the initialization stage.

5.5.4 FER Performance of TURBO-EM Receiver

Figure 5.11 depicts the frame error rate (FER) performance in the ITU Veh-A channel. This figure also shows that the receiver with channel estimation update performs very well in terms of FER. Hence, in order to achieve a good performance, it is necessary to refine the CSI, especially when the number of pilot tones is small and the normalized maximum Doppler frequency is large. From Figure 5.9 to Figure 5.11, we also observe that the receivers with $G = 2$ and $G = 4$ have nearly identical BER performance.

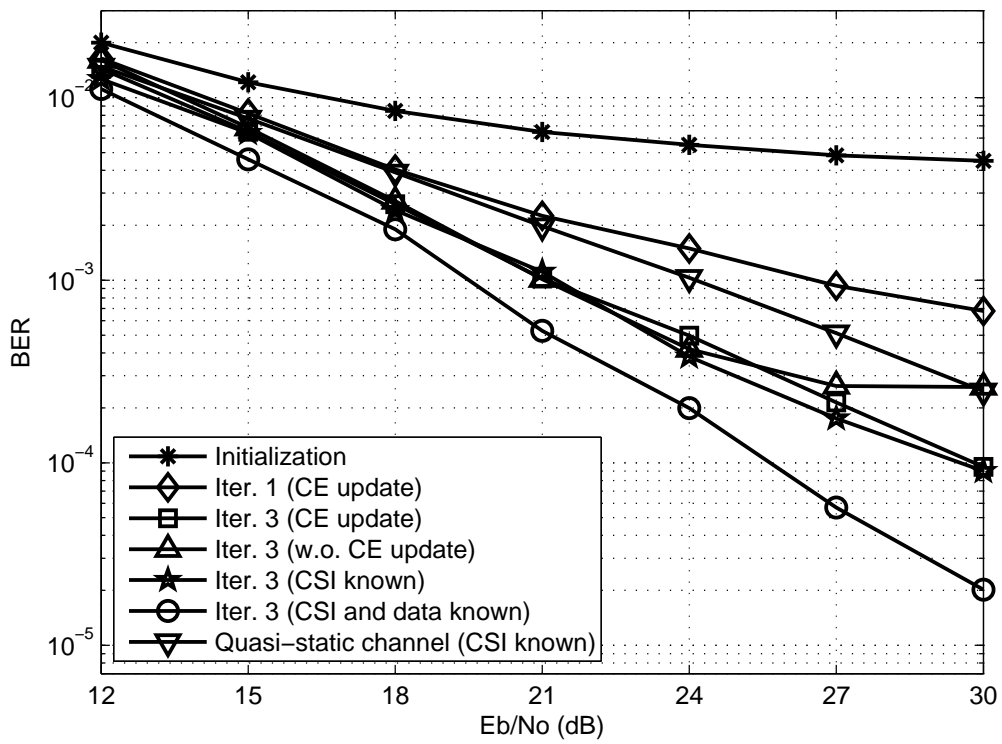


Figure 5.6: BER performance of the ML-EM receiver in the two-path channel ($N_{ML} = 3$ and $[G, Q] = [4, 4]$).

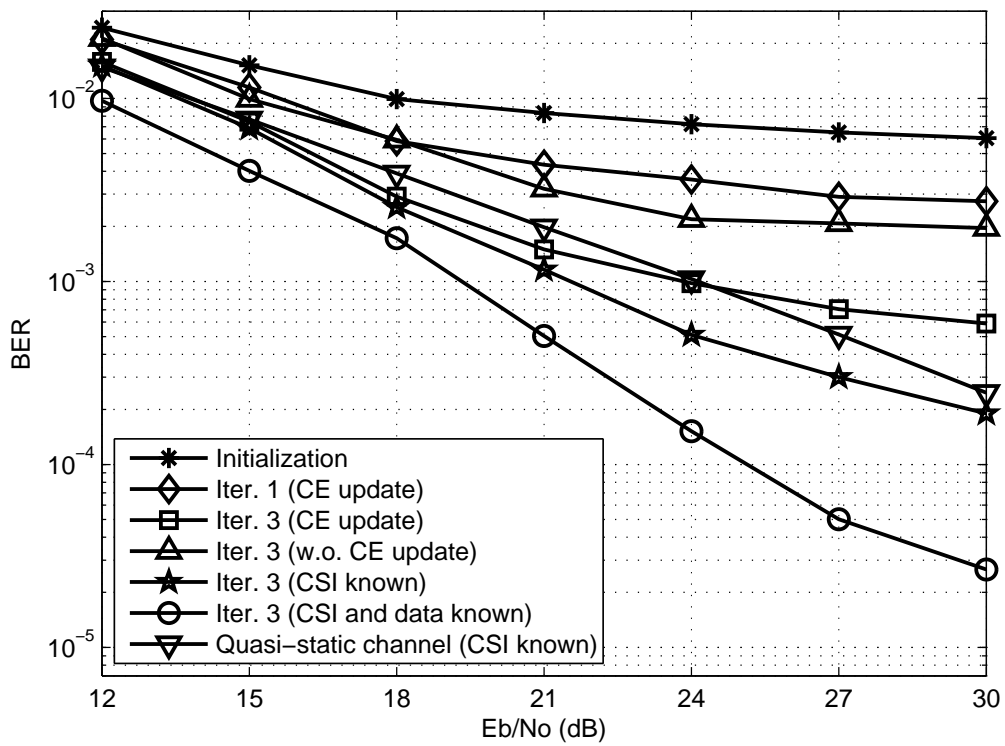


Figure 5.7: BER performance of the ML-EM receiver in the ITU Veh-A channel ($N_{ML} = 3$ and $[G, Q] = [4, 4]$).

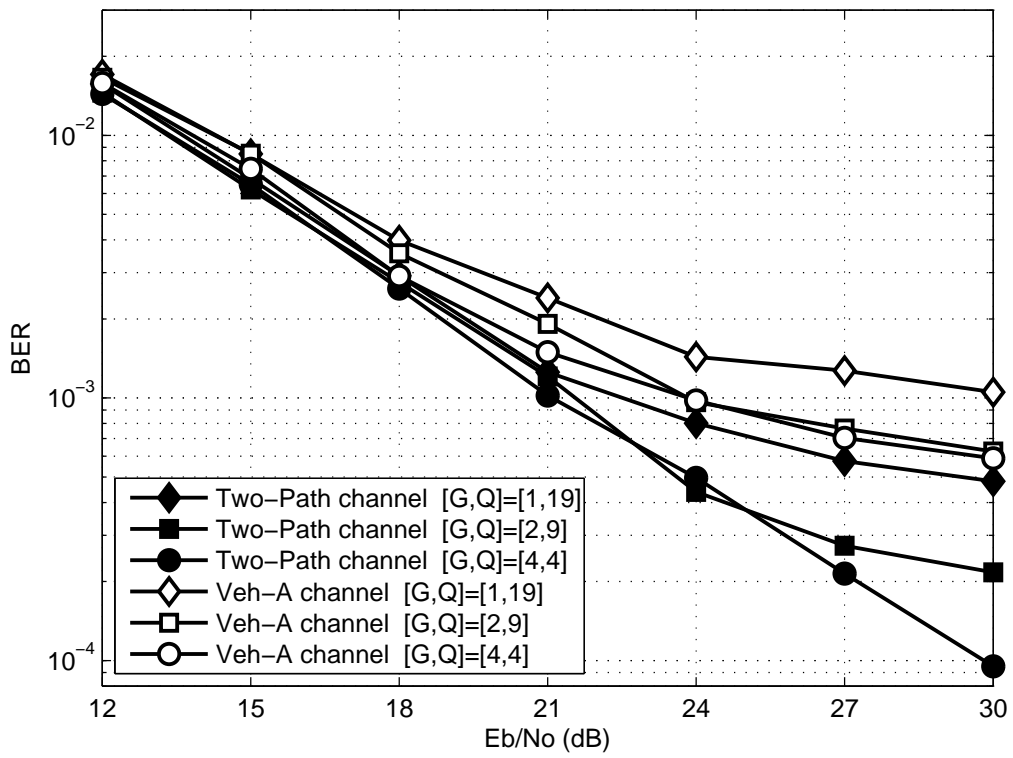


Figure 5.8: BER performance of the ML-EM receiver with channel estimation update for various $[G, Q]$ ($N_{ML} = 3$).

