



Title: System capacity enhancement for 5g network
and beyond

Name: Hua Luo

This is a digitised version of a dissertation submitted to the University of
Bedfordshire.

It is available to view only.

This item is subject to copyright.

SYSTEM CAPACITY ENHANCEMENT FOR 5G
NETWORK AND BEYOND

HUA LUO

Ph.D

2017

UNIVERSITY OF BEDFORDSHIRE

SYSTEM CAPACITY ENHANCEMENT FOR 5G NETWORK AND BEYOND

by

Hua Luo

A thesis submitted to the University of Bedfordshire, in fulfilment of the requirements
for the degree of Doctor of Philosophy

August 2017

Academic Thesis: Declaration of Authorship

I, Hua Luo, declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

System Capacity Enhancement For 5G Network and Beyond

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have cited the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as indicated on Page iv.

Name of candidate: Hua Luo

Signature:

Date: 31/08/2017

Abstract

The demand for wireless digital data is dramatically increasing year over year. Wireless communication systems like Laptops, Smart phones, Tablets, Smart watch, Virtual Reality devices and so on are becoming an important part of people's daily life. The number of mobile devices is increasing at a very fast speed as well as the requirements for mobile devices such as super high-resolution image/video, fast download speed, very short latency and high reliability, which raise challenges to the existing wireless communication networks. Unlike the previous four generation communication networks, the fifth-generation (5G) wireless communication network includes many technologies such as millimetre-wave communication, massive multiple-input multiple-output (MIMO), visual light communication (VLC), heterogeneous network (HetNet) and so forth. Although 5G has not been standardised yet, these above technologies have been studied in both academia and industry and the goal of the research is to enhance and improve the system capacity for 5G networks and beyond by studying some key problems and providing some effective solutions existing in the above technologies from system implementation and hardware impairments' perspective.

The key problems studied in this thesis include interference cancellation in HetNet, impairments calibration for massive MIMO, channel state estimation for VLC, and low latency parallel Turbo decoding technique. Firstly, inter-cell interference in HetNet is studied and a cell specific reference signal (CRS) interference cancellation method is proposed to mitigate the performance degrade in enhanced inter-cell interference coordination (eICIC). This method takes carrier frequency offset (CFO) and timing offset (TO) of the user's received signal into account. By reconstructing the interfering signal and cancelling it afterwards, the capacity of HetNet is enhanced.

Secondly, for massive MIMO systems, the radio frequency (RF) impairments of the hardware will degrade the beamforming performance. When operated in time duplex

division (TDD) mode, a massive MIMO system relies on the reciprocity of the channel which can be broken by the transmitter and receiver RF impairments. Impairments calibration has been studied and a closed-loop reciprocity calibration method is proposed in this thesis. A test device (TD) is introduced in this calibration method that can estimate the transmitters' impairments over-the-air and feed the results back to the base station via the Internet. The uplink pilots sent by the TD can assist the BS receivers' impairment estimation. With both the uplink and downlink impairments estimates, the reciprocity calibration coefficients can be obtained. By computer simulation and lab experiment, the performance of the proposed method is evaluated.

Channel coding is an essential part of a wireless communication system which helps fight with noise and get correct information delivery. Turbo codes is one of the most reliable codes that has been used in many standards such as WiMAX and LTE. However, the decoding process of turbo codes is time-consuming and the decoding latency should be improved to meet the requirement of the future network. A reverse interleave address generator is proposed that can reduce the decoding time and a low latency parallel turbo decoder has been implemented on a FPGA platform. The simulation and experiment results prove the effectiveness of the address generator and show that there is a trade-off between latency and throughput with a limited hardware resource.

Apart from the above contributions, this thesis also investigated multi-user precoding for MIMO VLC systems. As a green and secure technology, VLC is achieving more and more attention and could become a part of 5G network especially for indoor communication. For indoor scenario, the MIMO VLC channel could be easily ill-conditioned. Hence, it is important to study the impact of the channel state to the precoding performance. A channel state estimation method is proposed based on the

signal to interference noise ratio (SINR) of the users' received signal. Simulation results show that it can enhance the capacity of the indoor MIMO VLC system.

Acknowledgements

During my doctoral period, many people have helped me with my research as well as my daily life. Firstly, I would like to give a lot of thanks to Dr. Yue Zhang and Prof. Dayou Li for their supervision and valuable advice on my research. Secondly, I also want to thank my colleague Dr. Wei Li for his help on many technical details in this research. Many special thanks to Dr. Li-ke Huang, Kexuan Sun and Hongyuan Huang in Cobham Wireless for their advice and support on building the testbed and experiment system in this research.

Moreover, I would like to give many thanks to my parents for their support of my study abroad. Most importantly, I want to thank my wife Sisi for her love and encouragement during this research. Her companion is one of the most important reason that I can finish my PhD.

Publications to date

1. H. Luo, W. Li, Y. Zhang, L. K. Huang, J. Cosmas and Q. Ni, "CRS interference cancellation algorithm for heterogeneous network," in *Electronics Letters*, vol. 52, no. 1, pp. 77-79, 18 2016.
2. H. Luo, Y. Zhang, L. K. Huang and J. Cosmas, "A new reciprocity calibration method for massive MIMO systems," *2016 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, Nara, 2016, pp. 1-4.
3. H. Luo, Y. Zhang, L. K. Huang, J. Cosmas and A. Aggoun, "A Closed-Loop Reciprocity Calibration Method for Massive MIMO in Terrestrial Broadcasting Systems," in *IEEE Transactions on Broadcasting*, vol. 63, no. 1, pp. 11-19, March 2017.
4. H. Luo, Y. Zhang, L. k. Huang and J. Cosmas, "Low latency turbo decoder implementation for future broadcasting systems," *2017 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, Cagliari, Italy, 2017, pp. 1-4.
5. H. Luo; Y. Zhang; W. Li; L. K. Huang; J. Cosmas; D. Li; C. Maple; X. Zhang, "Low Latency Parallel Turbo Decoding Implementation for Future Terrestrial Broadcasting Systems," in *IEEE Transactions on Broadcasting*, vol.64, no.1, pp.96-104, Mar. 2018

Contents

Abstract.....	i
Acknowledgement	iv
Publications to date	v
List of Figures	ix
List of Tables	xii
Abbreviations.....	xiii
1 Introduction.....	1
1.1 Research Background.....	1
1.1.1 1G-5G history	1
1.1.2 MmWave communication.....	5
1.1.3 Massive MIMO	6
1.1.4 Visual light communication.....	6
1.1.5 Heterogeneous network	7
1.2 Research Problem.....	8
1.2.1 Inter-cell interference cancellation	8
1.2.2 RF impairments calibration for massive MIMO.....	9
1.2.3 Low latency parallel decoding methods	10
1.2.4 Channel state estimation of VLC systems	11
1.3 Research objectives	13
1.4 Main contributions	15
1.5 Outline of the thesis.....	16
2 Background and state of the art	18
2.1 Wireless channel model	18
2.2 OFDM signal model.....	20
2.3 HetNet inter-cell interference background.....	24
2.3.1 LTE signal model.....	24
2.3.2 HetNet inter-cell interference	26
2.4 Beamforming and hardware impairments fundamentals	28
2.4.1 Beamforming	28
2.4.2 Precoding	31
2.4.3 Hardware impairments.....	37
2.5 Channel coding/decoding method.....	41

2.5.1	Background	41
2.5.2	Turbo encoding	43
2.5.3	Turbo decoding	49
2.6	VLC systems introduction.....	58
2.6.1	Architecture of VLC	59
2.6.2	Physical layer of VLC.....	60
2.6.3	Modulation techniques.....	61
3	CRS interference cancellation for Heterogeneous networks	68
3.1	Interference model.....	68
3.1.1	Time domain eICIC	68
3.1.2	Frequency domain eICIC	69
3.1.3	Power control techniques	70
3.2	CRS Interference cancellation.....	74
3.2.1	System model.....	74
3.2.2	TO/CFO estimation.....	75
3.2.3	Interfering channel estimation	76
3.2.4	Interfering signal reconstruction and reduction	78
3.3	Simulation results.....	79
3.4	Summary	83
4	Impairments Calibration for massive MIMO systems	84
4.1	Massive MIMO impairments	84
4.2	MIMO systems Hardware impairments model	87
4.2.1	Single antenna DCT impairments model	87
4.2.2	MIMO and massive MIMO systems hardware impairments model.....	90
4.3	Impairments calibration for MIMO systems	94
4.3.1	Transmit beamforming calibration	95
4.3.2	Reciprocity calibration for massive MIMO systems	96
4.4	Simulation and experiment results	102
4.4.1	Computer simulation.....	102
4.4.2	Lab experiment	110
4.5	Summary	118
5	Low latency Parallel Turbo decoding and its implementation	120
5.1	Parallel turbo decoding method.....	120

5.1.1	Parallel decoding structure.....	121
5.1.2	Parallel interleave/deinterleave	122
5.1.3	Proposed reverse address generator	124
5.2	Parallel turbo decoder implementation	127
5.2.1	Extraction.....	127
5.2.2	Extrinsic information	128
5.2.3	Interleaver/Deinterleaver	129
5.2.4	Modulo operation.....	130
5.2.5	Double buffering.....	131
5.3	Simulation and experiment results	132
5.4	Summary	139
6	Channel state estimation for indoor VLC MIMO Systems	141
6.1	VLC channel model	141
6.1.1	Indoor SISO Channel model.....	141
6.1.2	Indoor MIMO channel model	144
6.2	Channel state estimation.....	145
6.2.1	System model.....	146
6.2.2	Proposed CSE	147
6.3	Simulation results.....	149
6.4	Summary	153
7	Conclusions.....	154
7.1	Discussions.....	154
7.2	Further work.....	155
	Reference	157

List of Figures

Figure 1.1 1G-4G summery	4
Figure 1.2 5G expectation and scope	5
Figure 2.1 Propagation of wireless radio signal	19
Figure 2.2 OFDM signal structure	22
Figure 2.3 OFDM transmitter simplified structure	23
Figure 2.4 LTE frame structure	25
Figure 2.5 LTE resource grid	25
Figure 2.6 Heterogeneous network	27
Figure 2.7 A picocell working with a macrocell	28
Figure 2.8 LTE reference signal structure	28
Figure 2.9 Linear array beamforming	29
Figure 2.10 Transmit beamforming schematic	30
Figure 2.11 Beamforming result diagram	31
Figure 2.12 Point-to-point MIMO system	32
Figure 2.13 Multi-user MIMO system	34
Figure 2.14 Multi-user precoding	35
Figure 2.15 Direct conversion transceiver structure	39
Figure 2.16 Convolutional encoding	44
Figure 2.17 Turbo encoder: (a) Encoder structure; (b) Trellis diagram.	45
Figure 2.18 Interleaving	46
Figure 2.19 Turbo decoding structure	49
Figure 2.20 Decoding trellis	53
Figure 2.21 VLC spectrum	58
Figure 2.22 Physical layer of VLC systems	60
Figure 2.23 DC balanced OOK	62
Figure 2.24 CIE 1931 colour space chromaticity diagram	63
Figure 2.25 Block diagram of CSK modulation of VLC system	64
Figure 2.26 DCO-OFDM VLC system	65
Figure 2.27 DCO-OFDM output: (a) Before adding DC bias; (b) After adding DC bias and zero clipping	66
Figure 3.1 HetNet frequency sharing	70
Figure 3.2 Illustration of almost blank subframe	70

Figure 3.3 Non-colliding scenario (left) and colliding scenario (right).....	71
Figure 3.4 non-colliding CRS interference with TO/CFO.....	73
Figure 3.5 CRS interference cancellation scheme	75
Figure 3.6 BLER performance of different CFO and propagation delay:	81
Figure 3.7 BLER performance of different TO and propagation delay:.....	82
Figure 3.8 BLER performance of SNR.....	82
Figure 4.1 DCT demodulation model	90
Figure 4.2 Impaired transmit beamforming	92
Figure 4.3 MIMO precoding impairments model.....	93
Figure 4.4 Massive MIMO reciprocity and impairments model	94
Figure 4.5 Transmit beamforming calibration	95
Figure 4.6 TD based closed-loop reciprocity calibration diagram	100
Figure 4.7 OFDM frame and pilot structure	103
Figure 4.8 Antenna pattern comparison after beamforming in Cartesian coordinate: (a) Test case 1; (b) Test case 2	104
Figure 4.9 Antenna pattern comparison of beamforming in Polar coordinate: (a) Test case 1; (b) Test case2	105
Figure 4.10 Uplink simulation diagram	106
Figure 4.11 Downlink simulation diagram	106
Figure 4.12 Downlink frame structure for data transmission	107
Figure 4.13 Constellation comparison: (a) without calibration; (b) Relative calibration; (c) TD based calibration.....	108
Figure 4.14 EVM performance over SNR	109
Figure 4.15 EVM performance over different antenna number	110
Figure 4.16 Beamforming implementation setup	111
Figure 4.17 Beamforming calibration implementation: (a)(b) first subcarrier; (c)(d) middle subcarrier; (e)(f) last subcarrier	112
Figure 4.18 Experiment setup: (a) The experiment setup scheme; (b) The experiment implementation setup.	114
Figure 4.19 BS hardware impairments' amplitude: (a) transmitters' impairments' amplitude (b) receivers' impairments' amplitude.....	115
Figure 4.20 Phase impairments: (a) transmitters phase impairments; (b) receivers phase impairments	116

Figure 4.21 Constellation of captured data over radio: (a) The constellation using traditional relative calibration; (b) the constellation using proposed TD based calibration.	118
Figure 5.1 Parallel RSC decoding structure.....	122
Figure 5.2 Timing diagram of each iteration; (a) original sequence interleaving; (b) reversed sequence interleaving	125
Figure 5.3 Extraction of received data.....	128
Figure 5.4 Parallel extrinsic information calculation.....	128
Figure 5.5 Parallel Interleaving.....	130
Figure 5.6 Modulo function flowchart.....	131
Figure 5.7 Double buffering.	132
Figure 5.8 Testbed system structure.	133
Figure 5.9 FPGA utilization, $P = 8$, (a) Area utilization; (b) Resource utilization. ..	134
Figure 5.10 FPGA utilization, $P = 64$, (a) Area utilization; (b) Resource utilization.	135
Figure 5.11 FPGA utilization, $P = 8$, with 8 decoders, (a) Area utilization; (b) Resource utilization.	136
Figure 5.12 Decoding results capture	137
Figure 5.13 BLER simulation results.....	139
Figure 6.1 VLC channel model: (a)LOS; (b) Diffuse; (c)Hybrid VLC link.....	142
Figure 6.2 Indoor VLC configuration.....	142
Figure 6.3 VLC MU-MIMO configuration	145
Figure 6.4 Block diagram of precoding with CSE.....	147
Figure 6.5 Channel gain of a single transmitter	150
Figure 6.6 Channel gain of 4 transmitters.....	151
Figure 6.7 System capacity with different SNR	152
Figure 6.8 System capacity with different terminal number.....	152

List of Tables

Table 3-1 Simulation Configuration	79
Table 4-1 OFDM parameters	102
Table 4-2 Beamforming phase and amplitude parameters	105
Table 5-1 Throughput and latency comparison	136
Table 5-2 Modulo operation comparison.....	137
Table 6-1 Simulation parameters	150
Table 6-2 Simulation parameters of MIMO VLC	150

Abbreviations

3GPP	3rd Generation Partnership Project
ABS	Almost Blank Subframe
ADC	Analog-to-Digital Conversion
AMPS	Advanced Mobile Phone System
AP	Access Point
ASIC	Application Specific Integrated Circuits
AWGN	Additive White Gaussian Noise
BCJR	Bahl-Cocke-Jelinek-Raviv
BER	Bit Error Rate
BLER	Block Error Rate
BS	Base Station
CDMA	Code Division Multiple Access
CFO	Carrier Frequency Offset
CIM	Colour-Intensity Modulation
CIR	Channel Impulse Response
CP	Cyclic Prefix
C-RAN	Centralized/Cloud Radio Access Network
CSI	Channel State Information
CSE	Channel State Estimation
CSK	Colour Shift Keying
CRS	Cell Specific Reference Signal
DAC	Digital-to-Analog Convert
DCT	Direct conversion transceivers
DOA	Direction of Arrival
DPC	Dirty Paper Coding
DVB	Digital Television Broadcasting
EPA	Extended Pedestrian A model
EVA	Extended Vehicular A model
EVM	Error Vector Magnitude
eICIC	Enhanced Inter Cell Interference Coordination
FD	Frequency-Dependent
FDD	Frequency-Division Duplexing