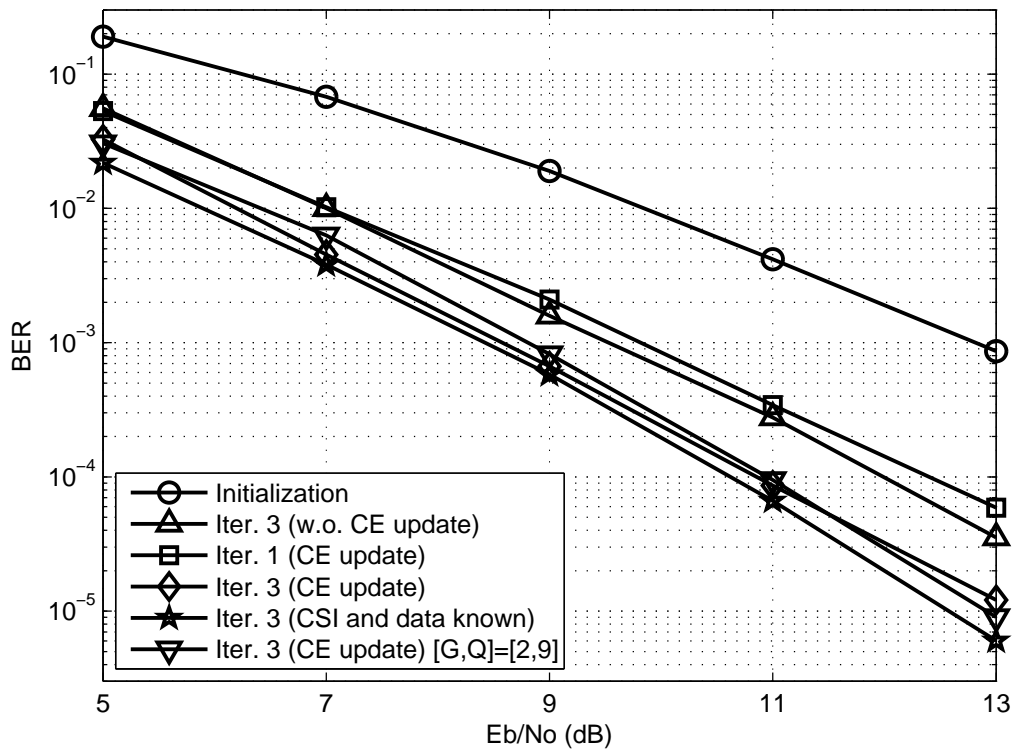


Figure 5.9: BER performance of the TURBO-EM receiver in the two-path channel ($N_{TB} = 3$ and $[G, Q] = [4, 4]$).

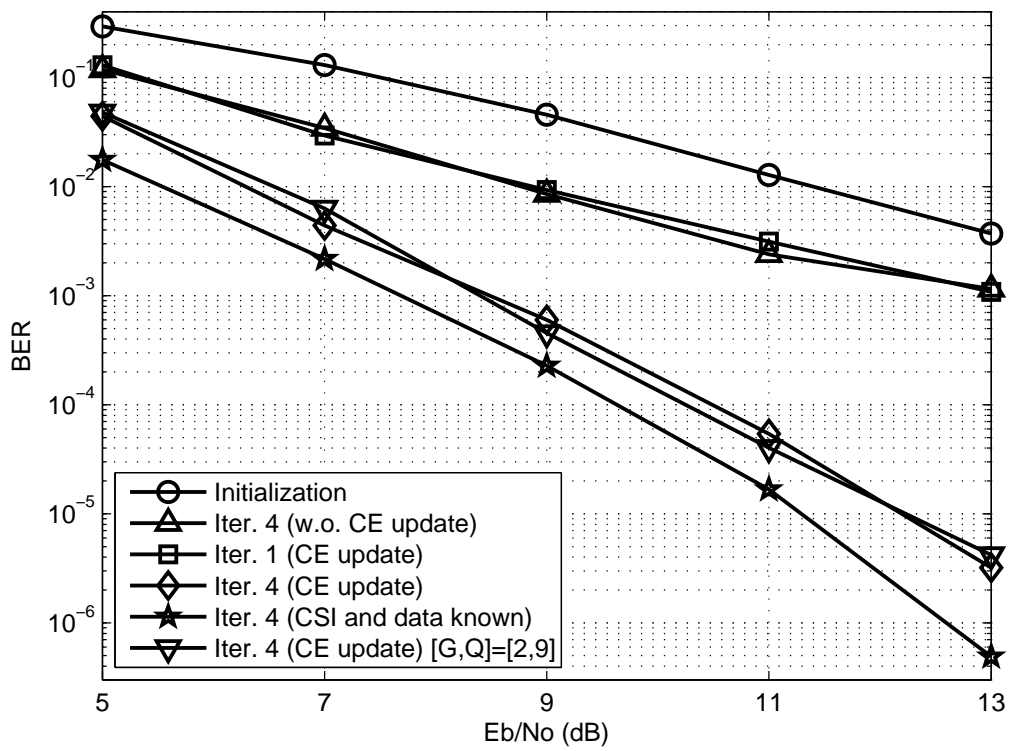


Figure 5.10: BER performance of the TURBO-EM receiver in the ITU Veh-A channel ($N_{TB} = 4$ and $[G, Q] = [4, 4]$).