FDM	Frequency Division Multiplexing
FDMA	Frequency division multiple access
FEC	Forward Error Correction
FeICIC	Further eICIC
FFT	Fast Fourier Transform
FI	Frequency-Independent
FOV	Field-of-View
FPGA	Field-programmable gate array
GSM	Global System for Mobile
HetNet	Heterogeneous Network
HSDPA	High-Speed Downlink Packet Access
ICI	Inter-Carrier Interference
IF	Intermediate Frequency
IFFT	Inverse Fast Fourier Transform
IP	Intellectual Property
IMT	International Mobile Telecommunications
ISI	Inter-Symbol Interference
ITU-R	International Telecommunications Union-Radio communications sector
LAN	Local Area Network
LD	Laser Diode
LDPC	Low Density Parity Check
LED	Light Emitting Diode
LLR	Log-Likelihood Ratio
LNA	Low Noise Amplifier
LO	Local oscillator
LS	Least-square
LTE	Long term Evolution
LUT	Look Up Table
MAC	Media Access Control
MAP	Maximum a Posteriori
MCS	Modulation and Coding Schemes
MIMO	Multiple-Input and Multiple-Output

MmWave	Millimetre-wave
MMSE	Minimum Mean-squared Error
MRT	Maximum Ratio Transmission
MUSIC	Multiple Signal Classification
NGMN	Next Generation Mobile Networks
OFDM	Orthogonal Frequency-Division Multiplexing
OFDMA	Orthogonal Frequency-Division Multiple Access
OOK	ON-OFF keying
OWC	Optical Communication
PA	Power Amplifier
PAPR	Peak-to-Average-Power Ratio
PBCH	Physical Broadcast Channel
PCIe	Peripheral Component Interconnect Express
PD	Photodiodes
PDCCH	Physical Downlink Control Channel
PDF	Probability Density Function
PDSCH	Physical Downlink Shared Channel
PHY	Physical Layer
PPM	Pulse Position Modulation
PPS	Pulse-per-Second
PMI	Precoding Matrix Indicator
PRB	Physical Resource Blocks
PSS	Primary Synchronization Signal
PWM	Pulse Width Modulation
QAM	Quadrature amplitude modulation
RAN	Radio Access Network
RB	Resource Block
RE	Resource Element
RF	Radio Frequency
RSC	Recursive Systematic Convolutional
SDMA	Space-Division Multiple Access
SFO	Sample Frequency Offset
SINR	Signal-to-Interference plus Noise Ratio

SMS	Short Message Service
SM	Spatial Multiplexing
SNR	Signal-to-Noise Ratio
SSS	Secondary Synchronization Signal
STO	Symbol Timing Offset
SISO	Single-input and single-output
TDD	Time-Division Duplexing
TO	Timing Offset
UE	User Equipment
ULA	Uniform Linear Array
UMTS	Universal Mobile Telecommunications System
VLC	Visual Light Communication
VLSI	Very-Large-Scale Integration
V5GTF	Verizon 5G Technology Forum
W-CDMA	Wideband Code Division Multiple Access
WiMAX	Worldwide Interoperability for Microwave Access
ZF	Zero Forcing

1 Introduction

1.1 Research Background

Nowadays, the demand for wireless digital data is dramatically increasing year over year. Wireless communication systems like Laptops, Smart phones, Tablets, Smart watch, Virtual Reality devices and so on are becoming an important part of people's daily life. Take mobile phone users as an example, the number of mobile phone users is increasing at a very fast speed, from 4.43 billion in 2015 to 4.61 billion in 2016 and is estimated to reach 5 billion in 2019 [1]. Along with the vast number of mobile users, the requirements for mobile devices such as super high-resolution image/video, fast download speed, very short latency and high reliability are still going to be the challenges for the existing wireless communication networks.

1.1.1 1G-5G history

The history of wireless communication network (cellular network) has been through several generations, from the first generation (1G) to the recent fourth (4G). The first-generation cellular system, named Advanced Mobile Phone System (AMPS), was introduced in the United States, Israel, and Australia in the 1980s. Although AMPS had several serious problems from the modern standard's perspective, it helped drive the development and usage of cellular technology. The second-generation cellular network (2G) was introduced in the 1990s, with two different standards, GSM and CDMA, which stand for the European Global System for Mobile Communications and the US Coding Division Multiple Access (CDMA), respectively. A new communication type called short message service (SMS) or text messaging was introduced in this generation. At the same time, apart from phone and SMS, 2G also gives the users the ability to access to media content such as the Internet. During the 2G period, the use of mobile phones became more and more popular in people's daily life. Moreover, the demand for data especially the Internet data started to grow. It became clear that the 2G technology cannot meet the demand of people's increasing

data usage which drove the development of the third cellular network (3G, Mobile broadband). Since then, data rate has become a most important target for the communication scientists. To meet the demand from the market, different types of technologies appeared which resulted in different standards supported by several countries and companies worldwide.

The first 3G network was launched in Japan by NTT DoCoMo in October 2001, using a new technology called Wide Band Code Division Multiple Access (W-CDMA). After that, another technology named CDMA2000 1xEV-DO was used in the 3G network by SK Telecom and KTF in South Korea and American Monet. CDMA2000, which represents Code Division Multiple Access, is a family of mobile technology first used in the middle of 2G (2.5G). 2.5G systems are the extensions of existing 2G systems at that time while 3G systems were still under development. For CDMA2000-1x, the ideal maximum data rate can reach up to 307 Kbit/s. And the connection time between the terminals and the base stations became very short which made it possible, for the first time, that media stream like radio or even television could be received by mobile handsets. In the mid-2000s, compared to CDMA2000 and WCDMA, an enhanced 3G, higher speed technology standard, High-Speed Downlink Packet Access (HSDPA), began to be implemented. It allows the communication network to be able to support much higher data rate and system capacity. The downlink speed of HSDPA can even reach up to 14.0 Mbit/s which gives the ability for the mobile phones to access the mobile website. Except for mobile phones, some network service devices appeared to use the 3G network to give Desktop computers wireless internet service. At this time, the wireless communication network is not just a traditional communication network only for voice and text service, but also a channel for unlimited Internet service. And it became very obvious that the wireless communication network in the future was going to cover every part of people's daily life.

The fourth generation of mobile communication network, 4G, was first deployed since 2009 in North Europe. The 4G standards were included in the IMT-Advanced specification. IMT-Advanced, stands for the International Mobile Telecommunications Advanced, was defined by the International Telecommunications Union-Radio communications sector (ITU-R). According to the 4G specifications, the peak data rate of service can be up to 100 Mbit/s for fast-moving scenarios and 1 Gbit/s for low-moving scenarios, which improved around 100 times compared to the 3G network. Long-Term Evolution (LTE) is one of the 4G standards. The capacity and data rate of LTE is improved greatly by utilizing different radio interface along with other network improvements. 3GPP (3rd Generation Partnership Project) is the main driver of LTE who specified different carrier frequencies (range from 1850 MHz to 3800 MHz) and bands for LTE in different countries. Unlike the previous generations, the terminals need to support multi-band to obtain LTE service in all countries. The original LTE technology cannot meet the requirements set by ITU-R at first. Therefore, LTE-Advanced was released so that the demand of the market could be met. The LTE-Advanced standards can support downlink peak data rate of 1000 Mbit/s and 500 Mbit/s uplink peak data rate with latency of only 1ms. This powerful ability enables LTE to manage and to support fast-moving mobile terminals and also multi-cast/broadcast data streams. There are two kinds of duplexing strategies in LTE, frequency division duplexing (FDD) and time-division duplexing (TDD). LTE-FDD and LTE-TDD use different spectrum and the way of uploading/downloading data is different as well. LTE-FDD uses paired frequencies while LTE-TDD uses single carrier frequency by separating uploading and downloading time. From the speed point of view, the 4G network has similar ability as local area network (LAN).

Figure 1.1 below shows the development history of 1G-4G, including the standards, frequency band and what they can support etc.





 <p>1G Year: early 80s Standards: AMPS,TACS Technology: Analog Peak speed: 14.4 kbps</p>	<p>3G Year: 2001 Standards: UMTS/HSPA Technology: Digital Bandwidth: Broad Band Peak speed: 14.4 Mbps Support: SMS/MMS, Internet Access, Video calls, Mobile TV</p> 
 <p>2G Year: 1991 Standards: GSM,GPRS,EDGE Technology: Digital Bandwidth: Narrow Band Peak speed: 171.2 Kbps Support: SMS/MMS</p>	<p>4G Year: 2010 Standards: LTE,LTE-Advanced Technology: Digital Bandwidth: Mobile Broad Band Peak Speed: 300Mbps Support: SMS/MMS, Internet Access, Video calls, Mobile TV, Gaming services, Cloud computing</p> 

Figure 1.1 1G-4G summary

Although 4G network can support a rather high system throughput, the demand for wireless communication from the market still keeps rising at a very fast speed. It can be easily imagined that several years later from now, the 4G network cannot meet the demand because of the requirements from not only the mobile phones but also other systems like high-resolution televisions, Virtual Reality devices, automotive communication systems and so forth. The next generation wireless communication network, 5G, is still under research and development. 5G aims to support much higher system capacity than the current 4G and serve much more mobile users in a single cell. Apart from the communication between the base stations and the terminals, 5G is believed to support device to device (D2D) communication, and the D2D could be the foundation technique of the Internet of Things (IoT). Although 5G technology has not been standardized yet at this moment, the Next Generation Mobile Networks (NGMN) Alliance, a mobile telecommunication association organized by several mobile operators, vendors, manufacturers and research institutes, defined some general requirements for 5G. These requirements can be separated into three dimensions: connections number, latency and system throughput.

A 5G network should support up to a million connections for an area of a square kilometre because of the demand of Everything on Mobile. Due to certain applications such as self-driving cars requiring very low latency, the 4G network cannot meet the demand, therefore, the latency of 5G should be less than 1ms. The system throughput of 5G should be 10 gigabits (Gbps) to better meet the market which needs to support many new demands like high-resolution video, Virtual Reality and so on. Figure 1.2 below shows the expectation and scope of the next generation communication network. Unlike 1G-4G, 5G is a kind of mobile broadband network supporting not only regular cell-to-device communication but also device-to-device. The system capacity is going to increase at least 100 times compared to 4G.

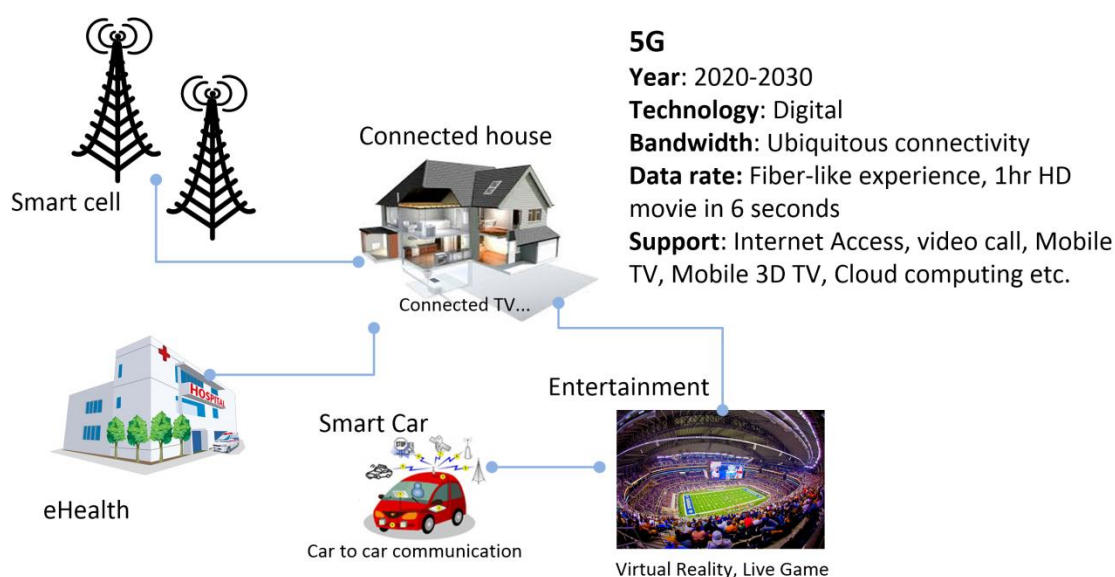


Figure 1.2 5G expectation and scope

5G network is expected to be deployed by 2020. Before that, many research institutes have reported their research results towards 5G. Among them, many new technologies and solutions for future mobile wireless communication have been proposed.

1.1.2 MmWave communication

Millimetre-wave (mmWave) wireless communication is a very well-known candidate that uses millimetre-wave as the carrier frequency. Due to the wide frequency band available in that spectrum, mmWave communication systems can easily reach a very

high data rate by using large bandwidth. Many research results about radio propagation measurements and channel models have been published [2][3]. Different mmWave frequency bands, 28, 38, 60 and 72-73 GHz, have been used to cover outdoor and indoor scenarios. Due to the much smaller wavelength of mmWave frequencies, more antennas can be configured at the base station which gives more flexibility when spatial signal processing techniques such as beamforming are utilized.

1.1.3 Massive MIMO

Massive multiple-input multiple-output (MIMO) is another potential technology for 5G with many (hundreds) antennas at the base station side that can work with multiple terminals at the same time. MIMO technology was first used in LTE-Advanced as a wireless communication technology which helped to improve the system capacity. A massive MIMO system is also known as a large-scale antenna system because the antenna number at the base station will increase significantly and is assumed to be 10 times larger than the terminals. In this way, a significant beamforming gain can be obtained, and more user terminals can be served simultaneously [53]. By beamforming, the transmission energy can focus to small regions which will improve both throughput and energy efficiency dramatically. Massive MIMO has many benefits such as inexpensive components, simplified MAC layer, and robust against interferences. These benefits will provide an energy-efficient, secure and robust future mobile broadband network. It also offers the ability to create a new digital society infrastructure by connecting people and things with clouds and other networks together.

1.1.4 Visual light communication

Visual light communication (VLC) is another promising alternative technology that can help meet the future wireless data demand especially for indoor communication scenarios [4]. VLC, also related with Li-Fi and optical communication (OWC), offers

services with several advantages over radio frequency (RF) systems. First, the spectrum ranging from 385 to 800 THz, gives a huge bandwidth to deliver license-free fast data services. Second, VLC has lower energy cost compared to RF because the visual light as a propagation media can also be used for illumination purposes. Third, since visual light cannot pass walls for indoor cases, the same spectrum can be reused in other spaces. This feature increases the spatial reusability and also creates a confidential and safe communication technology [5]. However, VLC suffers from the line of sight character of the channel when MIMO technology is applied. The MIMO VLC channel can easily become ill-conditioned which is a challenge to improve the system throughput [142].

1.1.5 Heterogeneous network

Since LTE Release 9, heterogeneous network (HetNet) has become a promising solution for mobile communication networks. HetNet is going to be a key technique to enhance the mobile network capacity, coverage performance and energy efficiency [4]. Different types of base station deployment, such as macrocells, microcells, picocells, and femtocells, has been deployed in the existing cellular systems. Based on these solutions, future 5G HetNet can also be integrated with current networks and new technologies such as massive MIMO and mmWave. For example, the outdoor point-to-point backhaul can use mmWave technology to maintain high volume data relay to support indoor high-rate wireless applications. Hence, the mmWave communication technology, as well as massive MIMO, can potentially combine with the existing HetNet technology and meet the future requirements of the 5G network together.

1.2 Research Problem

1.2.1 Inter-cell interference cancellation

While new technologies are being applied to the next generation communication network, some problems arise that need us to do future research on them. When using HetNet structure, the interferences among different types of cells becomes a concern. More specifically, when the macrocell is providing basic coverage, the picocell can provide services as a complementary cell. The off-load speed can be improved compared to the case that uses a single cell. However, since the macrocell and picocell are working together to serve the users, the signals they send to the users are the same frequency and bandwidth which cause inter-cell interferences between these two signals. The users suffer from the interferences by the neighbour high power macrocells, which is going to be even more serious if the users are within the coverage of macrocells [9]. Technically, a part of the reason for the inter-cell interferences is the system hardware non-ideality. The user terminal cannot be perfectly frequency synchronized with the base station which generates carrier frequency offset (CFO). The imperfect timing synchronization brings timing offset (TO). These two factors will affect the cancellation of those interferences hence need to be considered when designing inter-cell interference cancellation algorithms. To the best of the author's knowledge, there is no interference cancellation algorithm yet that considered CFO and TO.

Many people have studied the CRS inter-cell interference problem. In [41] and [42], the authors introduced a traditional CRS interference cancellation (IC) method that the interference channel is first estimated and then the cancellation of the interference is performed. A log-likelihood ratio muting algorithm was studied in [43] and [44]. A receiver algorithm was presented in [45] for the non-colliding scenario that the interference cancelation utilized the direct decision of channel estimation. In [46], a robust equalization scheme was proposed whose performance is similar with the

traditional method, but the complexity and computation latency are smaller. These methods are under the condition that the interfering CRS from the macrocell is aligned with the RE from the picocell both in time and frequency domain. However, practically, this condition cannot be met because of hardware impairments because the relative timing and carrier frequency offset will bring some error of the interference estimation which will make the overall system performance much poorer.

1.2.2 RF impairments calibration for massive MIMO

For massive MIMO systems, despite their huge advantages, there are still many challenges. One of them is caused by the RF impairments when low-precision/low-cost components working together that causes the downlink and uplink channel not reciprocal in time-division duplex operation. To serve multiple users simultaneously, massive MIMO systems rely on spatial multiplexing beamforming which requires a good knowledge of channel state information (CSI). In conventional MIMO systems like LTE, CSI is obtained by sending out pilots from the base station and feeding back the channel indicators by the terminals. This cannot be feasible when the number of antennas becomes very large as the time-frequency resources needed by the pilots will be a burden for the system [10].

One solution is to get the CSI is utilizing the reciprocity between the uplink and downlink channel in TDD mode. In this case, channel estimation is performed at the BS side while the pilots are sent by the terminals. The pilots from different terminal of different cell could cause a pilot contamination problem which is beyond the discussion of this work. Although the propagation channel itself is reciprocal, the hardware chains between the base station and terminal transceivers are not ideal nor reciprocal, especially when direct conversion transceivers (DCT) and cheap components are used to reduce the hardware costs of massive MIMO system [11]. Different specific levels of impairments also result in the MIMO channels not coherent with each other. As we know that beamforming techniques rely on the

adjustment of weightings, which is simply a factor of amplitude and phase, to control and switch the beam pattern. Generally speaking, the RF impairments can be represented by a factor of amplitude and phase [83]. Therefore, with RF impairments, the beamforming performance is going to be affected hence they need to be calibrated.

Many articles have discussed reciprocity calibration of massive MIMO systems. A relative calibration method was first proposed and implemented in [56]. The hardware impairments are generally equivalent to the frequency response of the hardware. R. Rogalin et al. in [55] studied an over-the-air calibration protocol which can scale well with the network size for distributed massive MIMO. A least square (LS) problem was generated for reciprocity calibration which took the noise effect into account. In [69], a hardware-based calibration setup, which can support online calibration, was studied. However, additional hardware-based circuitries are needed which will increase the cost of BS. All the above methods are all based on the "relatively accurate estimation", in which one of the antennas is used as a reference for calibration. Although this will not result in beam pattern distortion as mentioned in [56], it will degrade the performance (such as EVM) of the base station. A calibration method based on parameter estimation was carried out in [70] which can be performed without any aid of users or external devices. For each pair of transmit (TX) and receive (RX) hardware chains, a parameter was estimated as a function of analog transmit and receive frontends. However, the calibration accuracy is limited because the impairments' parameters of the TX and RX are not estimated separately.

1.2.3 Low latency parallel decoding methods

Apart from different new technologies, some existing techniques are still being researched to enhance the system capacity of the 5G network and beyond. Channel decoding is one technique that is under study. Among all the channel codes, Turbo code and Low Density Parity Check (LDPC) code are two of the most commonly used channel codes that have strong error correction capability with transmission

throughput close to Shannon channel capacity. However, the decoding algorithms are normally time-consuming. Take turbo decoding as an example, the Logarithmic Bahl-Cocke-Jelinek-Raviv (Log-BCJR) algorithm is used when decoding the received data. Due to its serial decoding nature because of the data dependencies of its forward-backward recursions [12], the decoding data rate is hard to meet the system demand. Therefore, it is essential to reduce the decoding process time by designing new decoding algorithms or new decoding implementation methods. Parallelization of traditional decoding is one of the practical and effective methods to reduce the decoding time and improve the decoding throughput.

Several parallel turbo decoders have been proposed previously, and most of them mainly tried to improve the level of parallelism to get a higher throughput and lower latency. In [101], a fully-parallel turbo decoder was implemented using analog decoder, but only short message lengths are supported. According to [102], a parallel turbo decoder algorithm that operates based on stochastic bit sequences was proposed which requires more processing time than Log-BCJR algorithm. A high performance parallel turbo decoder was introduced in [99] with configurable interleaving network which is implemented on very-large-scale integration (VLSI). A fully-parallel turbo decoding algorithm was studied in [103] which can support all LTE and WiMAX standards. However, the computing complexity is too high and is not practical for hardware platform like Field-programmable gate array (FPGA).

1.2.4 Channel state estimation of VLC systems

Using white light emitting diodes (LEDs) in VLC is gaining much attention in both academia and industry due to their advantages such as long lifetime, compact size and low power consumption especially for indoor communication [13][14]. However, the modulation bandwidth of the LED is very limited. Only several megahertz can be used because of the response speed of the phosphorescent component of visual light within the current LED production technology [15]. As a result, the throughput of

VLC systems is quite limited and not large enough to meet the modern communication demand. To get a higher throughput, spectral-efficient techniques such as orthogonal frequency division multiplexing (OFDM) and MIMO can be introduced in VLC. Since VLC is non-coherent form of communication, different forms of OFDM, such as DC biased optical OFDM (DCO-OFDM), are used [130]. LEDs illuminate the area in an indoor VLC system, in the meantime, MIMO techniques can help achieve parallel data transmission and spatial multiplexing (SM), thanks to the mature MIMO techniques in RF wireless communication [16]. The throughput can be greatly improved by using precoding technique, such as conjugate precoding, zero forcing (ZF) precoding, and minimum mean-squared error (MMSE) precoding, for multiuser MIMO VLC systems. Multiuser precoding for MIMO VLC systems requires the base station to know the condition of the channel, by which the performance of precoding methods is affected.

Many literatures have discussed multiuser precoding for MIMO VLC systems. In [139], linear ZF and ZF dirty paper coding (ZF-DPC) are studied for MU-MIMO broadcasting for indoor VLC system. The performances of these two techniques are compared. According to [140], a robust MMSE linear precoding is introduced with imperfect channel estimation taken into account. A multiuser MIMO-OFDM for VLC with ZF and MMSE precoding is studied in [141]. All these methods require the knowledge of the channel and the performances of these precoding methods will be affected by the condition of the channel. As mentioned in [141], zero forcing is a good approach for high power or low noise scenarios, but its performance will be reduced if the channel matrix is ill-conditioned. Furthermore, according to [142], the channel matrix can be ill conditioned when the receivers are deployed at the centre of the LEDs, along the axes or when the users are carrying out unconstrained movement. Hence, the channel of indoor MIMO VLC system can easily be ill-conditioned.

In summary, the research problems of this thesis are:

1). The existing inter-cell interference cancellation algorithms for HetNet systems did not take CFO and TO into consideration which can affect the system performance. For massive MIMO reciprocity calibration, the calibration accuracy needs to be improved without increase the cost of the BS and the calibration method should support online calibration.

2). Parallel decoding is an effective way to enhance the capacity of next generation communication network. For the 5G network, the strict demand of throughput and latency is a challenge for the existing decoding methods, hence parallel decoding needs to be addressed. The more parallelism of the decoder means more hardware resource are used. This research will try to improve the decoding throughput and latency performance and investigate the relationship between throughput, latency and hardware resource on FPGA.

3). MIMO VLC system is another subset of the 5G indoor network. Using multi-user beamforming to improve the system throughput is necessary. Since the indoor MIMO channel is easily ill-conditioned, the MIMO VLC channel estimation needs to be studied.

These problems are very important regard to the capacity of the 5G network and beyond. Of course, there are many other aspects of 5G since it includes more technologies than the previous generations, this thesis deals with the proceeding problems with effective and practical solutions.

1.3 Research objectives

The objective of this thesis is to give some practical solutions that can help improve the system capacity of the 5G communication network and beyond.

1) Design a novel inter-cell interference cancellation method for HetNet systems

The timing offset and carrier frequency offset between the macrocell, picocell and the user terminal are considered. The common reference symbols of the serving and dominant interfering cell are not overlapped for the non-colliding scenario [17]. The interference signal is rebuilt first and then be cancelled. Based on the statistics character of the channel estimation, the mean value of the interfering signal is close to 0 and the interfering channel can be estimated by a moving average. After reconstruction of the interfering signal, the interference can be cancelled simply by subtracting the interfering signal from the original received signal directly.

2) Propose a closed-loop reciprocity calibration method for massive MIMO systems

A closed loop reciprocity calibration method will be proposed to calibrate the non-reciprocity of massive MIMO systems caused by RF impairments. The error vector magnitude (EVM) performance can be improved by this method. The calibration method has higher calibration accuracy compared to the existing methods and can support online calibration as well.

3) Develop and implement a low latency, high throughput parallel Turbo decoder

A low latency, high throughput parallel Turbo decoder will be introduced in this dissertation. By improving the interleaving process, the time of the decoding recursion can be reduced, hence this decoder uses less decoding time and the throughput can be increased. The decoder is also implemented on a field-programmable gate array (FPGA) platform on which the performance of the parallel decoder can be measured precisely.

4) Develop a channel state estimation method to detect the real-time channel condition of MIMO VLC system

A channel state estimation (CSE) method will be proposed to choose a precoding method with better performance no matter if the channel condition is ill or not. Firstly,

the signal-to-interference-plus-noise ratios (SINRs) are evaluated for different precoding methods. Secondly, the SINRs are compared and the precoding method with a larger SINR is chosen. By using CSE, the system can achieve larger capacity compared to that with conjugate (CJ) precoding or minimum mean-squared error (MMSE) precoding only.

1.4 Main contributions

This thesis has the main contributions as below:

- 1) Proposed a novel CRS interference cancellation method for HetNet that can mitigate the effect of timing offset and carrier frequency offset. This method takes CFO and TO into account which, to the best of the author's knowledge, is by far the first.
- 2) Designed a new reciprocity calibration method for massive MIMO systems which can improve the EVM performance. The method is not only able to improve the calibration accuracy but also support online over-the-air calibration without raising the cost of hardware.
- 3) Designed and implemented a low latency parallel Turbo decoder. A reverse interleave address generator can reduce the decoding time hence the latency is decreased. The decoder is implemented on an FPGA platform in a reasonable complexity. The trade-off between latency and throughput with limited hardware resource is also addressed.
- 4) Proposed a channel condition estimating method for MIMO VLC systems. The SINR of the system can be improved as well as the capacity, by scheduling an appropriate precoding method according to the channel state, compared to the method with only one of the precoding method mentioned in this thesis.

1.5 Outline of the thesis

This thesis consists of 7 chapters. Apart from this chapter, the other 6 chapters are organized as below:

Chapter 2 gives a general background and state of the art of 5G systems and beyond. Since 5G is expected to consist of several different technologies like massive MIMO, mmWave and so forth, this chapter mainly focuses on the following aspects. Firstly, the basics of wireless channel model are introduced. Secondly, the OFDM signal model is presented. Thirdly, the background of HetNet inter-cell interference is introduced. Then, the background introduction of beamforming and hardware impairment, channel coding/decoding method and the VLC system are presented.

Chapter 3 carries out the research of inter-cell interference cancellation of HetNet. A novel interference cancellation method will be introduced. This method is mainly based on the reconstruction of the interfering signal. Therefore, the interfering signal reconstruction, along with the interfering channel estimation will be given. Simulation results tell that the proposed method is robust and effective. Some passages in this chapter have been quoted verbatim from the author's published paper which can be found in [9].

Chapter 4 studies the reciprocity calibration of massive MIMO systems. A new reciprocity calibration method is presented in this chapter. This method includes offline-over-cable calibration and over-the-air online calibration. Inspired by traditional relative calibration, the first part of calibration, offline-over-cable calibration uses the same calibration algorithm but with a better reference hardware chain. In the second part, a test device is introduced to support online, over-the-air calibration. The calibration accuracy of the method will be improved compared to the existing methods, hence the performance such as EVM is better. What's more, with this calibration method, the BS does not need any additional hardware parts, which

can reduce the BS system cost. Some passages in this chapter have been quoted verbatim from the author's published paper which can be found in [11].

Chapter 5 presents the research of the parallel decoding methods. A low latency parallel Turbo decoder is introduced as well as the implementation of the proposed decoder. Based on the testbed system used, the latency of the parallel Turbo decoder after implementation can be less than 100 microseconds (us) and the throughput can reach up to multi-Gbps. More specifically, a reverse interleave address generator will be proposed which can reduce the iteration time. And a practical method of modulo operation is designed to reduce the computation complexity in FPGA. Some passages in this chapter have been quoted verbatim from the author's published paper which can be found in [145][146].

Chapter 6 investigates multi-user beamforming of visual light communication systems. A channel state detecting method will be proposed to detect the real-time channel state between the VLC transmitters and the receivers. This detecting method can help to enhance the system performance by evaluating the condition of the MIMO VLC channel and decide which precoding method to choose to obtain a better SINR. Using eigen decomposition of the MIMO channel estimation, the SINR of the channel can be evaluated. Choosing the precoding method with better SINR performance means that the system capacity will be larger with this channel state decoding method.

2 Background and state of the art

As discussed before, 5G network and beyond will not be just the evolution of 4G. Many technologies, such as HetNet, massive MIMO, mmWave communication and so on, are going to be included to define 5G. Since the definition of the standards have yet to be completed, the exact physical layer model of 5G remains unknown.

Based on the objectives of this thesis, this chapter mainly focuses on the following parts as a foundation of further research. First of all, the wireless channel model corresponding to different technologies will be introduced. Second, from the baseband signal model's perspective, the OFDM signal model is presented. After that, the background of HetNet inter-cell interference, beamforming and hardware impairments, channel coding/decoding method, and the introduction of VLC systems will be briefly given out.

2.1 Wireless channel model

The mobile wireless channel can be defined as the variations of the channel gain over time and frequency. Two kinds of channel are normally found to describe a wireless channel according the extent of fading, large-scale fading channel and small-scale fading [18]. Large-scale fading mainly happens when the radio signal is blocked and reflected by large objects like a building which is normally frequency independent. Small-scale fading, which is normally frequency dependent, appears when the signal is reflected by multiple smaller objects and results in multiple paths signal. These multiple paths signals can be constructive or destructive.

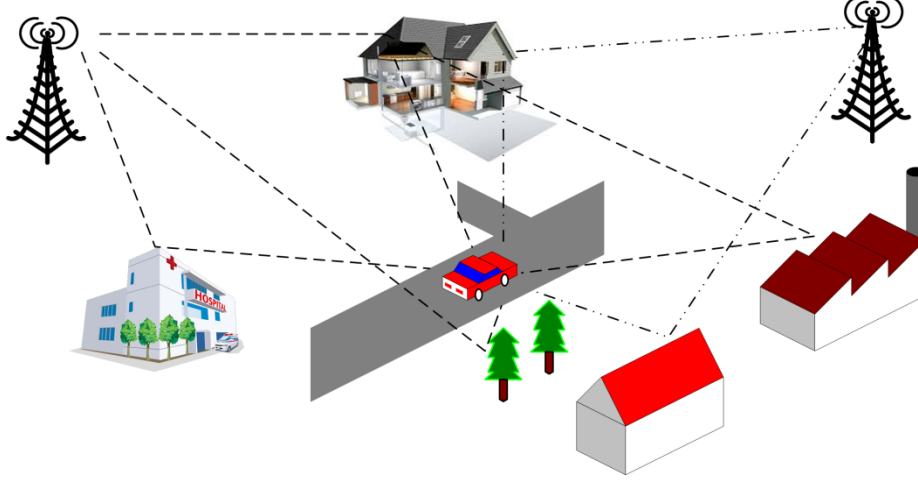


Figure 2.1 Propagation of wireless radio signal

A system function can be defined for a wireless channel as [2]:

$$H(f) = \frac{\alpha(\theta, \psi, f) e^{-j2\pi f r / c}}{r}, \quad (2.1)$$

where $\alpha(\theta, \psi, f)$ is the product of transmit antenna pattern and receive antenna pattern in a given direction; r represents the distance between the transmitter and receiver; f stands for the carrier frequency; and c is a constant value of the speed of light.

The channel model can also be viewed as a filter and a function of time t and propagation delay τ can describe its channel impulse response (CIR) [19].

$$h(\tau, t) = \sum_{k=1}^{K(t)} \alpha_k(t) e^{-j2\pi f \tau_k(t)} \delta(\tau - \tau_k(t)), \quad (2.2)$$

where $\tau_k(t)$ is the delay of the k^{th} path signal from the transmitter to the receiver; $\alpha_k(t)$ represents the attenuation level for the signal on the k^{th} path; $K(t)$ means the effective number of signal paths at the time t . From the statistics perspective, a wireless channel model can be described by its statistic character. In the case that there are many reflection objects in the channel and there exist many signal paths, the

CIR can be modelled using a complex Gaussian process. When there are many small reflectors, the probability density is close to Gaussian and the magnitude of $h(\tau, t)$ is a Rayleigh random variable, which is also called Rayleigh fading.

In another case, the transmit signal passes through several different paths with one path much stronger than the others, which is also named line of sight (LOS) signal. This model is called Rician fading because the gain of $h(\tau, t)$ is similar with Rician distribution. Compared to Rician fading, Rayleigh fading is a special case where there is no line of sight signal. For example, the channel of indoor VLC system is a Rician fading channel due to the strong direct signal between the transmitter and receiver. While for radio communication in an urban area, it is normally a Rayleigh fading channel due to multiple reflections of those buildings.

From the frequency-selective point of view, a wireless channel also includes flat fading and frequency-selective fading. Generally speaking, flat fading means the channel's frequency response is flat so that all the frequency-domain components are passed to the receiver side with same magnitude fading. However, the frequency components can also be affected independently and experience uncorrelated fading, makes the channel frequency-selective. To deal with this scenario, orthogonal frequency-division multiplexing (OFDM) can be utilized to provide robust service to fading.