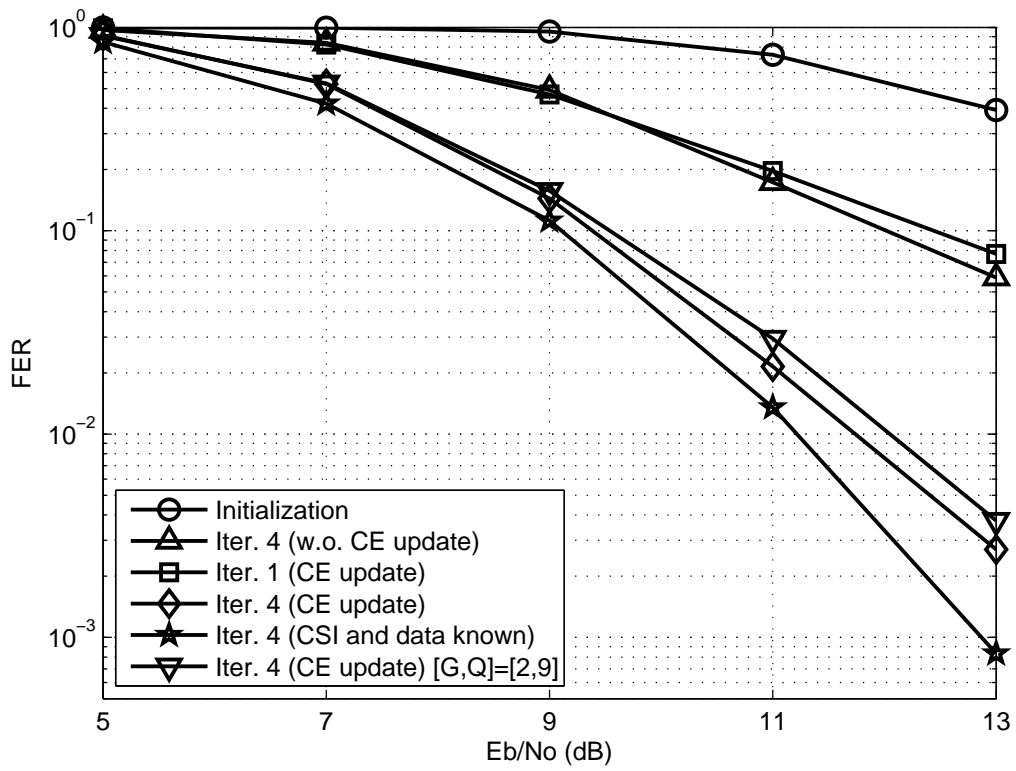


Figure 5.11: FER performance of the TURBO-EM receiver in the ITU Veh-A channel ($N_{TB} = 4$ and $[G, Q] = [4, 4]$).

5.6 Summary

In this chapter, we have investigated two EM-based iterative receivers for OFDM and BICM-OFDM systems in doubly selective channels. Based on the proposed EM algorithm for data detection, both receivers use groupwise processing with ICI cancellation to reduce computational complexity and to explore time diversity inherent in time-variant channels. For OFDM systems, the ML-EM receiver significantly outperforms the conventional one-tap equalizer, and its BER performance even approaches the BER performance without Doppler effect. Compared with the matched-filter bound, an E_b/N_o gap appears because of the error propagation effect. For BICM-OFDM systems, a TURBO-EM receiver, which iterates between the MAP EM detector and the SOVA decoder, is then introduced. This receiver effectively solves the error propagation problem, and it attains a performance close to the low bound in terms of both BER and FER. Simulation results indicate that in order to attain a good performance, the channel estimation update is required when we use low-density pilot tones at high Doppler frequencies. As a final remark, a group size of two to four is large enough to guarantee an acceptable performance under practical channel conditions.

Chapter 6

Conclusions

In this dissertation, we have studied channel estimation and data detection methods for OFDM systems in time-varying multipath channels. The scope of our research encompasses the design of pilot signals for MIMO channels, channel estimation and tracking for low-mobility channels as well as data detection for high-mobility channels. There are four main contributions in our works. First of all, we design CC pilot signals for optimal channel estimation in MIMO systems. It is worth noting that the CC pilot signals not only exhibit the properties of both impulse-like auto-correlation and zero cross-correlation, but also have the characteristic of minimum PAPR in time domain. We also propose a CC pilot-based STBC-OFDM system to achieve bandwidth-efficient transmission at high vehicle speed. A receiver architecture for channel estimation and data detection is developed and its BER performance is analyzed and simulated. The second contribution is to present the equivalence between the DF DFT-based channel estimation method and the Newton's method. Thus, the relationship between them

is well established. We clarify that the DF DFT-based channel estimation method can be further improved through the use of this equivalence. As to the third contribution, a refined channel estimation method is investigated for STBC-OFDM systems. The proposed channel estimation method is accomplished in two stages. In order to reduce computational complexity and improve channel estimation performance, we propose an MPIC-based decorrelation method to catch significant channel paths in the initialization stage. The tracking stage considers the use of a gradient vector derived from pilot tones to track temporal channel variation at the first iteration, followed by the DF DFT-based channel estimation method at the subsequent iterations. In addition, an analytic formula for adaptively determining the optimum step size is investigated. The final contribution of this dissertation is the development of OFDM receivers that enable to deal with the ICI in the presence of Doppler spread in multipath channels. In this work, the EM algorithm is derived and performed for data detection using ML criterion. By integrating the groupwise ICI cancellation with the proposed EM algorithm, we study the design of a low-complexity iterative ML-EM receiver for OFDM systems. Based on the turbo processing principle, a TURBO-EM receiver, for joint detection and decoding in BICM-OFDM systems, is proposed to further improve system performance.

Bibliography

- [1] M. L. Roberts, M. A. Temple, R. F. Mills, and R. A. Raines, “Evolution of the air interface of cellular communications systems toward 4G realization,” *IEEE Commun. Surv. and Tutorials*, vol. 8, no. 1, pp. 2–23, First Quarter 2006.
- [2] T. Hwang, C. Yang, G. Wu, S. Li, and Y. Li, “OFDM and its wireless applications: a survey,” *IEEE Trans. Veh. Technol.*, vol. 58, no. 4, pp. 1673–1694, May 2009.
- [3] F. Wang, A. Ghosh, C. Sankaran, P. J. Fleming, F. Hsieh, and S. J. Benes, “Mobile WiMAX systems: performance and evolution,” *IEEE Commun. Mag.*, vol. 46, no. 10, pp. 41–49, Oct. 2008.
- [4] K. Etemad, “Overview of mobile WiMAX technology and evolution,” *IEEE Commun. Mag.*, vol. 46, no. 10, pp. 31–40, Oct. 2008.
- [5] H. Ekstrom, A. Furuskar, J. Karlsson, M. Meyer, S. Parkvall, J. Torsner, and M. Wahlqvist, “Technical solutions for the 3G long-term evolution,” *IEEE Commun. Mag.*, vol. 44, no. 3, pp. 38–45, Mar. 2006.

- [6] G. J. Foschini, “Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas,” *Bell Labs Tech. J.*, vol. 1, no. 2, pp. 41–59, Autumn 1996.
- [7] L. Zheng and D. N. C. Tse, “Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels,” *IEEE Trans. on Inform. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.
- [8] S. Haykin, M. Sellathurai, Y. de Jong, and T. Willink, “Turbo-MIMO for wireless communications,” *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 48–53, Oct. 2004.
- [9] H. Sampath, S. Talwar, J. Tellado, V. Erceg, and A. Paulraj, “A fourth-generation MIMO-OFDM broadband wireless system: design, performance, and field trial results,” *IEEE Commun. Mag.*, vol. 40, no. 9, pp. 143–149, Sept. 2002.
- [10] H. Yang, “A road to future broadband wireless access: MIMO-OFDM-based air interface,” *IEEE Commun. Mag.*, vol. 43, no. 1, pp. 53–60, Jan. 2005.
- [11] B. Vucetic and J. Yuan, *Space-time coding*. John Wiley and Sons, 2003.
- [12] S. M. Alamouti, “A simple transmit diversity technique for wireless communications,” *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

- [13] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1456–1467, July 1999.
- [14] K. F. Lee and D. B. Williams, "A space-time coded transmitter diversity technique for frequency selective fading channels," in *Proc. IEEE Workshop on Sensor Array and Multichannel Signal Processing*, Mar. 2000, pp. 149–152.
- [15] P. Garg, R. K. Mallik, and H. A. Gupta, "Performance analysis of space-time coding with imperfect channel estimation," in *Proc. IEEE Int. Conf. Personal Wireless Comm.*, Dec. 2002, pp. 71–75.
- [16] S.-A. Yang and J. Wu, "Optimal binary training sequence design for multiple-antenna systems over dispersive fading channels," *IEEE Trans. Veh. Technol.*, vol. 51, no. 5, pp. 1271–1276, Sept. 2002.
- [17] X. Ma, L. Yang, and G. B. Giannakis, "Optimal training for MIMO frequency-selective fading channels," in *Proc. 36th Asilomar Conf. Signals, Systems, Comput.*, Nov. 2002, pp. 1107–1111.
- [18] C. Fragouli, N. Al-Dhahir, and W. Turin, "Training-based channel estimation for multiple-antenna broadband transmissions," *IEEE Trans. Wireless Commun.*, vol. 2, no. 2, pp. 384–391, Mar. 2003.
- [19] G. Kang, E. Costa, M. Weckerle, and E. Schulz, "Optimum channel estimation over frequency-selective fading channel in multiple antenna