2.2 OFDM signal model

OFDM is a modulation technique that encodes the information on multiple carrier frequencies. OFDM is a special FDM scheme that needs all sub-carriers to be orthogonal to each other. By this orthogonality, the cross-talk between the sub-carriers is removed, which simplifies the transceiver design. Multiple data sub-streams are modulated on these sub-carriers simultaneously and in parallel at a low symbol rate. Because of the independence of each sub-stream, OFDM can cope with

severe channel condition, such as frequency-selective fading without complex equalization. It has become very popular in wideband digital communication like digital television broadcasting (DVB), DSL Internet access, 4G communication network and so forth. And it will be the main modulation scheme for 5G network and beyond.

An time-domain FDM signal can be expressed as:

$$x(t) = \sum_{i=0}^{N-1} X_i \Phi_i(t), \quad (2.3)$$

where X_i is the coefficient of i th subcarrier; i and N denotes the index of a subcarrier and the number that subcarrier respectively. $\Phi_i(t)$ represents the i th subcarrier whose time-domain representation can be given as:

$$\Phi_i(t) = \omega(t) e^{j\pi f_i t}, \quad (2.4)$$

where $\omega(t)$ is its window function and f_i denotes the sub-carrier frequency. For an OFDM signal, a rectangular window is used as the window function with the length of T which is also the symbol period. To make the subcarriers orthogonal to each other, the following condition needs to be satisfied:

$$\int_{-\infty}^{\infty} \Phi_l(t) \Phi_l^*(t) dt = \begin{cases} 1 & \text{for } l = m \\ 0 & \text{for } l \neq m \end{cases}. \quad (2.5)$$

Replace $\Phi_l(t)$ by (2.4), (2.5) will be:

$$\begin{aligned} \int_{-\infty}^{\infty} \Phi_l(t) \Phi_l^*(t) dt &= \frac{1}{T} \int_{-T}^T e^{j2\pi(f_l - f_m)t} dt \\ &= \frac{\sin(\pi(f_l - f_m)T)}{\pi(f_l - f_m)T} . \\ &= \begin{cases} 1 & \text{for } l = m \\ 0 & \text{for } l \neq m \end{cases} \end{aligned} \quad (2.6)$$

From the (2.6) above, we can see that to make the subcarriers orthogonal, the subcarrier frequency difference should be a time of $1/T$. Or stated in another way, the spacing between two subcarriers should be n/T , where n is an integer. Figure 2.2 below shows the subcarrier relationship of an OFDM system in the frequency domain. Point A represents the peak amplitude of the i th subcarrier, while the value of Point B is zero which means there is no interference from other sub-channels. The spacing between Point B and C is the subcarrier spacing $1/T$, which can be viewed as the bandwidth of a subcarrier.

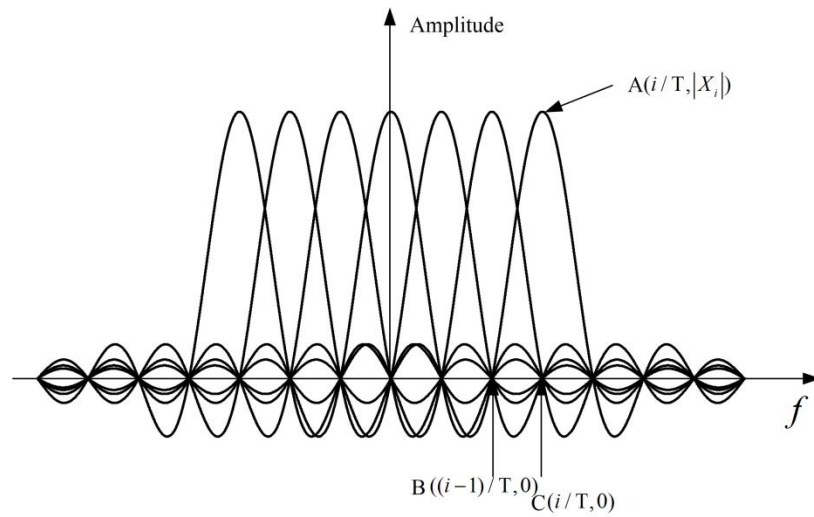


Figure 2.2 OFDM signal structure

From the mathematical description's perspective, an OFDM signal is the summation of multiple *sinc* functions where each *sinc* function shifts by i/T in the frequency domain. Since the null position of a *sinc* function lies at the point of i/T , the summation of different subcarriers at the null point is still zero, therefore, there is no interference between every subcarrier. This is how the orthogonality succeeds.

However, OFDM technology suffers from inter-symbol interference (ISI) that is because of the spreading of a symbol pulse in the time domain, especially in multipath propagation. ISI introduces errors at the receiver side, therefore, to minimize the errors, a guard interval is inserted between each symbol. The cyclic prefix of an

OFDM symbol is normally used as the guard interval so that the receiver can easily perform OFDM demodulation with FFT. A cyclic prefix of length T_g is put at the beginning of an OFDM symbol between time interval $-T_g \leq t < 0$, and it equals the signal from the “tail” of the OFDM signal between time interval $T - T_g \leq t < T$. The basic structure of an OFDM transceiver is shown below in Figure 2.3. The inverse discrete Fourier transform (IDFT) is utilized to modulate the data sub-stream while the discrete Fourier transform (DFT) is used for the demodulating the data sub-streams. The guard interval is inserted after IDFT at the transmitter side while the guard interval is removed before DFT at the receiver side.

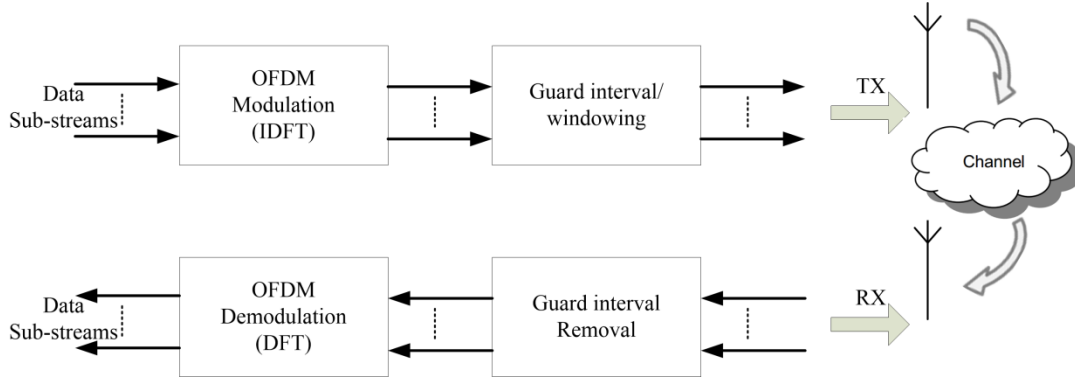


Figure 2.3 OFDM transmitter simplified structure

Taking the channel model in Section 2.1 into consideration, the received OFDM signal can be modelled as:

$$y(t) = h(\tau(t), t) * x(t) + n(t), \quad (2.7)$$

where $n(t)$ is the time variant white noise and “*” stands for convolution. The received digital signal is passed to the Fast Fourier Transform (FFT) module which converts the received data to the frequency domain $Y(n) = H(n) \cdot X(n) + N(n)$. The original transmitted OFDM signal estimation is given as:

$$\hat{X}(n) \approx Y(n) / \hat{H}(n), \quad (2.8)$$

where $\hat{H}(n)$ is the channel estimation frequency response.

OFDM has many advantages compared to other modulation technologies, such as high spectral efficiency, being robust against severe channel conditions and ISI, efficient implementation and so on. Nevertheless, it is very sensitive to Doppler shift and frequency synchronization error. The frequency synchronization error due to carrier frequency offset will break the orthogonality and introduce inter-carrier interference (ICI) to the system. Non-linearity components like an amplifier can cause intermodulation distortion which may raise inter-carrier interference as well.

Generally speaking, the performance of OFDM systems can be degraded by hardware impairments. Carrier frequency offset, due to the carrier frequency difference between the transmitter and receiver and Doppler shift, destroys the orthogonality. Timing offset (or Timing error), generated from timing synchronization, will introduce ICI as well as ISI. I/Q imbalance, produced by RF front-end, will cause a sub-carrier being interfered by its mirror sub-carrier. DC offset, caused by the carrier leakage of RF front-end, will break the orthogonality and result in severe ICI. Phase noise from the oscillator also brings ICI and makes the constellation fuzzy. The above kinds of hardware impairments are the basic types that affect OFDM performance, other impairments like power amplifier distortion also bring influence on the performance.

2.3 HetNet inter-cell interference background

2.3.1 LTE signal model

As the 5G heterogeneous network is not standardized yet, the interference model in this dissertation is based on LTE standard. The first step of this section gives the signal structure of LTE downlink signal.

The 3GPP LTE downlink uses Orthogonal Frequency Division Multiple Access (OFDMA) that performs very well in terms of efficiency and latency. A specific number of subcarriers, which are also referred to as physical resource blocks (PRBs) for a predefined period, are allocated to the users [38][39]. PRBs are described on

both time and frequency domain. The time domain frame structure is shown in Figure 2.4 below [39]. One radio frame, of which the whole duration is 10 milliseconds, is consisted of two subframes. Each subframe includes two slots, where each slot is 0.5 millisecond. For the short cyclic prefix scenario, each slot is made up of seven symbols and their corresponding cyclic prefix (CP).

To get a better understanding of LTE frequency structure, the downlink resource grid is given in [38][39]. According to the LTE specification, the whole bandwidth is 20MHz and can be divided into 100 physical resource blocks. The bandwidth of each PRB is 180KHz, which includes 12 subcarriers and each subcarrier is 15KHz. One resource element represents a single subcarrier for one symbol.

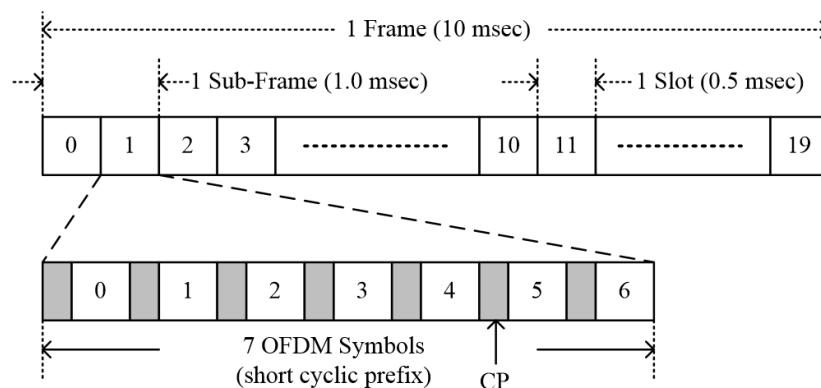


Figure 2.4 LTE frame structure

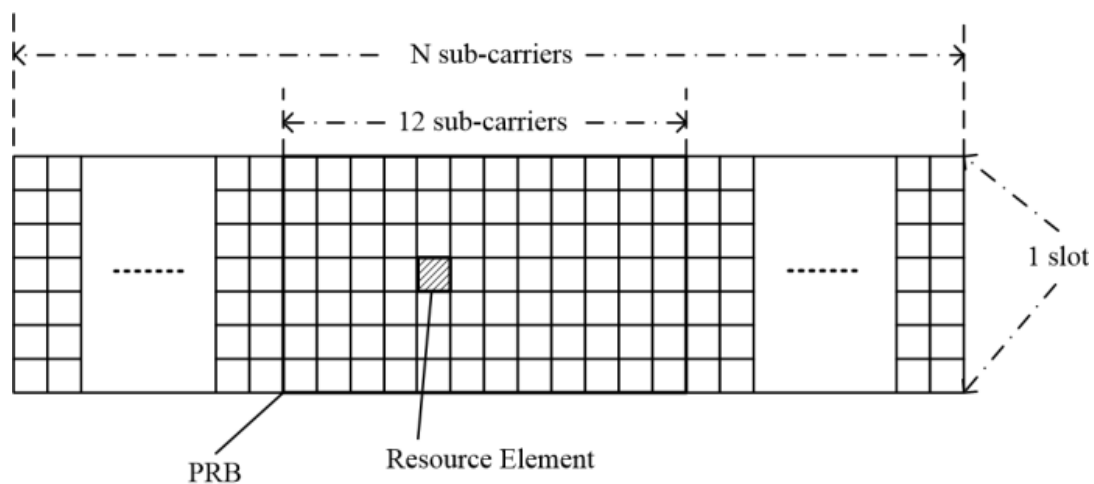


Figure 2.5 LTE resource grid

To demodulate the received signal properly, some reference signals are designed in LTE to perform timing synchronization, carrier offset estimation, channel estimation and so forth. For example, when short CP is used, reference signals are transmitted during the 1st and 5th OFDM symbol of every slot and the reference REs are inserted in the reference symbol at every sixth subcarrier. The channel response can be obtained based on these reference symbols. The reference structure can be seen in Figure 2.8 below [39].

2.3.2 HetNet inter-cell interference

The heterogeneous network was introduced to release data traffic pressure and can extend the network coverage range. A HetNet normally consists a high power macrocell and some low power small cells, such as femtocells, picocells, and microcells, see Figure 2.6. A macrocell is a high power (1~40W) cellular base station that can cover a larger area (around 5 kilometres) than small cells. The antennas of a macrocell are deployed in a high position like ground-based masts or on some existing structure such as a rooftop so that it can cover a clear view over the surrounding areas. A macrocell is able to support a wide range so it normally can be found in rural areas. At the same time, small cells are deployed to support a smaller cell area.

For a densely populated urban area, a microcell can be placed to provide network services. For areas that even smaller still such as a large office, a shopping mall or train station, a picocell can be used. What's more, for an indoor scenario like a home or small office, a femtocell, which is known to give the smallest coverage, can be arranged. By controlling the transmitting power of a microcell, the coverage area (typically less than two kilometres wide) can be adjusted. A picocell is a low-cost, small size base station that covers a small area (200 metres or less) like shopping malls, stock exchanges, in-airplanes and etc. Picocells are normally used to cover

indoor areas that outdoor coverage does not support well and add network capacity to some dense phone usage areas. Compared to picocells, a femtocell is even smaller and designed for indoor usage mainly, which is also known as femto access point (AP). They are typically connected to the wireless network via broadband and can support around 10 mobile phones. A home eNodeB nowadays is actually a 3G femtocell or a 4G LTE femtocell.

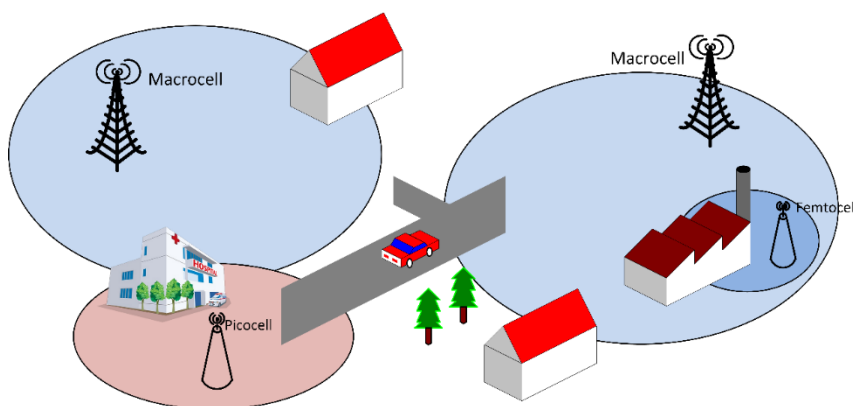


Figure 2.6 Heterogeneous network

For a heterogeneous network, there are many area requirements where combinations of macrocells, microcells, picocells and femtocells can work together. This dissertation mainly studies the scenario where a macrocell works with a picocell that the macrocell provides basic coverage while the picocell acts as a complementary cell. Cooperating with the macrocell, the picocell is able to extend the coverage, improve data capacity and spectrum efficiency compared to the network using macrocells only. It reduces the overall cost and also provides flexibility to the operators that they can deploy the picocells according to the market while it is not practical for macrocells. In this scenario, the user terminals served by the picocell can also receive a signal from the macro cell, which is a common problem for HetNet called inter-cell interference. The interference becomes more severe if the users stay in the range extension region [37] as shown in Figure 2.7 below.

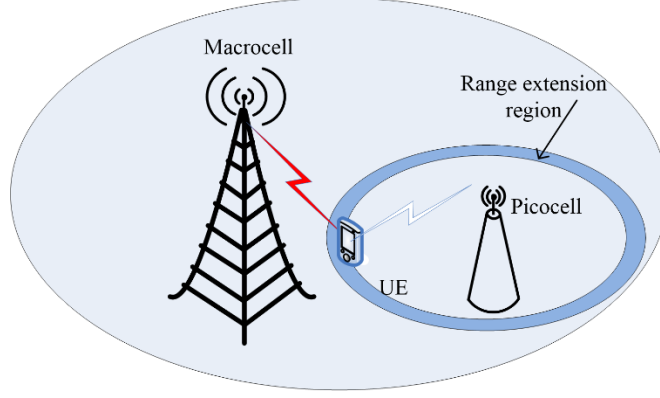


Figure 2.7 A picocell working with a macrocell

To deal with this issue, an interference mitigation scheme named enhanced inter-cell interference coordination (eICIC) was proposed in LTE release 10.