- systems,” in *Proc. IEEE Int. Conf. Commun. Technol., ICCT*, Apr. 2003, pp. 1799–1803.
- [20] Y. Li, N. Seshadri, and S. Ariyavisitakul, “Channel estimation for OFDM systems with transmitter diversity in mobile wireless channels,” *IEEE J. Sel. Areas Commun.*, vol. 17, no. 3, pp. 461–470, Mar. 1999.
- [21] Y. Li, “Simplified channel estimation for OFDM systems with multiple transmit antennas,” *IEEE Trans. Wireless Commun.*, vol. 1, no. 1, pp. 67–75, Jan. 2002.
- [22] H. Minn, D. I. Kim, and V. K. Bhargava, “A reduced complexity channel estimation for OFDM systems with transmit diversity in mobile wireless channels,” *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 799–807, May 2002.
- [23] S. Kang and J. S. Lehnert, “Channel estimation for OFDM systems with transmitter diversity for a quasi-static fading channel,” in *Proc. IEEE Military Commun. Conf., MILCOM*, Oct. 2003, pp. 309–313.
- [24] G. Gong and W.-C. Ge, “Research on an OFDM system using superimposed PN sequences in time domain,” in *Proc. IEEE Wireless Commun., Networking and Mobile Comput.*, Oct. 2008, pp. 1–5.
- [25] J. Li, J. Ma, and S. Liu, “RLS channel estimation with superimposed training sequence in OFDM systems,” in *Proc. IEEE Int. Conf. Commun. Technol.*, Nov. 2008, pp. 175–178.

- [26] J. P. Nair and R. V. R. Kumar, "An iterative channel estimation method using superimposed training for IEEE 802.16e based OFDM systems," in *Proc. IEEE Int. Symp. on Consumer Electronics*, Apr. 2008, pp. 1–4.
- [27] Q. Yang, K. S. Kwak, and F. Fu, "Channel estimation for STBC MB-OFDM UWB systems with superimposed training," in *Proc. IEEE Int. Symp. on Commun. and Inform. Technol.*, Oct. 2008, pp. 238–241.
- [28] J. P. Nair and R. V. R. Kumar, "A bandwidth efficient channel estimation method using superimposed training for MIMO-OFDM systems," in *Proc. IEEE TENCON 2008 Region 10 Conf.*, Nov. 2008, pp. 1–5.
- [29] N. Chen and G. T. Zhou, "Superimposed training for OFDM: a peak-to-average power ratio analysis," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2277–2287, June 2006.
- [30] M. Golay, "Multislit spectroscopy," *J. Opt. Soc. Amer.*, vol. 39, pp. 437–444, 1949.
- [31] M. J. E. Golay, "Complementary series," *IEEE Trans. Inf. Theory*, vol. 7, no. 2, pp. 82–87, Apr. 1961.
- [32] R. V. Nee, "OFDM codes for peak-to-average power reduction and error correction," in *Proc. IEEE Global Commun. Conf.*, Nov. 1996, pp. 740–744.
- [33] W. C. Jakes, *Microwave mobile communications*. New York: Wiley, 1974.

- [34] J. Laiho, A. Wacker, and T. Novosad, *Radio network planning and optimisation for UMTS*. New York: Wiley, 2002.
- [35] S. Werner, M. Enescu, and V. Koivunen, “Low-complexity time-domain channel estimators for mobile wireless OFDM systems,” in *Proc. IEEE Workshop Signal Processing Systems Design and Implementation*, Nov. 2005, pp. 245–250.
- [36] J.-H. Park, M.-K. Oh, and D.-J. Park, “New channel estimation exploiting reliable decision-feedback symbols for OFDM systems,” in *Proc. IEEE Int. Conf. on Commun.*, June 2006, pp. 3046–3051.
- [37] L. Deneire, P. Vandenameele, L. van der Perre, B. Gyselinckx, and M. Engels, “A low-complexity ML channel estimator for OFDM,” *IEEE Trans. Commun.*, vol. 51, no. 2, pp. 135–140, Feb. 2003.
- [38] M. Morelli and U. Mengali, “A comparison of pilot-aided channel estimation methods for OFDM systems,” *IEEE Trans. Signal Process.*, vol. 49, no. 12, pp. 3065–3073, Dec. 2001.
- [39] O. Edfors, M. Sandell, J. J. van de Beek, S. K. Wilson, and P. O. Borjesson, “Analysis of DFT-based channel estimators for OFDM,” *Wireless Pers. Commun.*, vol. 12, no. 1, pp. 55–70, Jan. 2000.
- [40] V. Tarokh, N. Seshadri, and A. Calderbank, “Space-time codes for high data rate wireless communication: performance criterion and code construction,” *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.

- [41] J. S. Arora, *Introduction to optimum design*. Elsevier, 2004.
- [42] M.-L. Ku and C.-C. Huang, "A refined channel estimation method for STBC/OFDM systems in high-mobility wireless channels," *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4312–4320, Nov. 2008.
- [43] IEEE std. 802.16-2004, "IEEE local and metropolitan area networks part 16: Air interface for fixed broadband wireless access systems," Tech. Rep., Oct. 2004.
- [44] S. Colieri, M. Ergen, A. Puri, and A. Bahai, "A study of channel estimation in OFDM systems," in *Proc. IEEE Veh. Technol. Conf.*, Sept. 2002, pp. 894–898.
- [45] K. S. Ahn and H. K. Baik, "Decision feedback detection for space-time block coding over time-selective fading channels," in *Proc. IEEE Personal, Indoor and Mobile Radio Commun.*, Sept. 2003, pp. 1983–1987.
- [46] A. Chini, Y. Wu, M. El-Tanany, and S. Mahmoud, "Filtered decision feedback channel estimation for OFDM-based DTV terrestrial broadcasting system," *IEEE Trans. Broadcast.*, vol. 44, no. 1, pp. 2–11, Mar. 1998.
- [47] M.-L. Ku and C.-C. Huang, "A derivation on the equivalence between Newton's method and DF DFT-based method for channel estimation in OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 7, no. 10, pp. 3982–3987, Oct. 2008.

- [48] Y. Gong and K. B. Letaief, "Low complexity channel estimation for space-time coded wideband OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 2, no. 5, pp. 876–882, Sept. 2003.
- [49] R. R. Muller and S. Verdu, "Design and analysis of low-complexity interference mitigation on vector channels," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 8, pp. 1429–1441, Aug. 2001.
- [50] T. Y. Al-Naffouri, O. Awoniyi, O. Oteri, and A. Paulraj, "Receiver design for MIMO-OFDM transmission over time variant channels," in *Proc. IEEE Global Commun. Conf.*, Dec. 2004, pp. 2487–2492.
- [51] Z. Liu, X. Ma, and G. B. Giannakis, "Space-time coding and Kalman filtering for time-selective fading channels," *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 183–186, Feb. 2002.
- [52] R. van Nee and R. Prasad, *OFDM for wireless multimedia communications*. Norwell, MA: Artech House, 2000.
- [53] E. Akay and E. Ayanoglu, "Achieving full frequency and space diversity in wireless systems via BICM, OFDM, STBC and Viterbi decoding," *IEEE Trans. Commun.*, vol. 54, no. 12, pp. 2164–2172, Dec. 2006.
- [54] IEEE std. 802.16e 2005 and IEEE 802.16-2004/Cor1 2005, "IEEE local and metropolitan area networks part 16: air interface for fixed and mobile broadband wireless access systems," IEEE-SA standards board, Tech. Rep., 2006.