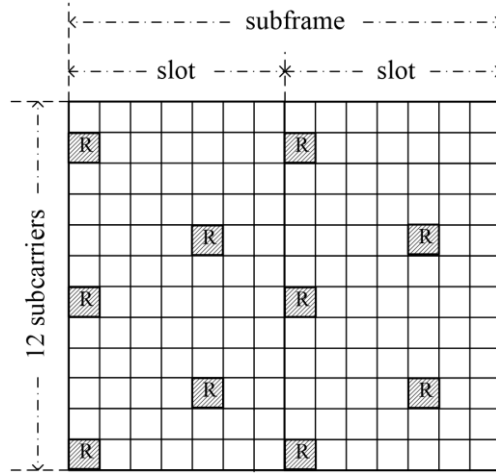


Figure 2.8 LTE reference signal structure

2.4 Beamforming and hardware impairments fundamentals

2.4.1 Beamforming

Beamforming is an array signal processing algorithm and the MIMO system antennas are a group of sensors or a sensor array, which is usually configured in a certain geometry pattern such as a uniform linear array (ULA), circular array, a rectangular array, or cylindrical array. There are sensor arrays where the sensor spacing is arbitrary, but only uniform sensor spacing is considered here. Beamforming requires the sensor spacing be smaller than half of the wavelength $\lambda/2$ so that the combined

antenna pattern can construct a more narrow and directional beam pattern. As a receiver, the receive beamforming can be accomplished by several beamforming algorithms such as delay-and-sum beamforming, spectrum-based beamforming, Multiple Signal Classification (MUSIC) beamforming and so on. The receive beamforming improves the SNR of the received signal and reduces the interference from other directions.

When beamforming is used at the transmitter side, the delay-and-sum beamforming is normally used to generate a narrow beam pattern and focus the maximum power to the target direction. The linear array transmitting diagram is shown in Figure 2.9 below:

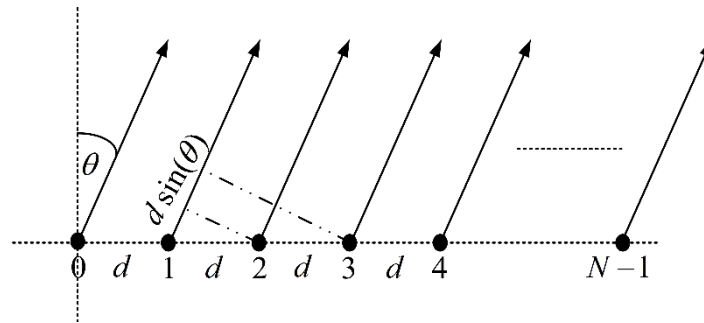


Figure 2.9 Linear array beamforming

Each element of the antenna array is transmitting RF signal simultaneously with a different delay. According to the electromagnetic theory, the wavefront of each element can accumulate together at the direction θ if the certain delay is applied and generate a narrower beam. The array propagation vector can be defined as [58]:

$$\mathbf{v} = \left[1 \quad e^{j\xi d \sin \theta} \quad \dots \quad e^{j\xi nd \sin \theta} \quad \dots \quad e^{j(N-1)\xi d \sin \theta} \right]^T, \xi = 2\pi / \lambda, \quad (2.9)$$

where the antenna 0 is used as a reference. And the weight vector of beamforming can be defined as:

$$\mathbf{w} = \left[w_0 \quad w_1 \quad \dots \quad w_n \quad \dots \quad w_{N-1} \right]^T, w_n = P_n e^{jn\alpha}, \quad (2.10)$$

where P_n and $n\alpha$ are the amplitude and phase factor of each antenna element weighting. The beamforming schematic is shown in Figure 2.10 below:

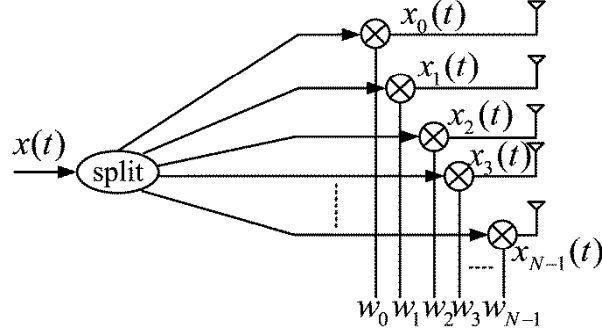


Figure 2.10 Transmit beamforming schematic

The single stream $x(t)$ is split into N sub-streams which are multiplied by different weightings respectively:

$$\mathbf{x}(t) = [x_0(t) \ x_1(t) \ \cdots \ x_{N-1}(t)] = \mathbf{w}x(t). \quad (2.11)$$

Suppose that each antenna is omnidirectional, the overall beam pattern of this array will be:

$$F(\theta) = \sum_{n=0}^{N-1} P_n e^{j(\xi n d \sin \theta + n\alpha)}. \quad (2.12)$$

At direction θ_0 , the emitted power of the antenna array is:

$$F(\theta_0) = \sum_{n=0}^{N-1} P_n e^{j(\xi n d \sin \theta_0 + n\alpha)}.$$

To make the emitted power at direction θ_0 maximum, α needs to meet the condition that $\alpha = -\xi d \sin \theta_0$. Hence, by controlling the weighting factors, the direction of beam pattern can be adjusted. As long as the direction of the UE is known, which can be estimated via direction of arrival (DOA) algorithm, the base station can adjust the transmit antenna beam to point to the UE and maximize the received signal SNR

which can enhance the connectivity between the UE and base station as shown in Figure 2.11 below. This beamforming model can be expanded to other geometry configuration as well.

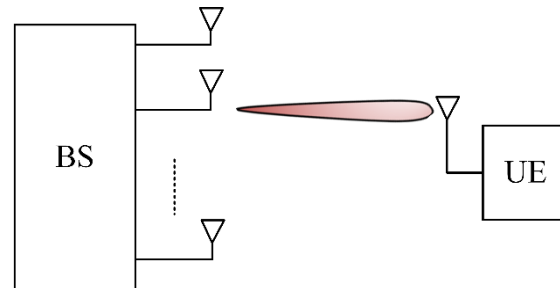


Figure 2.11 Beamforming result diagram

2.4.2 Precoding

Compared to beamforming, precoding is more general, and it can support multiple streams in MIMO systems. When the receiver is equipped with multiple antennas, traditional beamforming is not able to maximize the SNR of all the antennas' received signal. Hence precoding is designed not just to transmit multi-streams but also maximize the overall system capacity for the point-to-point MIMO system. Another more common scenario is that the receiver has only one antenna which is also known as multi-user MIMO. In this situation, the data stream to each user is independent and different with one and another. The throughput of a multi-user MIMO system with precoding can be maximized.

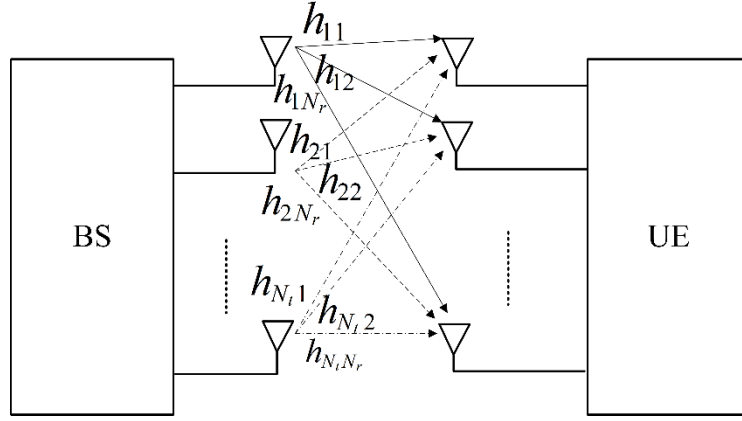


Figure 2.12 Point-to-point MIMO system

A point-to-point MIMO system is a single user MIMO system where the transmitter has N_t antennas communicating with the receiver which has N_r antennas, as shown in Figure 2.12 above. Precoding assumes that the channel is slow fading which means that the channel can be represented by a channel matrix, which represents the $N_t N_r$ paths between the transmitter and the receiver, for a certain period of time. If the channel state information can be obtained, precoding can achieve the MIMO capacity based on the known channel matrix. The signal model of a MIMO system can be given as:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \quad (2.13)$$

where \mathbf{y} and \mathbf{x} denote the receive and transmit signals respectively; \mathbf{H} and \mathbf{n} represent the channel matrix and the noise respectively. If the channel state information is perfectly known to both the transmitter and receiver, the channel capacity will be [59]:

$$C_{p\text{-}CSI} = E \left(\max_{\mathbf{Q}: \text{tr}(\mathbf{Q}) \leq 1} \log_2 \det(\mathbf{I} + \rho \mathbf{H} \mathbf{Q} \mathbf{H}^H) \right), \quad (2.14)$$

where $()^H$ represents the Hermitian transpose; ρ means the SNR of the received signal. \mathbf{Q} denotes the signal covariance.

If the transmitter does not know the channel perfectly, but the statistical character of the channel is known, the channel capacity will be reduced which can be modelled as [59]:

$$C_{s-CSI} = \max_{\mathbf{Q}} E \left(\log_2 \det(\mathbf{I} + \rho \mathbf{H} \mathbf{Q} \mathbf{H}^H) \right). \quad (2.15)$$

Another situation is where there is no CSI available and the capacity will be:

$$C_{no-CSI} = E \left(\log_2 \det \left(\mathbf{I} + \frac{\rho}{N_t} \mathbf{H} \mathbf{H}^H \right) \right). \quad (2.16)$$

According to the above equations, the MIMO channel capacity depends on the knowledge of the channel state information.

Multi-user MIMO can simultaneously support multiple receivers, which have one or multiple antennas, with only one multi-antenna transmitter. The system configuration diagram is shown in Figure 2.13 below. Precoding for multi-user is also known as space-division multiple access (SDMA) that includes two types, linear precoding and nonlinear precoding. Nonlinear precoding algorithms like dirty paper coding (DPC) are optimized algorithms that aim to achieve the channel capacity but with non-practical computation complexity. Linear precoding algorithms like zero-forcing (ZF) precoding and maximum ratio transmission (MRT) can achieve a reasonable performance but with much lower and practical complexity.

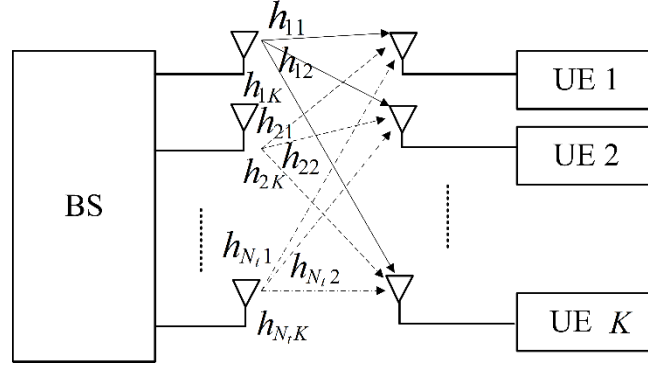


Figure 2.13 Multi-user MIMO system

Multi-user MIMO precoding also requires the knowledge of the CSI. Different level of knowledge will result in different channel capacity. Consider a multi-user MIMO system with N_t transmit antennas and K single-antenna users, the system model can be described as:

$$y_k = \mathbf{h}_k^H \mathbf{x} + n_k, k = 1, 2, \dots, K, \quad (2.17)$$

where \mathbf{h}_k is a $N_t \times 1$ vector that represents the channel coefficients from the N_t transmit antennas to the user k ; \mathbf{x} denotes the $N_t \times 1$ transmitted signal vector after precoding; y_k and n_k are the received signal and additional noise of user k , respectively. When linear precoding is applied to the transmitter, the transmitted signal vector can be represented by [60]:

$$\mathbf{x} = \sum_{j=1}^K \mathbf{w}_j s_j, \quad (2.18)$$

where \mathbf{w}_j and s_j denote the $N_t \times 1$ linear precoding vector and the data symbol respectively. The precoding system diagram is shown in Figure 2.14 below.

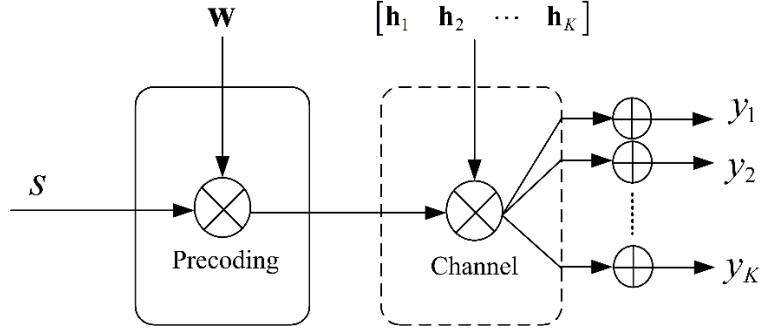


Figure 2.14 Multi-user precoding

In a multi-user MIMO system, the signal of one user can be treated as an interference to other users. Hence, signal-to-interference-and-noise ratio (SINR) is an important factor that can represent the performance of a single channel of a multi-user MIMO system. The SINR at user k can be given as [61]:

$$\text{SINR}_k = \frac{|\mathbf{h}_k^H \mathbf{w}_k|^2}{1 + \sum_{j \neq k} |\mathbf{h}_k^H \mathbf{w}_j|^2}. \quad (2.19)$$

Based on the SINR of each user, the corresponding overall channel capacity of the multi-user MIMO system will be:

$$C_{\text{multi-user}} = \sum_{k=1}^K \log_2(1 + \text{SINR}_k). \quad (2.20)$$

The above precoding analysis is based on the assumption that knowledge of CSI is perfectly known. In practice, the knowledge of CSI is limited and will result in system capacity loss. Hence (2.17) should be modified as:

$$y_k = \mathbf{h}_k^H \sum_{j=1}^K \hat{\mathbf{w}}_j s_j + n_k, k = 1, 2, \dots, K, \quad (2.21)$$

where $\hat{\mathbf{w}}_j$ is the distorted precoding vector and $\hat{\mathbf{w}}_j = \mathbf{w}_j + \mathbf{e}_j$; \mathbf{e}_j is the error vector due to inaccurate channel estimation or CSI quantization. Hence, additional inter-user interference is brought to the receiver k which is:

$$\mathbf{h}_k^H \sum_{j \neq k} \mathbf{e}_j \mathbf{s}_j$$

and the additional interference reduces the SINR of the receivers which in turn reduce the capacity of a multi-user MIMO system.

In general, there exist three types of linear precoding approaches, MMSE precoding, maximum ratio transmission (MRT) precoding, and zero-forcing (ZF) precoding. This work mainly uses zero-forcing precoding that aims at nulling the inter-user interference. ZF precoding can be close to optimal in terms of capacity when the user number is large. ZF precoding is actually using the concept of generalized inverse of the channel matrix \mathbf{H} and the pseudo-inverse of the channel matrix can satisfy that [62]:

$$\mathbf{H}^\dagger = \mathbf{H}(\mathbf{H}^H \mathbf{H})^{-1}. \quad (2.22)$$

Using \mathbf{H}^\dagger as the precoding matrix, the received signal can be modified as:

$$y_k = \mathbf{h}_k^H \sum_{j=1}^K \mathbf{h}_j^\dagger \mathbf{s}_j + n_k, k = 1, 2, \dots, K. \quad (2.23)$$

Change (2.23) to matrix representation, it will be:

$$\begin{aligned} \mathbf{y} &= \mathbf{H}\mathbf{H}^\dagger \mathbf{s} + \mathbf{n} \\ &= \mathbf{H}(\mathbf{H}^H (\mathbf{H}\mathbf{H}^H)^{-1}) \mathbf{s} + \mathbf{n}. \\ &= \mathbf{s} + \mathbf{n} \end{aligned} \quad (2.24)$$

It shows that this approach can achieve a linear capacity improvement of $\min(N_t, K)$ ideally. The drawbacks of this precoding method are that the channel may be ill-conditioned, and the matrix inversion will result in low SNR at some receivers [63]. There exist some other precoding approaches that can optimize the performance of ZF precoding, such as MMSE precoding or regularized ZF precoding. In this work, the original ZF precoding is used for the sake of simplicity.

2.4.3 Hardware impairments

According to [20], the information-theoretic capacity limits the spectral efficiency of a wireless channel and the capacity relies on both the signal-to-noise ratio (SNR) and on propagation environment, signal processing resources, transceiver hardware impairments and channel estimation accuracy etc. For massive MIMO systems, since each antenna channel uses less expensive hardware to reduce the system cost, hardware impairments become a very important issue that needs our attention. And impairments such as amplifier nonlinearities, I/Q imbalance, CFO, TO, phase noise and quantization errors exist in any transceiver. This section will introduce the models of these common impairments and how they are generated.

A. Amplifier nonlinearity

An amplifier, which amplifies the power of RF signal to reach a distant region, is an important module of a wireless communication system. An ideal amplifier operates in the linear region to amplify the input signal with different instant power by the same level. However, a practical amplifier behaves differently that it may work in the nonlinear region, especially when the power level is high, which will result in distortions such as in-band and out-of-band distortion and these distortions can cause higher bit error rate (BER) and reduce the system capacity [21]. What's more, the out-of-band distortion can cause adjacent channel interference.

A general modulated RF signal can be represented by [22]:

$$\begin{aligned} s(t) &= x_i(t)\cos(2\pi f_c t) - x_q(t)\sin(2\pi f_c t) \\ &= x(t)\cos(2\pi f_c t + \theta(t)) \end{aligned} \quad , \quad (2.25)$$

where $x(t) = (x_i^2(t) + x_q^2(t))^{1/2}$ is the envelope of the RF signal; $\theta(t) = \tan^{-1}(x_q(t)/x_i(t))$ represents the phase of the RF signal; and f_c denotes its

carrier frequency. $x_i(t)$ and $x_q(t)$ are its in-phase element (I) and quadrature element (Q) respectively.

After the process of the amplifier, the output signal is given as:

$$s_o(t) = G(x(t)) \cos(2\pi f_c t + \theta(t) + \Phi(x(t))), \quad (2.26)$$

where $G(x(t))$ (AM/AM) is the gain function and $\Phi(x(t))$ (AM/PM) the phase function. Consider the memoryless nonlinear model, and note that AM/PM effects are usually negligible ($\Phi(x(t)) \approx 0$), the effect of AM/AM is more worthy of attention [23][24]. The distortion caused by the amplifier nonlinearity can be compensated by using pre-distortion, which can be realized after the amplifier modelling. The modelling of amplifier nonlinearity is beyond the scope of this section, although pre-distortion can reduce the amplifier nonlinearity, different channels of a MIMO system may also perform in a different way.

B. I/Q imbalance

A modulated RF signal is demodulated at the receiver side. There are two kinds of receivers, superheterodyne receiver, and super-regenerative receiver. Almost all modern receivers use the superheterodyne structure which performs better compared to the super-regenerative structure. Frequency mixing is utilized to convert the original received signal from carrier frequency to a fixed intermediate frequency (IF), which is easier to process [25][26].

Based on superheterodyne, a direct-conversion receiver, also named as homodyne, synchrodyne, or zero-IF receiver was designed, which is a special case of superheterodyne where IF is zero. The structure of its receiver is simplified since there is no need for the conversion to an intermediate frequency. This reduces the system circuit complexity and cost. The structure of a direct-conversion receiver and its corresponding transmitter is shown in Figure 2.15 below. The baseband digital

information is converted to analog signal via DAC and then passed to a lowpass filter at the transmitter side, which will be modulated independently by the mixers into a passband RF signal. This RF signal is passed to the power amplifier (PA) for amplification afterwards and transmitted via an antenna. For the receiving of this signal, a low noise amplifier (LNA) is firstly used to amplify the low power signal and then fed to the mixers for down-conversion. The outputs of the mixers are the baseband in-phase signal and quadrature signal. Ideally, the I channel and Q channel are with the same amplitude and a phase difference of exactly 90-degree phase.

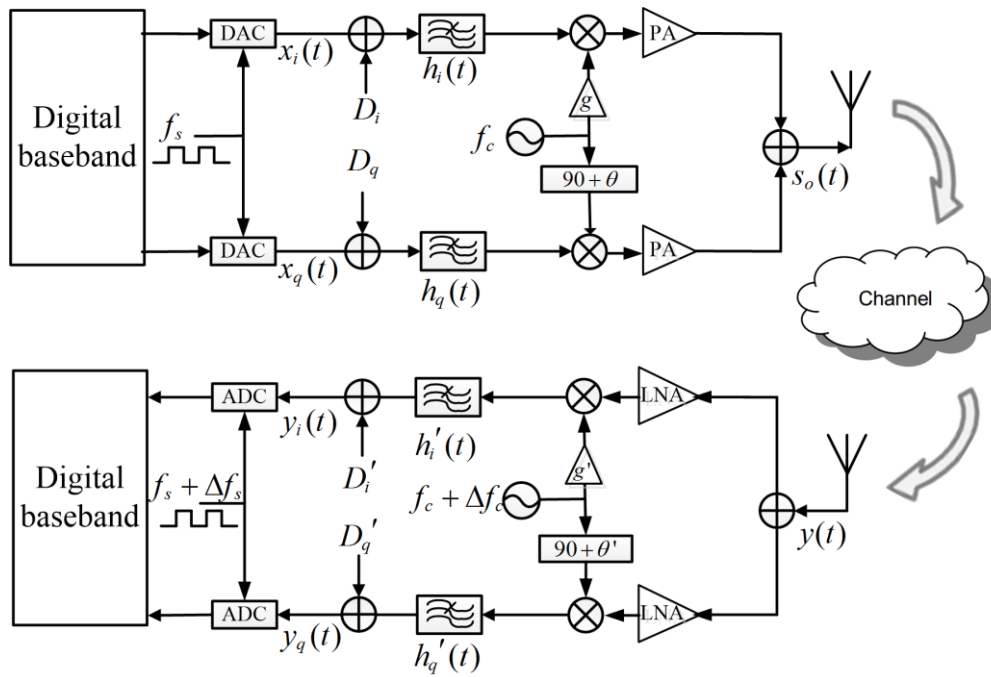


Figure 2.15 Direct conversion transceiver structure

For a practical system, the I and Q channel are not orthogonal because of the non-ideality of the hardware in terms of the mismatch of filter, mixer and amplifier in I and Q channel, and also the mismatch between the transmitter and receiver [27]. Hence, I/Q imbalance is introduced into the system that will result in image frequency interference and degrade the overall system performance.

C. CFO and TO

As shown in Figure 2.15 above, the transmitter and receiver are using different oscillators because of the physical separation, which means the frequency of the carriers generated by the oscillators are not the same. Moreover, for a mobile terminal, the Doppler effect exists between the transmitter and receiver. The two main reasons of carrier frequency offset are due to carrier frequency mismatch and Doppler effect. The received signal is shifted in frequency if CFO is applied. For an OFDM system, the orthogonality among the sub-carriers is broken. Frequency sources' separation means that the transmitter and receiver are working in the different digital clock, hence the sampling clock ADC and DAC are not synchronized, which will cause timing offset.

D. Phase noise

A transceiver system, or a digital system, needs a clock source to harmonize actions of the circuits. The clock signal, often generated from an oscillator, works like a metronome. Ideally, the oscillator would produce a pure and stable sine wave, that has a single frequency. Thus, all practical oscillators have phase noise components that expand the signal power to adjacent frequencies. Phase noise can be modelled by a stochastic process as:

$$v(t) = A \cos(2\pi f_0 t + \varphi(t)), \quad (2.27)$$

where A and f_0 are the amplitude and ideal frequency of the output sine wave respectively; $\varphi(t)$ denotes the stochastic process of the phase noise, which is typically expressed in dBc/Hz that means the noise power level relative to f_0 contained in a 1Hz bandwidth centred at a specific offset from f_0 . Phase noise can be measured via a spectrum analyser that has a less noisy local oscillator.

For an OFDM wireless system, phase noise will cause common phase errors and inter-carrier interference, which will result in a rotational constellation of received

signal, therefore needs to be compensated. Compensation can be made after estimation and many references have proposed the compensation methods of phase noise [28]-[31].

Since there are many compensation or calibration methods existing for hardware impairments mentioned above, this dissertation won't pay attention to the research of some new calibration algorithm. For 5G network or beyond, although compensation can be applied to the system, there still exist some residual impairments that need our attention. A massive MIMO system contains many hardware chains that may be running with cheap components, which will cause non-reciprocity between uplink and downlink. The cell members in a HetNet system are working with different oscillators hence it may generate a more complex CFO and TO situation.

2.5 Channel coding/decoding method

2.5.1 Background

Claude Shannon published an article named "A Mathematical Theory of Communication" in 1948, which became the foundation of information theory [84]. This article studied the problem of how to encode the information to the best. Information theory describes the data using a statistical description, which is also known as the information entropy. Coding theory is one of the most important applications of information theory. It includes source coding theory and channel coding theory. And channel coding theory includes the techniques like error detection and correction and error control. These techniques can help the system achieve reliable information transmission in noisy communication channels.

Channel noise will introduce errors into the data during the data delivery because many channels can be affected by it. Error detection can detect those errors and the errors can be reconstructed via error correction techniques in many cases. In general, error correction may be accomplished by two methods, forward error correction (FEC)

and automatic repeat request (ARQ). ARQ is also referred to as backward error correction. These two methods are often combined as a hybrid automatic repeat-request (HARQ) so that major errors can be recovered via retransmission while minor errors can be corrected without retransmission.

By adding some redundant data (parity data) to a message, a forward error correction (FEC) code can recover a number of errors of the data received by the receiver. FEC codes can be distinguished between convolutional codes and block codes. The former process the data bit-by-bit while the latter process the data block-by-block. Many types of codes are introduced in the past such as Viterbi code, Hamming code, Reed-Solomon code, Turbo code, low-density parity-check code, and so forth. Among them, Turbo code and LDPC code are believed to be able to provide channel capacity approaching performance. In 2009, another linear block error correcting code called Polar code was invented by Erdal Arıkan [85]. Polar code is the first channel capacity achieving code and is promoted by Huawei to be a part of the coding schemes of control channels of the 3GPP 5G standard.

Turbo code, which was developed in the 1990s, is the first practical code that can closely approach the channel capacity. Turbo coding uses two or more simple convolutional codes encoder and an interleaver to generate a block code. It gives similar performance with LDPC code and was applied to many applications such as CDMA2000, UMTS, LTE, and so on.

Turbo code can be decoded by iterative soft-decision decoding algorithms such as the BCJR algorithm [86], an algorithm that aims for maximum a posteriori decoding of FEC codes. The serial nature of BCJR algorithm makes the decoding process time-consuming thus the decoding latency is normally very large, which is a bottleneck for modern communication systems. The latency problem also holds back the improvement of data throughput. Therefore, there is a demand for reducing the

decoding latency. Parallelization of the BCJR decoding algorithm is a practical and effective way to improve the throughput and reduce the system latency at the decoding stage.

As a high performance forward error correction code, turbo code [87][88] is believed to be one of the most robust channel coding methods for wireless communications. In particular, turbo code is able to facilitate near-capacity transmission throughputs, leading to a wide deployment of the state-of-the-art communication standards such as WiMAX [89] and LTE[90] and could be employed in future 5G networks and beyond standard [91].

The detail of turbo codes encoding/decoding will be given in Chapter 5. Furthermore, parallel decoding methods will be introduced as well as the implementation of a low latency parallel turbo decoder.

2.5.2 Turbo encoding

There are many ways of encoding for Turbo code, which can be achieved by using different component encoders, interleavers, puncturing patterns, and input/output ratios. Turbo encoding can be accomplished by concatenating two convolutional encoders together. This section presents the Turbo encoder in the LTE standard and demonstrates the design of parallel turbo codes.

2.5.2.1 RSC encoding

The encoder of convolutional codes includes two types, non-recursive encoder and recursive encoder. General speaking, non-recursive codes are typically non-systematic because the output does not include the input data. While recursive codes are systematic on the contrary, hence it is also named recursive systematic convolutional (RSC) codes. An RSC encoder is a state machine with finite states. For an encoder with n binary cells, it will have 2^n states.

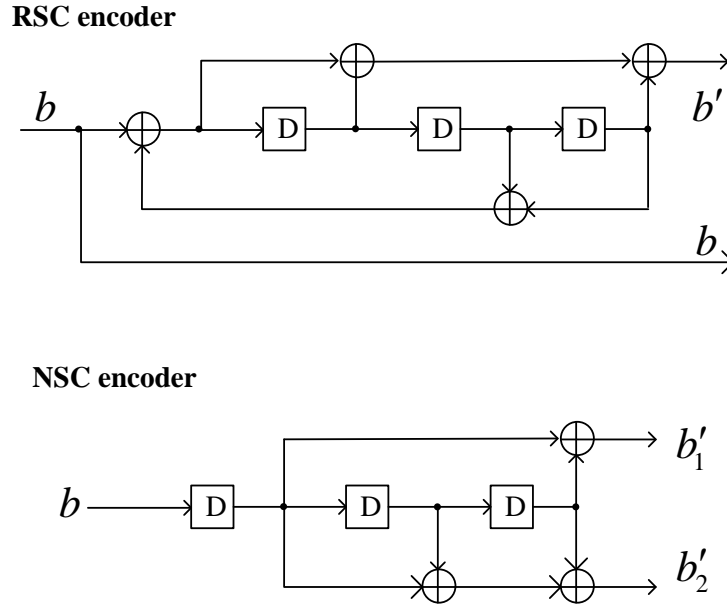


Figure 2.16 Convolutional encoding

Figure 2.16 shows a single RSC encoder with 8-state along with a non-recursive convolutional (NSC) for comparison. Each register (Delay) start with a value of 0 and hold 1 input bit. The input data bit b is firstly passed to the left register. Then it is shifted to the right and another input is going to happen until there are no more input bits and all registers return to 0. For the RSC encoder, the outputs of the second and third register are going to feed back to the first input, hence RSC is recursive.

The LTE Turbo encoder is made up of two tail-biting recursive systematic convolutional (RSC) encoders in parallel, as shown in Figure 2.17(a). The second RSC encoder is placed after an interleaver (Π). These two encoders generate two N - bit encoded frames, named a parity frame and a systematic frame. Each RSC coding rate is $R=1/2$ with a codeword length of N and a constraint length of $l=4$. The encoder can also be represented by a trellis diagram as shown in Figure 2.17(b) below. Since the message frame uses three encoded frames, the systematic frame (b_i), the two parity frames ($p_{1,i}$ and $p_{2,i}$), the turbo encoder produces a total length of $3N$ bits frame x_i and the overall coding rate is $R=1/3$. Following the Turbo encoder, the encoded frames are modulated and then transmitted to the receiver via a noisy channel.

The LTE RSC encoder operates on the basis of an $M=8$ -state transition diagram which is shown in Figure 2.17(b). The encoder begins from an initial state of $S_0=0$ and transits into each subsequent state $S_i \in \{0,1,2,\dots,M-1\}$ according to the corresponding message bit $b_i \in \{0,1\}$. Since the message bit $b_i \in \{0,1\}$ has two possibilities, there will be 2 potential transitions from the previous state S_{i-1} to the current state S_i .

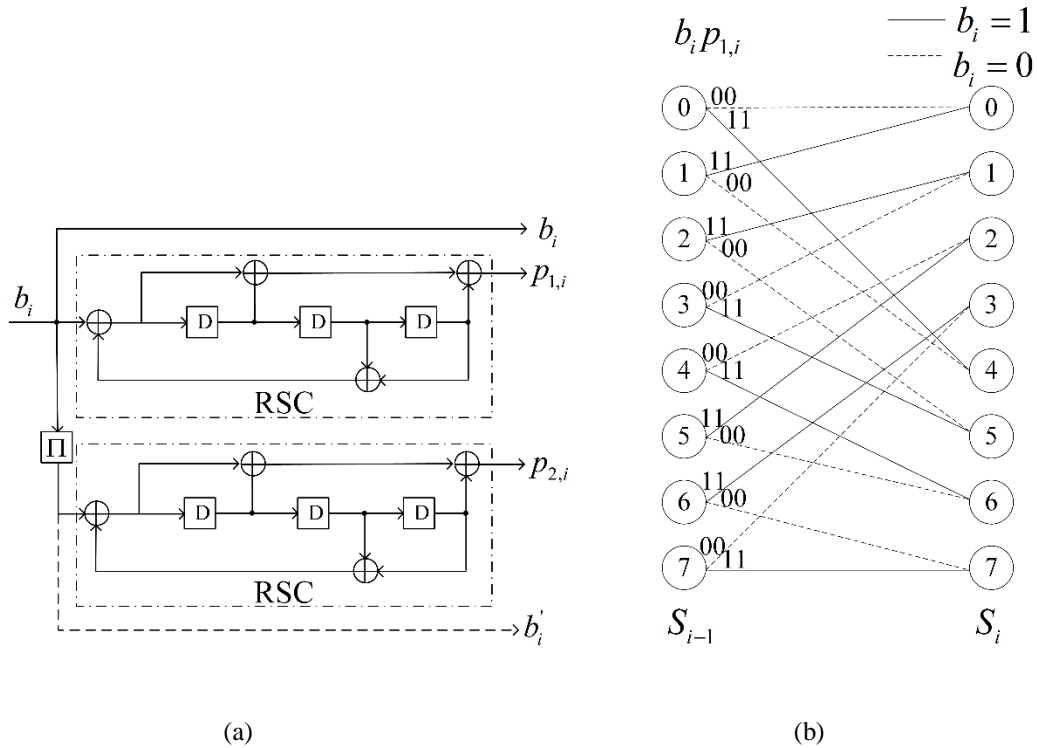


Figure 2.17 Turbo encoder: (a) Encoder structure; (b) Trellis diagram.