- [55] Y. Li and J. L. J. Cimini, "Bounds on the interchannel interference of OFDM in time-varying impairments," *IEEE Trans. Commun.*, vol. 49, no. 3, pp. 401–404, Mar. 2001.
- [56] X. Cai and G. Giannakis, "Bounding performance and suppressing inter-carrier interference in wireless mobile OFDM," *IEEE Trans. Commun.*, vol. 51, no. 12, pp. 2047–2056, Dec. 2003.
- [57] M.-X. Chang, "A novel algorithm of inter-subchannel interference self-cancellation for OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 6, no. 8, pp. 2881–2893, Aug. 2007.
- [58] A. Seyedi and G. J. Saulnier, "General ICI self-cancellation scheme for OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 54, no. 1, pp. 198–210, Jan. 2005.
- [59] H.-C. Wu, X. Huang, Y. Wu, and X. Wang, "Theoretical studies and efficient algorithm of semi-blind ICI equalization for OFDM," *IEEE Trans. Wireless Commun.*, vol. 7, no. 10, pp. 3791–3798, Oct. 2008.
- [60] P. Schniter, "Low-complexity equalization of OFDM in doubly selective channels," *IEEE Trans. Signal Process.*, vol. 52, no. 4, pp. 1002–1011, Apr. 2004.
- [61] T. Wang, J. G. Proakis, and J. R. Zeidler, "Techniques for suppression of intercarrier interference in OFDM systems," in *Proc. IEEE Wireless Commun. and Networking Conf.*, Mar. 2005, pp. 39–44.

- [62] G. Li, H. Yang, L. Cai, and L. Gui, "A low-complexity equalization technique for OFDM system in time-variant multipath channels," in *Proc. IEEE Veh. Technol. Conf.*, Oct. 2003, pp. 2466–2470.
- [63] Y.-S. Choi, P. J. Voltz, and F. A. Cassara, "On channel estimation and detection for multicarrier signals in fast and selective Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 49, no. 8, pp. 1375–1387, Aug. 2001.
- [64] W. G. Jeon, K. H. Chang, and Y. S. Cho, "An equalization technique for orthogonal frequency-division multiplexing systems in time-variant multipath channels," *IEEE Trans. Commun.*, vol. 47, no. 1, pp. 27–32, Jan. 1999.
- [65] K. Kim and H. Park, "A low complexity ICI cancellation method for high mobility OFDM systems," in *Proc. IEEE Veh. Technol. Conf.*, May 2006, pp. 2528–2532.
- [66] S. Kim and G. Pottie, "Robust OFDM in fast fading channel," in *Proc. IEEE Global Commun. Conf.*, Dec. 2003, pp. 1074–1078.
- [67] A. Gorokhov and J. P. Linnartz, "Robust OFDM receivers for dispersive time-varying channels: equalization and channel acquisition," *IEEE Trans. Commun.*, vol. 52, no. 4, pp. 572–583, Apr. 2004.
- [68] S. Tomasin, A. Gorokhov, H. Yang, and J. P. Linnartz, "Iterative interference cancellation and channel estimation for mobile OFDM," *IEEE Trans. Wireless Commun.*, vol. 4, no. 1, pp. 238–245, Jan. 2005.

- [69] K. Chang, Y. Han, J. Ha, and Y. Kim, "Cancellation of ICI by Doppler effect in OFDM systems," in *Proc. IEEE Veh. Technol. Conf.*, May 2006, pp. 1411–1415.
- [70] Y. Mostofi and D. C. Cox, "ICI mitigation for pilot-aided OFDM mobile systems," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 765–774, Mar. 2005.
- [71] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society Series B-Methodological*, vol. 39, no. 1, pp. 1–38, 1977.
- [72] C. P. Robert and G. Casella, *Monte carlo statistical methods*. New York: Springer-Verlag, 1999.
- [73] B. Lu, X. Wang, and Y. Li, "Iterative receivers for space-time block-coded OFDM systems in dispersive fading channels," *IEEE Trans. Wireless Commun.*, vol. 1, no. 2, pp. 213–225, Apr. 2002.
- [74] T. Y. Al-Naffouri, "An EM-based forward-backward Kalman filter for the estimation of time-variant channels in OFDM," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3924–3930, July 2007.
- [75] K. Muraoka, K. Fukawa, H. Suzuki, and S. Suyama, "Channel estimation using differential model of fading fluctuation for EM algorithm applied to OFDM MAP detection," in *Proc. IEEE Personal, Indoor and Mobile Radio Commun.*, Sept. 2007, pp. 1–5.

- [76] S. M. Kay, *Fundamentals of statistical signal processing: estimation theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [77] F. Peng and W. E. Ryan, “A low-complexity soft demapper for OFDM fading channels with ICI,” in *Proc. IEEE Wireless Commun. and Networking Conf.*, Apr. 2006, pp. 1549–1554.
- [78] B. M. Hochwald and S. ten Brink, “Achieving near-capacity on a multiple-antenna channel,” *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 389–399, Mar. 2003.
- [79] C. Berrou and A. Glavieux, “Near optimum error correcting coding and decoding: turbo-codes,” *IEEE Trans. Commun.*, vol. 44, no. 10, pp. 1261–1271, Oct. 1996.
- [80] J. A. Fessler and A. O. Hero, “Space-alternating generalized expectation-maximization algorithm,” *IEEE Trans. Signal Process.*, vol. 42, no. 10, pp. 2664–2677, Oct. 1994.
- [81] R. L. Frank, “Polyphase complementary codes,” *IEEE Trans. Inf. Theory*, vol. 26, no. 6, pp. 641–647, Nov. 1980.

Appendix A

Complementary Sequences

Complementary sequences are defined as a pair of sequences having the sum of their autocorrelation values equal to a Kronecker delta function. Binary complementary sequences, also widely named Golay sequences, were first introduced by Marcel J. E. Golay in 1949 [30,31]. Later, complementary sequences were generalized to polyphase or multilevel complementary sequences by other authors [81].