The transfer function can be used to describe the 8-state constituent code [92]:

$$T(D) = \begin{bmatrix} 1, \frac{g_1(D)}{g_0(D)} \end{bmatrix}, \quad (2.28)$$

where $g_0(D) = 1 + D^2 + D^3$ and $g_1(D) = 1 + D + D^3$; D denotes the value of shift register or delay. The initial value of the 8-state encoders are zeros. And the output of the Turbo encoder is:

$$b_0, p_{1,0}, p_{2,0}, b_1, p_{1,1}, p_{2,1}, \dots, b_{N-1}, p_{1,N-1}, p_{2,N-1}$$

where bits b_0, b_1, \dots, b_{N-1} are the input of the Turbo encoder, which are also the input of both the first RSC encoder and the internal interleaver; bits $p_{1,0}, p_{1,1}, \dots, p_{1,N-1}$ and $p_{2,0}, p_{2,1}, \dots, p_{2,N-1}$ are the output of the first and second RSC encoder; bits $b'_0, b'_1, \dots, b'_{N-1}$ are the input bits after interleaving which are used at the tail of the Turbo encoder output. At the end of the trellis, some tail bits are joint at the end of the transmitted bits, which are:

$$b_N, p_{1,N}, b_{N+1}, p_{1,N+1}, b_{N+2}, p_{1,N+2}, b'_N, p_{2,N}, b'_{N+1}, p_{2,N+1}, b'_{N+2}, p_{2,N+2}.$$

As shown in Figure 2.22(a), the input bits should pass an interleaver Π to generate the second parity frame.

2.5.2.2 Interleaving

To improve the performance of FEC codes, an interleaver is frequently used. An interleaver performs interleaving by shuffling the source bits to separate the distribution of errors, which helps the recovery of the errors.

The index relationship of the data between the input and output of this interleaver can be represented as:

$$i' = \pi(i), i = 0, 1, \dots, N-1, \quad (2.29)$$

where i is the original sequence index and the interleaving process is shown in Figure 2.18.

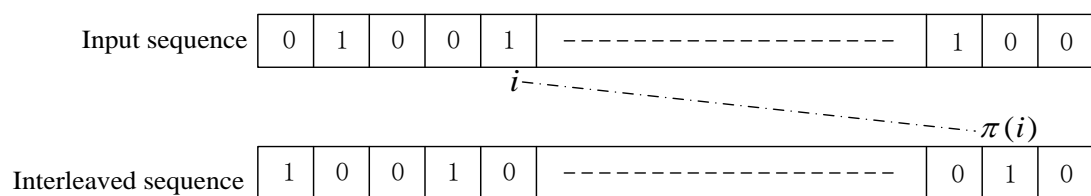


Figure 2.18 Interleaving

And the output sequence after the interleaver will be:

$$b'_i = b_{\pi(i)}, i = 0, 1, \dots, N-1, \quad (2.30)$$

where the shuffling is achieved by $\pi(i)$, which changes the index of the input bits.

According to [105], there are two criteria when designing an interleaver. The first one is the distance spectrum of the code and the second is the correlation of the parity bits and the information data sequence. The noise floor can be reduced if the minimum effective free distance is increased [106] or choose some appropriate interleavers when the block size is fixed. There are many types of interleavers, random interleavers, S-random interleavers, Berrou-Glavieux interleavers, UMTS interleavers, and so forth.

A. Random interleaver

A random interleaver can be utilized in Turbo code to perform a random permutation. And a random interleaver is believed to be a factor for the Turbo codes to reduce the error-floor at low BER region [107] by increasing the distance spectrum of the code. However, it is not practical for implementation because the interleaving pattern needs to be stored for decoding. Hence it is important to have some deterministic interleavers with random-like properties, which can be generated by some algorithms.

B. S-random interleaver

S-random interleaver, or semi-random interleaver, is a type of interleaver that requires the distance between the output index $\pi(i)$ and the input index i to be larger than a parameter s . The output position $\pi(i)$ is chosen randomly but if the index difference $|i - \pi(i)|$ is smaller than s , the output is invalid and the interleaving index is going to regenerate. The index generation process is repeated until all the N output indices meet the requirement. Hence the interleaver is not just random in structure but also guarantees minimum spreading errors [111].

Obviously, the drawback of the S-random interleaver is the time consumption of the index generation because output index may not always satisfy the distance condition. In [111], it was found that the index generation time is reasonable if the parameter s is chosen to be less than $\sqrt{N/2}$. The BER performance of the S-random interleaver is very well due to the strict distance condition. But it is time-consuming especially when the frame size is big hence not practical in many applications.

C. Block interleaver

A block interleaver is believed to be the simplest non-trivial interleaver. The interleaving process is accomplished by reading the data into an $R \times C$ rectangular matrix row by row and reading out column by column. This interleaver can be expressed as:

$$\pi(i) = (R \cdot i) \bmod (N-1), 0 \leq i \leq N-1, \quad (2.31)$$

which is simple to construct. According to [112], the problem of the block interleaver is that it exhibits a very low degree of randomness.

Block interleavers and linear interleavers are known as two simplest deterministic interleavers. For short frame sizes if proper parameters are chosen, they can perform even better than random interleavers [108]. A deterministic interleaver based on permutation polynomials over the ring of integers is introduced in [108], the performance of which can be close to S-random interleavers. Permutation polynomial-based interleaver can be realized simply by an arithmetic computation which is essential from the practical perspective.

In LTE standard, the interleaver is chosen to be permutation polynomial-based and $\pi(i)$ is given by:

$$\pi(i) = (f_1 i + f_2 i^2) \bmod N, \quad (2.32)$$

where the value of f_1 and f_2 can be found in the 3GPP Technical Specification [93].

2.5.3 Turbo decoding

The decoding of turbo codes is built similarly as the encoding. Two RSC decoders are interconnected in a serial way with each other as shown in Figure 2.19.

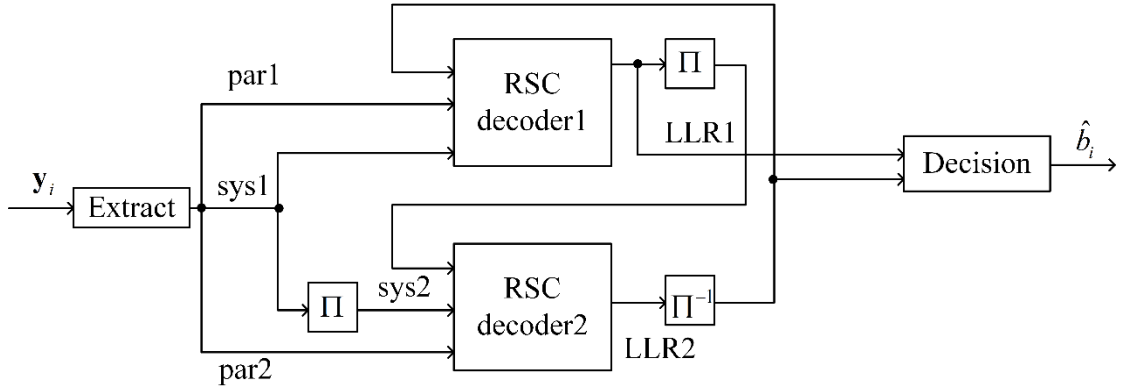


Figure 2.19 Turbo decoding structure

According to the encoding method, the received frame \mathbf{y}_i includes the one systematic frame and two parity frames. Hence \mathbf{y}_i is first extracted into three encoded frames, systematic frame (sys1), parity frame 1 (par1), and parity frame 2 (par2), based on the encoder. Two RSC decoders are included to perform iterative decoding. The sys1 and par1 will be transmitted into RSC decoder 1 while sys2, which is generated from sys1 by the interleaver, and par2 will input into RSC decoder 2. The whole decoding process works as follow: Firstly, RSC decoder 1 generates extrinsic information LLR1 according to systematic, parity and a priori bits. LLR1 is utilized as a priori information by RSC decoder 2 after interleaving. Secondly, the new extrinsic information LLR2 generated by RSC decoder 2 is fed back to decoder 1 after the process of deinterleaver (Π^{-1}). Therefore, the decoding iteration begins and after sufficient iteration, the performance of the decoder can approach to optimal.

2.5.3.1 MAP algorithm

A decoding algorithm, that is suitable for the RSC decoder and is based on a posteriori probabilities, was first published in 1974. This algorithm is known as Maximum a Posteriori (MAP), BCJR or forward-backward algorithm [94]. The MAP algorithm works in a recursive manner for the computation of the Log-Likelihood Ratio (LLR) of each bit according to the received data.

For the encoded sequence $\mathbf{x} = \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_N$, $\mathbf{x}_i = [x_{i1}, x_{i2}, x_{i3}]$ is the code word for each input bit b_i and x_{i1}, x_{i2}, x_{i3} are the sys1, par1 and par2 respectively. As the message bit b_i has two possible value 0 or 1, we can define the LLR as

$$L(b_i) = \ln \frac{P(b_i = 1)}{P(b_i = 0)}, \quad (2.33)$$

where $P(b_i = k), k \in \{0, 1\}$ is the posteriori probability of bit b_i which means the probability of deciding the received bit b_i as k . The received sequence $\mathbf{y} = \mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N$ is delivered to the decoder for the estimation of the original bit b_i . The decoding algorithm will compute a posteriori LLR given by

$$L(b_i | \mathbf{y}) = \ln \frac{P(b_i = 1 | \mathbf{y})}{P(b_i = 0 | \mathbf{y})}. \quad (2.34)$$

The $L(b_i | \mathbf{y})$, which represents the conditional probability, can be converted to a bit value through hard decision afterwards. More specifically, if $L(b_i | \mathbf{y}) < 0$, the estimation of the message bit will be $\hat{b}_i = 0$ and $\hat{b}_i = 1$ if $L(b_i | \mathbf{y}) > 0$. Therefore, the key problem of decoding is the calculation of LLR. After LLR calculation, the extrinsic information will be obtained. According to [95], the LLR can be defined by the joint probabilities of three parameters as:

$$\Lambda(b_i) = \ln \frac{\sum_{S_i} \sum_{S_{i-1}} \gamma_1(\mathbf{y}_i, S_{i-1}, S_i) \alpha_i(S_i) \beta_i(S_i)}{\sum_{S_i} \sum_{S_{i-1}} \gamma_0(\mathbf{y}_i, S_{i-1}, S_i) \alpha_{i-1}(S_{i-1}) \beta_{i-1}(S_{i-1})}, \quad (2.35)$$

where

$$\alpha_i(S_i) = P(S_i | \mathbf{y}_i) \quad (2.36)$$

$$\beta_i(S_i) = \frac{P(\mathbf{y}_{i+1} | S_i)}{P(\mathbf{y}_{i+1} | \mathbf{y}_i)} \quad (2.37)$$

and

$$\gamma_j(\mathbf{y}_i, S_{i-1}, S_i) = P(b_i = j, S_i, \mathbf{y}_i | S_{i-1}). \quad (2.38)$$

Based on $\gamma_j(\mathbf{y}_i, S_{i-1}, S_i)$, $\alpha_i(S_i)$ and $\beta_i(S_i)$ can be obtained by forward and backward recursions:

$$\alpha_i(S_i) = \frac{\sum_{S_{i-1}} \sum_{j=0}^1 \gamma_j(\mathbf{y}_i, S_{i-1}, S_i) \cdot \alpha_{i-1}(S_{i-1})}{\sum_{S_i} \sum_{S_{i-1}} \sum_{j=0}^1 \gamma_j(\mathbf{y}_i, S_{i-1}, S_i) \cdot \alpha_{i-1}(S_{i-1})} \quad (2.39)$$

and

$$\beta_i(S_i) = \frac{\sum_{S_{i-1}} \sum_{j=0}^1 \gamma_j(\mathbf{y}_i, S_{i-1}, S_i) \cdot \alpha_{i-1}(S_{i-1})}{\sum_{S_i} \sum_{S_{i-1}} \sum_{j=0}^1 \gamma_j(\mathbf{y}_i, S_{i-1}, S_i) \cdot \alpha_{i-1}(S_{i-1})}, \quad (2.40)$$

where S_i is the state of the encoder at the time i which can be the value between 0 and $2^M - 1$. $\gamma_j(\mathbf{y}_i, S_{i-1}, S_i)$ represents the encoder trellis and the transition probability of the wireless channel and it is given by

$$\gamma_j((y_i^s, y_i^p), S_{i-1}, S_i) = p(y_i^s | b_i = j) \cdot p(y_i^p | b_i = j, S_i, S_{i-1}) \cdot q(b_i = j | S_i, S_{i-1}) \cdot P_r\{S_i | S_{i-1}\}, \quad (2.41)$$

where y_i^s and y_i^p refer to the received systematic symbol and parity symbol, respectively. The term $q(b_i = j | S_{i-1}, S_i)$ is either one or zero which depends on where b_i is related to the transition from S_{i-1} to S_i .

The decoding process of the MAP algorithm includes the following steps:

- 1) Initialize $\alpha_0(S_0)$ and $\beta_N(S_N)$ as:

$$\alpha_0(S_0) = \begin{cases} 1 & \text{if } S_0 = 0 \\ 0 & \text{otherwise} \end{cases} \quad (2.42)$$

$$\beta_N(S_N) = \begin{cases} 1 & \text{if } S_N = 0 \\ 0 & \text{otherwise} \end{cases}, \quad (2.43)$$

- 2) When received bit b_i and the corresponding \mathbf{y}_i , $\gamma_j(\mathbf{y}_i, S_{i-1}, S_i)$ is computed along with $\alpha_i(S_i)$ in the forward recursion according to (2.39). The computed vectors $\gamma_j(\mathbf{y}_i, S_{i-1}, S_i)$ and $\alpha_i(S_i)$ are stored in the memory for $1 \leq i \leq N$.
- 3) After received all the N data sequence, $\beta_i(S_i)$ is calculated in the backward recursion based on (2.40) for $1 \leq i \leq N-1$.
- 4) Based on the results of the first three steps, the $\Lambda(b_i)$ can be computed according to (2.35) for $1 \leq i \leq N$. And then, the decoding output can be obtained by soft decision.

From the decoding trellis's perspective, the decoding process can be shown in Figure 2.20.

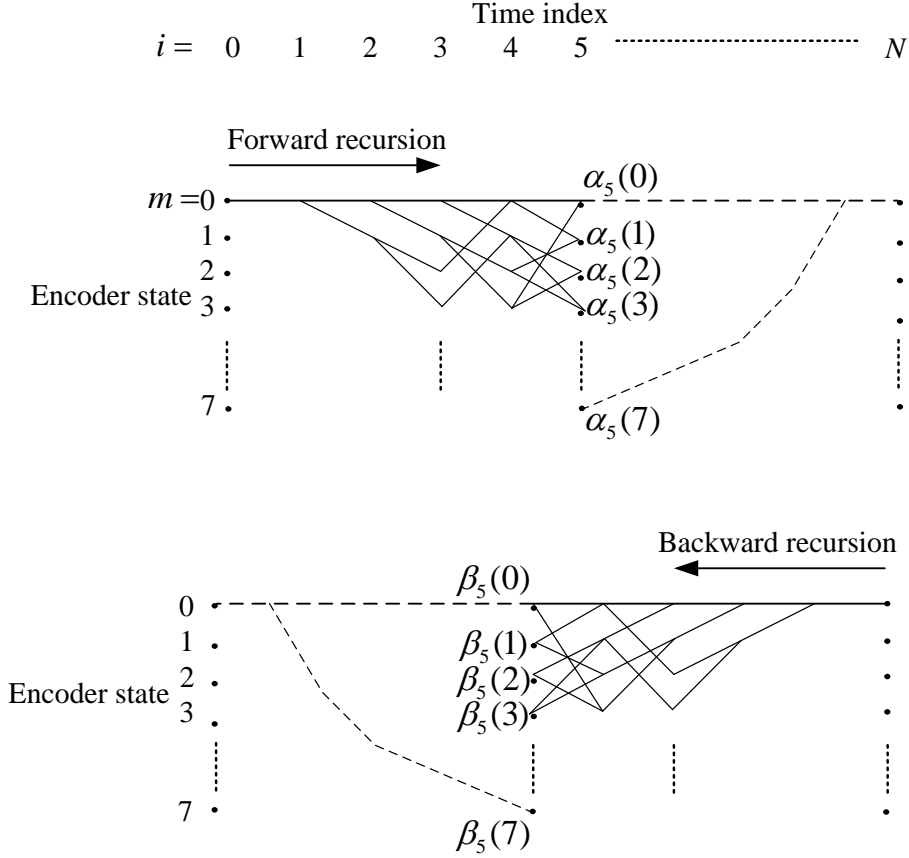


Figure 2.20 Decoding trellis

The MAP algorithm is very computationally complex for most applications and not suitable for chip design [95]. This is because MAP requires accurate estimation of the noise power and the performance of MAP is very sensitive to SNR estimation error. For the sake of practical implementation, based on MAP algorithms, log-MAP and max-log-MAP algorithms were introduced.