Let $\{\alpha[0], \dots, \alpha[N-1]\}$ and $\{\beta[0], \dots, \beta[N-1]\}$ be a pair of binary complementary sequences, i.e., the values of the two sequences are either "+1" or "-1". The two sequences are complementary if we have

$$\begin{aligned}\gamma[n] &\equiv \sum_{m=0}^{N-1} \{\alpha[m] \alpha^*[((m-n))_N] \\ &\quad + \beta[m] \beta^*[((m-n))_N]\} \\ &= 2N \cdot \delta[n] \\ &= \begin{cases} 2N, & \text{for } n = 0 \\ 0, & \text{for } n \neq 0 \end{cases} \end{aligned} \tag{A.1}$$

where $\delta[n]$ is a Kronecker delta function. This ideal autocorrelation property

makes complementary sequences attractive for many applications like radar pulse compression and spread spectrum communication. From (A.1), the complementary sequences in frequency domain representation have complementary power spectrum as follows:

$$\Gamma[k] = \sum_{k=0}^{N-1} |A[k]|^2 + |B[k]|^2 = 2N \quad (\text{A.2})$$

where $\{A[k]\}$ and $\{B[k]\}$ are the DFT of $\{\alpha[n]\}$ and $\{\beta[n]\}$, respectively. For the simplest example, we have binary complementary sequences of length two, given by $\{\alpha[n]\} = [+1, +1]$ and $\{\beta[n]\} = [+1, -1]$. Given a pair of complementary sequences $\{\alpha[n]\}$ and $\{\beta[n]\}$, a new pair of complementary sequences can be generated from the following rules if

1. Any of the two sequences is multiplied by $e^{j\phi}$.
2. Any of the two sequences is time-reversed.
3. Any of the two sequences is circular-shifted.
4. The two sequences are interchanged.
5. Both sequences are decimated in time by K .
6. Both sequences are multiplied by $e^{j\pi kn/N}$, where k is a constant.
7. The two sequences are concatenated to form $[\alpha[0], \dots, \alpha[N-1], \beta[0], \dots, \beta[N-1]]$ and $[\alpha[0], \dots, \alpha[N-1], -\beta[0], \dots, -\beta[N-1]]$.
8. The two sequences are concatenated and interleaved to form $[\alpha[0], \beta[0], \dots, \alpha[N-1], \beta[N-1]]$ and $[\alpha[0], -\beta[0], \dots, \alpha[N-1], -\beta[N-1]]$.

9. The two sequences are added and subtracted to form $[\alpha [0] + \beta [0], \dots, \alpha [N - 1] + \beta [N - 1]]$ and $[\alpha [0] - \beta [0], \dots, \alpha [N - 1] - \beta [N - 1]]$.



Appendix B

Proof of (2.37) and (2.38)

By substituting the definition of \tilde{D}_e in (2.36) into $\mathbb{E}[\tilde{D}_e]$, we can obtain

$$\begin{aligned}\mathbb{E}[\tilde{D}_e] &= (1 - BER_C(\zeta))(1 - BER_C(\zeta)) \mathbb{E}[\Lambda_1 \hat{X}_{FC} \\ &\quad + \Lambda_2 \hat{X}_{SC}] \\ &\quad + (1 - BER_C(\zeta)) BER_C(\zeta) \mathbb{E}[\Lambda_1 \hat{X}_{FC} + 2H_2 X_S \\ &\quad + \Lambda_2 \hat{X}_{SC}] \\ &\quad + BER_C(\zeta)(1 - BER_C(\zeta)) \mathbb{E}[2H_1 X_F + \Lambda_1 \hat{X}_{FC} \\ &\quad + \Lambda_2 \hat{X}_{SC}] \\ &\quad + BER_C(\zeta) BER_C(\zeta) \mathbb{E}[2H_1 X_F + \Lambda_1 \hat{X}_{FC} \\ &\quad + 2H_2 X_S + \Lambda_2 \hat{X}_{SC}] \\ &= 2BER_C(\zeta)(H_1 X_F + H_2 X_S)\end{aligned}\tag{B.1}$$

Moreover, we can calculate the second moment of \tilde{D}_e as follows:

$$\begin{aligned}
\mathbb{E} \left[\left| \tilde{D}_e \right|^2 \right] &= (1 - BER_C(\zeta)) (1 - BER_C(\zeta)) \mathbb{E} \left[\left| \Lambda_1 \right|^2 \left| \hat{X}_{FC} \right|^2 \right. \\
&\quad \left. + \left| \Lambda_2 \right|^2 \left| \hat{X}_{SC} \right|^2 \right] \\
&\quad + (1 - BER_C(\zeta)) BER_C(\zeta) \mathbb{E} \left[\left| \Lambda_1 \right|^2 \left| \hat{X}_{FC} \right|^2 + 4 \left| H_2 \right|^2 \left| X_S \right|^2 \right. \\
&\quad \left. + \left| \Lambda_2 \right|^2 \left| \hat{X}_{SC} \right|^2 \right] \\
&\quad + BER_C(\zeta) (1 - BER_C(\zeta)) \mathbb{E} \left[4 \left| H_1 \right|^2 \left| X_F \right|^2 + \left| \Lambda_1 \right|^2 \left| \hat{X}_{FC} \right|^2 \right. \\
&\quad \left. + \left| \Lambda_2 \right|^2 \left| \hat{X}_{SC} \right|^2 \right] \\
&\quad + BER_C(\zeta) BER_C(\zeta) \mathbb{E} \left[4 \left| H_1 \right|^2 \left| X_F \right|^2 + \left| \Lambda_1 \right|^2 \left| \hat{X}_{FC} \right|^2 \right. \\
&\quad \left. + 4 \left| H_2 \right|^2 \left| X_S \right|^2 + \left| \Lambda_2 \right|^2 \left| \hat{X}_{SC} \right|^2 \right. \\
&\quad \left. + 4 H_1^* X_F^* H_2 X_S + 4 H_2^* X_S^* H_1 X_F \right] \\
&= 2 BER_C(\zeta) \zeta E_b + \sigma_\lambda^2 E_b \\
&\quad + 4 BER_C^2(\zeta) (H_1 H_2^* X_F X_S^* + H_1^* H_2 X_F^* X_S) \tag{B.2}
\end{aligned}$$

Appendix C

Proof of (3.21)

Define $\mathbf{g}^{(j,i)} = [\mathbf{g}_1^T, \mathbf{g}_2^T]^T$ and $\mathbf{q}^{(j,i)} = [\mathbf{q}_1^T, \mathbf{q}_2^T]^T$, where \mathbf{g}_1 , \mathbf{g}_2 , \mathbf{q}_1 , and \mathbf{q}_2 are of size $G \times 1$. From (3.15)–(3.17), we convert $\mathbf{g}^{(j,i)} = \mathbf{E}^{(j,i)^{-1}} \mathbf{q}^{(j,i)}$ into block matrix representation as follows

$$\begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{A} - \mathbf{B} \\ \mathbf{B} \ \mathbf{A} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{q}_1 \\ \mathbf{q}_2 \end{bmatrix} \quad (\text{C.1})$$

where \mathbf{A} and \mathbf{B} are $G \times G$ sub-matrices of $\mathbf{E}^{(j,i)}$. It is clear that the matrix $\mathbf{E}^{(j,i)^{-1}}$ holds the same structure as the matrix $\mathbf{E}^{(j,i)}$, given by

$$\mathbf{E}^{(j,i)^{-1}} = \begin{bmatrix} \mathbf{C} - \mathbf{D} \\ \mathbf{D} \ \mathbf{C} \end{bmatrix} \quad (\text{C.2})$$

where $\mathbf{A}\mathbf{C} - \mathbf{B}\mathbf{D} = \mathbf{I}_G$, $\mathbf{B}\mathbf{C} + \mathbf{A}\mathbf{D} = \mathbf{0}_G$, and $\mathbf{0}_G$ is a zero matrix of size $G \times G$. Hence, we have

$$\begin{aligned} \mathbf{g}_1 + j\mathbf{g}_2 &= (\mathbf{C}\mathbf{q}_1 - \mathbf{D}\mathbf{q}_2) + j(\mathbf{D}\mathbf{q}_1 + \mathbf{C}\mathbf{q}_2) \\ &= (\mathbf{C} + j\mathbf{D})(\mathbf{q}_1 + j\mathbf{q}_2) \\ &= (\mathbf{A} + j\mathbf{B})^{-1}(\mathbf{q}_1 + j\mathbf{q}_2) \end{aligned} \quad (\text{C.3})$$

Appendix D

Explanation of Hessian Matrix

In this appendix, we provide an explanation of $\tilde{\mathbf{E}}^{(j,i)-1}$. For simplicity, we assume that the DF data symbols are all correct, i.e., $\hat{\mathbf{X}}[k] = \mathbf{X}[k]$, and neglect noise terms. Therefore, the LS estimate in (3.20) for channel $H^{(j,i)}[k]$ given in (3.1) becomes

$$C[k] H^{(j,i)}[k] = C[k] \sum_{l=1}^G \check{\mu}_l^{(j,i)} e^{-j \frac{2\pi(k-1)(l-1)}{K}} \quad (\text{D.1})$$

for $k \in \Theta$, where $C[k] = \sum_{t=1}^{N_L} |X^{(i)}[t, k]|^2$ and $\check{\mu}_l^{(j,i)}$ is the complex gain of the l th path. Taking the IDFT of (D.1), we get the estimate for the l' th channel path gain as follows

$$\hat{\eta}_{l'}^{(j,i)} = \sum_{l=1}^G \check{\mu}_l^{(j,i)} \sum_{k \in \Theta} C[k] e^{j \frac{2\pi(k-1)(l'-l)}{K}} \quad (\text{D.2})$$

where $1 \leq l' \leq G$. By rewriting (D.2) in a vector form, we have

$$\hat{\boldsymbol{\eta}}^{(j,i)} = \frac{1}{2} \tilde{\mathbf{E}}^{(j,i)} \check{\boldsymbol{\mu}}^{(j,i)} \quad (\text{D.3})$$

where $\tilde{\mathbf{E}}^{(j,i)}$ is defined as in (3.21), $\hat{\boldsymbol{\eta}}^{(j,i)} = [\hat{\eta}_1^{(j,i)}, \dots, \hat{\eta}_G^{(j,i)}]^T$, and $\check{\boldsymbol{\mu}}^{(j,i)} = [\check{\mu}_1^{(j,i)}, \dots, \check{\mu}_G^{(j,i)}]^T$. As can be seen in (D.2) and (D.3), $\check{\mu}_l^{(j,i)}$ from other paths

causes interference in $\hat{\eta}_\nu^{(j,i)}$ due to the effect of aliasing, and $\tilde{\mathbf{E}}^{(j,i)^{-1}}$ acts as a path decorrelator to mitigate this effect.



Appendix E

Review of EM Algorithm

Consider a general ML estimate problem in a missing data model. Suppose that we observe a vector \mathbf{y} generated from $P(\mathbf{y}|\mathbf{x},\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is the parameter vector and \mathbf{x} represents the missing data vector (or called unobserved latent data vector), and want to compute the ML estimate

$$\hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta}} P(\mathbf{y}|\boldsymbol{\theta}) \quad (\text{E.1})$$

The EM algorithm seeks to find the ML estimate of (E.1) by iteratively applying the following two steps:

E - step:

Compute

$$\Omega\left(\boldsymbol{\theta}|\mathbf{y},\hat{\boldsymbol{\theta}}^{(m-1)}\right) = \mathbb{E}_{\mathbf{x}|\mathbf{y},\hat{\boldsymbol{\theta}}^{(m-1)}} [L(\mathbf{y},\mathbf{x}|\boldsymbol{\theta})] \quad (\text{E.2})$$

where the expectation is with respect to $P(\mathbf{x}|\mathbf{y},\hat{\boldsymbol{\theta}}^{(m-1)})$.

M - step:

Maximize

$$\hat{\boldsymbol{\theta}}^{(m)} = \arg \max_{\boldsymbol{\theta}} \Omega\left(\boldsymbol{\theta}|\mathbf{y},\hat{\boldsymbol{\theta}}^{(m-1)}\right) \quad (\text{E.3})$$

We refer to $L(\mathbf{y}, \mathbf{x} | \boldsymbol{\theta})$ as the complete log-likelihood function which corresponds to the complete data of both \mathbf{y} and \mathbf{x} . The theoretical core of the EM algorithm is that the likelihood function $L(\mathbf{y} | \boldsymbol{\theta})$ is guaranteed to increase monotonically at each iteration by maximizing $\Omega(\boldsymbol{\theta} | \mathbf{y}, \hat{\boldsymbol{\theta}}^{(m-1)})$ at the M-step. In other word, we can ensure that

$$L(\mathbf{y} | \hat{\boldsymbol{\theta}}^{(m)}) \geq L(\mathbf{y} | \hat{\boldsymbol{\theta}}^{(m-1)}) \quad (\text{E.4})$$

The EM procedure is conducted until a fixed value of $\Omega(\hat{\boldsymbol{\theta}}^{(m)} | \mathbf{y}, \hat{\boldsymbol{\theta}}^{(m-1)})$ is obtained. In this case, we have $L(\mathbf{y} | \hat{\boldsymbol{\theta}}^{(m)}) = L(\mathbf{y} | \hat{\boldsymbol{\theta}}^{(m-1)})$. The theorem can be proved as follows. According to the Bayesian rule and taking logarithm, it is easy to show that

$$L(\mathbf{y} | \boldsymbol{\theta}) = L(\mathbf{y}, \mathbf{x} | \boldsymbol{\theta}) - L(\mathbf{x} | \mathbf{y}, \boldsymbol{\theta}) \quad (\text{E.5})$$

By taking the expectation of (E.5) with respect to $P(\mathbf{x} | \mathbf{y}, \hat{\boldsymbol{\theta}}^{(m-1)})$, it leads to

$$L(\mathbf{y} | \boldsymbol{\theta}) = \mathbb{E}_{\mathbf{x} | \mathbf{y}, \hat{\boldsymbol{\theta}}^{(m-1)}} [L(\mathbf{y}, \mathbf{x} | \boldsymbol{\theta})] - \mathbb{E}_{\mathbf{x} | \mathbf{y}, \hat{\boldsymbol{\theta}}^{(m-1)}} [L(\mathbf{x} | \mathbf{y}, \boldsymbol{\theta})] \quad (\text{E.6})$$

From the definition of the M-step in (E.3), it is straightforward to find

$$\mathbb{E}_{\mathbf{x} | \mathbf{y}, \hat{\boldsymbol{\theta}}^{(m-1)}} [L(\mathbf{y}, \mathbf{x} | \hat{\boldsymbol{\theta}}^{(m)})] \geq \mathbb{E}_{\mathbf{x} | \mathbf{y}, \hat{\boldsymbol{\theta}}^{(m-1)}} [L(\mathbf{y}, \mathbf{x} | \hat{\boldsymbol{\theta}}^{(m-1)})] \quad (\text{E.7})$$

Moreover, from Jensen's inequality, we can obtain

$$\begin{aligned}
& \mathbb{E}_{\mathbf{x}|\mathbf{y}, \hat{\boldsymbol{\theta}}^{(m-1)}} \left[\log \frac{P(\mathbf{x}|\mathbf{y}, \hat{\boldsymbol{\theta}}^{(m)})}{P(\mathbf{x}|\mathbf{y}, \hat{\boldsymbol{\theta}}^{(m-1)})} \right] \\
& \leq \log \mathbb{E}_{\mathbf{x}|\mathbf{y}, \hat{\boldsymbol{\theta}}^{(m-1)}} \left[\frac{P(\mathbf{x}|\mathbf{y}, \hat{\boldsymbol{\theta}}^{(m)})}{P(\mathbf{x}|\mathbf{y}, \hat{\boldsymbol{\theta}}^{(m-1)})} \right] \\
& = \log \int P(\mathbf{x}|\mathbf{y}, \hat{\boldsymbol{\theta}}^{(m)}) d\mathbf{x} \\
& = 0
\end{aligned} \tag{E.8}$$

Thus, we can verify (E.4) by using (E.6)-(E.8).



Appendix F

Proof of (5.24) and (5.25)

Let us first define an AR channel model

$$h[l, N-1] = \alpha h[l, 0] + u \quad (\text{F.1})$$

where α is the parameter of the AR model and u represents a complex Gaussian random variable, independent of $h[l, n]$, with zero mean and variance σ_u^2 . From Jakes model, α is evaluated by

$$\alpha = \mathbb{E}[h[l, N-1] h^*[l, 0]] = J_0\left(\frac{2\pi f_d (N-1)}{N}\right) \quad (\text{F.2})$$

where $J_0(\cdot)$ is the zeroth order Bessel function of the first kind, f_d denotes the normalized maximum Doppler frequency. Recall from (5.4) that we have $\mathbb{E}[h[l, n]] = 0$ and $\mathbb{E}[h^2[l, n]] = \Xi_l$. Following the energy conservation rule in (F.1), the variance of u can be calculated as $\sigma_u^2 = (1 - \alpha^2) \Xi_l$. According to (F.1), (F.2) and (5.5), we find the slope of the l th tap over the duration of one OFDM symbol as

$$\begin{aligned} a[l, 1] &= \frac{1}{N-1} \{h[l, N-1] - h[l, 0]\} \\ &= \frac{1}{N-1} \{(\alpha - 1) h[l, 0] + u\} \end{aligned} \quad (\text{F.3})$$

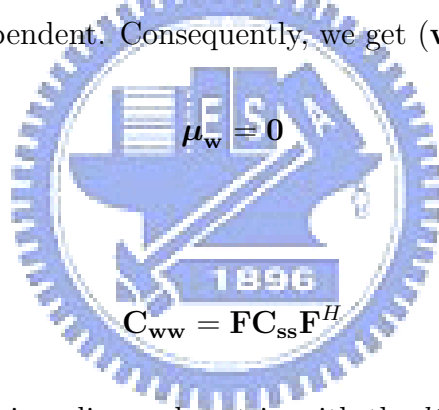
and its mean and variance is calculated by

$$\mathbb{E}[a[l, 1]] = 0 \quad (\text{F.4})$$

and

$$\begin{aligned} \mathbb{E}[a[l, 1] a^*[l, 1]] &= \frac{1}{(N-1)^2} \{(\alpha-1)^2 \Xi_l + (1-\alpha^2) \Xi_l\} \\ &= \frac{2(1-\alpha) \Xi_l}{(N-1)^2} \end{aligned} \quad (\text{F.5})$$

Besides, it is reasonable to assume that the slopes of channel taps are independent of each other, i.e., $\mathbb{E}[a[l, 1] a^*[l', 1]] = 0$ if $l \neq l'$ since $h[l, n]$'s for different l 's are independent. Consequently, we get ($\mathbf{w} = \mathbf{F}\mathbf{s}$):



$$\boldsymbol{\mu}_{\mathbf{w}} = \mathbf{0} \quad (\text{F.6})$$

and

$$\mathbf{C}_{\mathbf{w}\mathbf{w}} = \mathbf{F} \mathbf{C}_{\mathbf{s}\mathbf{s}} \mathbf{F}^H \quad (\text{F.7})$$

where $\mathbf{C}_{\mathbf{s}\mathbf{s}} = \mathbb{E}[\mathbf{s}\mathbf{s}^H]$ is a diagonal matrix with the l th diagonal entry given by $2(1-\alpha) \Xi_l / (N-1)^2$.

Appendix G

Calculation of (5.36)

Using the definition of LLR, we have

$$L(X[kG+j]=+1) = \lambda^{C,ext}[kG+j] - \ln(e^0 + e^{\lambda^{C,ext}[kG+j]}) \quad (G.1)$$

$$L(X[kG+j]=-1) = -\ln(1 + e^{\lambda^{C,ext}[kG+j]}) \quad (G.2)$$

The calculation of (G.1) and (G.2) can be simplified by using the rule:

$$\ln \sum_j e^{a_j} \approx \max_j a_j \quad (G.3)$$

When applied, straightforward manipulation yields

$$L(X[kG+j]=q) = \begin{cases} -\ln 2, & \text{if } \lambda^{C,ext}[kG+j] = 0 \\ \min(q\lambda^{C,ext}[kG+j], 0), & \text{otherwise} \end{cases} \quad (G.4)$$

where q is either $+1$ or -1 .

Vita

Meng-Lin Ku was born in Taoyuan, Taiwan, R.O.C. He received the B.S. degree in communication engineering from National Chiao Tung University, Hsinchu, Taiwan, in 2002. He pursued his M.S. and Ph.D. in the Department of Communication Engineering under the supervision of Professor Chia-Chi Huang in National Chiao Tung University, and was graduated in 2003 and 2009, respectively. Over the past several years of his graduate study, he built up very good academic background in the areas of both communication theory and applied mathematics. His current research interests are in the wide areas of wireless and mobile communications, optimization for communication engineering, and statistical signal processing.

Publication List

Journal Papers

1. Meng-Lin Ku and Chia-Chi Huang, "A Complementary Code Pilot-Based Transmitter Diversity Technique for OFDM Systems," *IEEE Trans. on Wireless Communications*, vol. 5, no. 2, pp. 504-508, Mar. 2006.
2. Meng-Lin Ku and Chia-Chi Huang, "A Derivation on the Equivalence between Newton's Method and DF-DFT-Based Method for Channel Estimation in OFDM Systems," *IEEE Trans. on Wireless Communications*, vol. 7, no. 10, pp. 3982-3987, Oct. 2008.
3. Meng-Lin Ku and Chia-Chi Huang, "A Refined Channel Estimation Method for STBC/OFDM Systems in High-Mobility Wireless Channels," *IEEE Trans. on Wireless Communications*, vol. 7, no. 11, pp. 4312-4320, Nov. 2008.
4. Hsiao-Yun Chen, Meng-Lin Ku, Shyh-Jye Jou and Chia-Chi Huang, "A Robust Channel Estimator for High-Mobility STBC-OFDM Systems," *IEEE Trans. on Circuits and Systems I*, accepted, June 2009.
5. Meng-Lin Ku, Wen-Chuan Chen and Chia-Chi Huang, "EM-based It-

erative Receivers for OFDM and BICM/OFDM Systems in Doubly Selective Channels,” *IEEE Trans. on Wireless Communications*, submitted, Apr. 2009.

Conference Papers

1. Meng-Lin Ku, Ching-Kai Li and Chia-Chi Huang, “A Novel Iterative Multi-layered Detection Method for MIMO Multi-code Multicarrier Systems,” in *Proc. 2006 National Symp. on Telecommun.*, Kaohsiung, Taiwan, 2006.
2. You-Tsai Jeng, Meng-Lin Ku and Chia-Chi Huang, “A Modified DVB-T System Architecture with Multi-carrier Multi-code Transmission and MPIC Based Reception,” in *Proc. IEEE Int. Conf. on Commun.*, Beijing, China, pp. 907-913, May 2008.



Patents

1. Meng-Lin Ku and Chia-Chi Huang (National Chiao-Tung University), “A Joint Channel Estimation and Data Detection Method for STBC/OFDM Systems,” filed for Taiwan patent, Publication No. 200830770, Jan. 2007.
2. Meng-Lin Ku and Chia-Chi Huang (National Chiao-Tung University), “A Joint Channel Estimation and Data Detection Method for STBC/OFDM Systems,” filed for US patent, Publication No. 20080165672, July 2008.