2.5.3.2 Max-log-MAP algorithm

The max-log-MAP algorithm is used in this Chapter 5 and it reduces the complexity of MAP algorithm but with similar performance. The Natural logarithm of the above variables is computed hence the multiplication becomes addition [109]. Note that

$$p(y_i^p | b_i = j, S_i, S_{i-1}) = \frac{1}{\sqrt{\pi N_0}} \cdot e^{-\frac{1}{N_0} (y_i^p - x_i^p(j, S_i, S_{i-1}))^2} \quad (2.44)$$

and

$$p(y_i^p | b_i = j) = \frac{1}{\sqrt{\pi N_0}} \cdot e^{-\frac{1}{N_0}(y_i^s - x_i^s(j))^2}. \quad (2.45)$$

Based on (2.44) and (2.45), taking the logarithm of $\gamma_j((y_i^s, y_i^p), S_{i-1}, S_i)$ can be derived as:

$$\ln(\gamma_j((y_i^s, y_i^p), S_{i-1}, S_i)) = \frac{2y_i^s x_i^s(i)}{N_0} + \frac{2y_i^p x_i^p(i, S_i, S_{i-1})}{N_0} + \ln(P_r \{S_i | S_{i-1}\}) + C. \quad (2.46)$$

The constant value C can be ignored and it cancels out the calculation of $\ln(\alpha_i(S_i))$ and $\ln(\beta_i(S_i))$. And then, to get $L(b_i)$, we can calculate the logarithm of α as:

$$\begin{aligned} \ln(\alpha_i(S_i)) = & \ln\left(\sum_{S_{i-1}} \sum_{j=0}^1 e^{\ln(\gamma_j((y_i^s, y_i^p), S_{i-1}, S_i)) + \ln(\alpha_{i-1}(S_{i-1}))}\right) \\ & - \ln\left(\sum_{S_i} \sum_{S_{i-1}} \sum_{j=0}^1 e^{\ln(\gamma_j((y_i^s, y_i^p), S_{i-1}, S_i)) + \ln(\alpha_{i-1}(S_{i-1}))}\right) \end{aligned} \quad (2.47)$$

For the sake of simplicity, the following approximation can be used:

$$\ln(e^{x_1} + e^{x_2} + e^{x_3} + \dots) \approx \max(x_1 + x_2 + x_3 + \dots), \quad (2.48)$$

Hence the following equations can be obtained. Let

$$\tilde{\gamma}_j((y_i^s, y_i^p), S_{i-1}, S_i) = \gamma_j((y_i^s, y_i^p), S_{i-1}, S_i), \quad (2.49)$$

then

$$\begin{aligned} \tilde{\alpha}_i(S_i) = \ln(\alpha_i(S_i)) \approx & \max_{(S_{i-1}, j)} (\tilde{\gamma}_j((y_i^s, y_i^p), S_{i-1}, S_i) + \tilde{\alpha}_{i-1}(S_{i-1})) \\ & - \max_{(S_i, S_{i-1}, j)} (\tilde{\gamma}_j((y_i^s, y_i^p), S_{i-1}, S_i) + \tilde{\alpha}_{i-1}(S_{i-1})) \end{aligned} \quad (2.50)$$

and similarly,

$$\begin{aligned}\tilde{\beta}_i(S_i) = \ln(\beta_i(S_i)) \approx & \max_{(S_{i-1}, j)} (\tilde{\gamma}_j((y_{i+1}^s, y_{i+1}^p), S_i, S_{i+1}) + \tilde{\beta}_{i+1}(S_{i+1})) \\ & - \max_{(S_i, S_{i+1}, j)} (\tilde{\gamma}_j((y_{i+1}^s, y_{i+1}^p), S_i, S_{i+1}) + \tilde{\alpha}_i(S_i))\end{aligned}\quad (2.51)$$

And then the approximation of the log-likelihood reliability of the received data bit b_i can be:

$$\begin{aligned}\Lambda(b_i) \approx & \max_{(S_i, S_{i-1})} (\tilde{\gamma}_1((y_i^s, y_i^p), S_{i-1}, S_i) + \tilde{\alpha}_{i-1}(S_{i-1}) + \tilde{\beta}_i(S_i)) \\ & - \max_{(S_i, S_{i-1})} (\tilde{\gamma}_0((y_i^s, y_i^p), S_{i-1}, S_i) + \tilde{\alpha}_{i-1}(S_{i-1}) + \tilde{\beta}_i(S_i))\end{aligned}\quad (2.52)$$

To apply this algorithm to a Turbo decoder, $\Lambda(b_i)$ should be separated into three terms, extrinsic, a priori and systematic components. By defining

$$\tilde{\gamma}'_j(y_i^p, S_{i-1}, S_i) = \ln(p(y_i^p | b_i = j, S_i, S_{i-1})) + \ln(q(b_i = j | S_i, S_{i-1})), \quad (2.53)$$

we can obtain

$$\begin{aligned}\Lambda(b_i) \approx & \max_{(S_i, S_{i-1})} (\tilde{\gamma}'_1(y_i^p, S_{i-1}, S_i) + \tilde{\alpha}_{i-1}(S_{i-1}) + \tilde{\beta}_i(S_i)) + \ln(p(y_i^s | b_i = 1)) + \ln(P_r \{b_i = 1\}) \\ & - \max_{(S_i, S_{i-1})} (\tilde{\gamma}'_0(y_i^p, S_{i-1}, S_i) + \tilde{\alpha}_{i-1}(S_{i-1}) + \tilde{\beta}_i(S_i)) + \ln(p(y_i^s | b_i = 0)) + \ln(P_r \{b_i = 0\})\end{aligned}\quad (2.54)$$

Equation (2.54) can be rewritten as:

$$\begin{aligned}\Lambda(b_i) \approx & \max_{(S_i, S_{i-1})} (\tilde{\gamma}'_1(y_i^p, S_{i-1}, S_i) + \tilde{\alpha}_{i-1}(S_{i-1}) + \tilde{\beta}_i(S_i)) \\ & - \max_{(S_i, S_{i-1})} (\tilde{\gamma}'_0(y_i^p, S_{i-1}, S_i) + \tilde{\alpha}_{i-1}(S_{i-1}) + \tilde{\beta}_i(S_i)) - \frac{4y_i^s}{N_0} - L(b_i)\end{aligned}\quad (2.55)$$

The first two components of (2.55) are the extrinsic component, the third component denotes the systematic information and the last one represents the a priori component [109].

In the next decoding stage, the extrinsic component will be used as a priori information. This a priori LLR for bit b_i is $L(b_i)$ and it can be determined as;

$$L(b_i) = \ln \left(\frac{P_r \{b_i = 1\}}{P_r \{b_i = 0\}} \right) = \ln \left(\frac{P_r \{S_i | S_{i-1}\}}{1 - P_r \{S_i | S_{i-1}\}} \right), \quad (2.56)$$

therefore $\ln(P_r \{S_i | S_{i-1}\}) = L(b_i) - \ln(1 + e^{L(b_i)})$. Using similar approximation way as (2.48),

$$\ln(P_r \{S_i | S_{i-1}\}) \approx L(b_i) - \max(0, L(b_i)). \quad (2.57)$$

Based on the Max-log-MAP algorithm described above, the initialization in step 1) above becomes:

$$\alpha_0(s) = \begin{cases} 0 & \text{if } s = 0 \\ -\infty & \text{otherwise} \end{cases}$$

and $\beta_i(s)$ can be represented by similar approximation and the soft decision can be made according to:

$$\begin{aligned} \tilde{L}(b_i) \approx & \max_{S_{i-1}} (\tilde{\gamma}_1(y_i^p, S_{i-1}, S_i) + \alpha_{i-1}(S_{i-1}) + \beta_i(S_i)) \\ & - \max_{S_{i-1}} (\tilde{\gamma}_0(y_i^p, S_{i-1}, S_i) + \alpha_{i-1}(S_{i-1}) + \beta_i(S_i)). \end{aligned} \quad (2.58)$$

After the calculation of $\tilde{L}(b_i)$, iteration is performed as shown in Figure 2.19.

2.5.3.3 Log-MAP algorithm

Compared to the MAP algorithm, the Max-Log-MAP algorithm is suboptimal because of the approximation of (2.56) which reduces the soft-decision accuracy. The problem occurs when calculate $\ln(e^{x_1} + e^{x_2} + e^{x_3} + \dots)$ which can be solved by the Jacobian logarithm [110]:

$$\begin{aligned} \ln(e^{x_1} + e^{x_2}) &= \max(x_1, x_2) + \ln(1 + e^{-|x_2 - x_1|}) \\ &= \max(x_1, x_2) + f_c(|x_2 - x_1|), \end{aligned} \quad (2.59)$$

where $f_c(\cdot)$ represents a correction function. The calculation of $\ln(e^{x_1} + e^{x_2} + e^{x_3} + \dots)$ will be accomplished by utilizing the Jacobian logarithm recursively.

$$\ln(e^{x_1} + e^{x_2} + e^{x_3} + \dots) = \ln(\Delta + e^{x_N}), \quad (2.60)$$

where $\Delta = e^{x_1} + e^{x_2} + \dots + e^{x_{N-1}} = e^x$ and

$$\begin{aligned} \ln(e^{x_1} + e^{x_2} + e^{x_3} + \dots) &= \max(\ln \Delta, x_N) + f_c(|\ln \Delta - x_N|) \\ &= \max(x, x_N) + f_c(|x - x_N|). \end{aligned} \quad (2.61)$$

By correcting the approximation each step recursively, the Log-MAP algorithm can approach the original MAP algorithm. However, calculating the correcting function is complex which can be solved by a pre-computed table.

2.5.3.4 Decoding latency and throughput

As discussed above, to decode one message bit, the decoder should take all the received signal into account. Put differently, the decoder needs to calculate in both forward and backward directions, which makes the decoding process to be sequential. What's more, the iteration takes multiple times of the decoding time of a single RSC decoder. The theoretical decoding latency can be given as:

$$D = 4N \cdot I, \quad (2.62)$$

where I is the iteration number and the decoding throughput will be:

$$T = \frac{Nf}{D}, \quad (2.63)$$

where f represents the decoding clock frequency. Obviously, the serial nature of the decoding algorithm will limit the throughput, which is the bottleneck of decoding. Therefore, the decoding time of the RSC decoder should be reduced to improve both the system throughput and latency performance. Parallel decoding is believed to be effective to reduce the decoding latency [96].

2.6 VLC systems introduction

As the demand for higher data rate and interconnected devices number keep growing, the radio frequency spectrum is very limited and becoming scarce, which led to the research and development of optical wireless communication (OWC) technology. The visual light spectrum has particularly got people's interest due to its advantages of large bandwidth and the ability to support both illumination and communication. Visible light communication (VLC), which is a subset of optical wireless communication, uses visible light between 430 and 790 THz (750-380 nm) for data communication [117] as shown in Figure 2.21. VLC can play a role as a communications medium for ubiquitous computing due to the extensive existence of the light-producing devices. Light-producing devices can be found in many applications such as traffic signs, indoor/outdoor lamps, car headlights and so forth. Moreover, visible light is a green medium that is less dangerous for high-power applications compared to radio frequency.

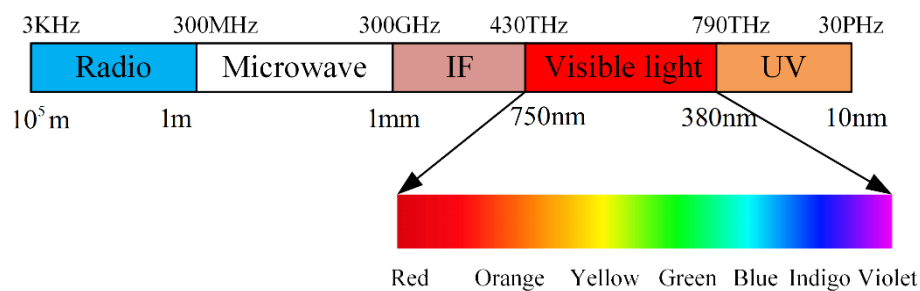


Figure 2.21 VLC spectrum

The development history of VLC can date back to 1880s when the photophone, which can transmit speech on modulated sunlight over hundreds of metres, was invented by Alexander Graham Bell. Since 2003, many research institutes, such as Oxford University, began to research on VLC. In 2010, a demo of VLC over a light emitting diode (LED) was shown by Siemens and its co-operators. The transmission data rate can reach up to 500 Mbit/s over a distance of 5 metres [118]. In July 2011, a live demonstration of high-definition video transmission over a LED lamp was presented

at TED Global [119]. Since then, a new term “Li-Fi” becomes popular in the industry. Li-Fi tends to support the Internet of Things (IoT) and the data transmission speed can be up to 10Gbits/s [120].

Compared to RF communication, VLC has a lot of advantages, such as its unlicensed wide spectrum, security to interference in indoor scenario, high data transmission rate since high SNR can be obtained indoor, and so forth. Because of these advantages, VLC systems are suitable for various applications like smart home/office network, hybrid RF-VLC systems, intelligent transport systems, underwater communications, and so on. VLC communication through LEDs can bring illumination and communication simultaneously for a home/office environment. This indoor application can release the RF spectrum band in 2.4GHz and 5GHz, which can be utilized in other scenarios. As discussed Section 2.3, the future communication networks are likely to be heterogeneous with multiple technologies together to enhance users’ experience. A hybrid RF-VLC system can join together which can be highly energy efficient and reliable. As auto-driving technologies become more and more popular nowadays, intelligent transport systems can support the communication between vehicle and vehicle through VLC. For underwater communication, VLC is believed to be more appropriate than RF because of its longer transmitting distance than radio waves [121]. More other applications can be achieved by VLC, this thesis focuses on the indoor scenario in terms of the smart home/office network.

2.6.1 Architecture of VLC

The architecture of VLC systems is similar to RF systems, which consists a transmitter and a receiver. And three common layers, physical layer, MAC layer, and application layer, are defined for VLC which is the same as RF systems. The physical layer of VLC systems is slightly different with RF systems because of the different transmit/receive front-end and modulation/demodulation scheme.

2.6.2 Physical layer of VLC

The typical physical layer block diagram is shown in Figure 2.22. First, the data bit stream is encoded by channel encoder to enhance the performance. Second, the encoded bit stream will be fed into the modulator which performs modulation like OOK, PPM, PWM, OFDM, etc. Finally, the modulated data will pass a LED or other light source and then the visible light is emitted over the optical channel. Depends on the actual implementation structure of the VLC system, the detail of the physical layer can be different.

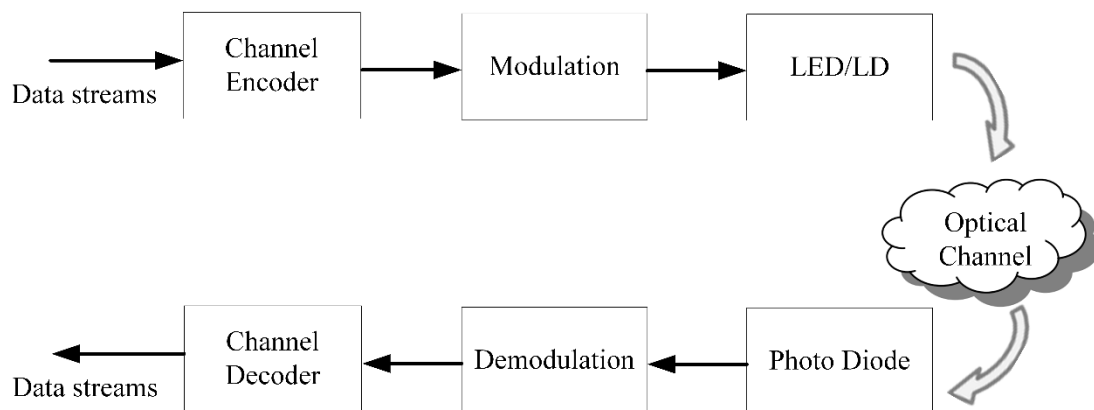


Figure 2.22 Physical layer of VLC systems

From the hardware architecture point of view, the main difference between VLC systems and RF systems is the front-end. Literally, the transmitter front-end of a VLC system can be any device that can generate visible light, such as a bulb, LED, a laser diode (LD), and so forth. LED is a most popular kind due to its reliability, power efficient, luminous efficiency. Despite most LEDs can be found in our daily life are white LEDs, some coloured LEDs (red, green and blue) can be used in VLC systems as well to modulate the data using colour shift keying (CSK). The LED selection is based on the channel model and the modulation scheme. The LED converts the electrical signal into an optical signal.

A photodiode is normally used at the receiver side. It converts the visible light into photo current. Different types of photodiodes, such as PIN diode, silicon photodiode, and avalanche photodiode, can be introduced to VLC [132]. Among them, the gain of an avalanche photodiode is higher but more expensive. An optical filter should be used to mitigate the interference from other light sources such as the sun light.

IEEE 802.15.7 defines three types of physical layers with different data rates. The PHY 1 can only achieve a data rate from 11.67 Kbit/s to 267.6 Kbit/s and it is mainly for outdoor application. The PHY 2 can support a minimum data rate of 1.25 Mbit/s and a maximum rate of 96 Mbit/s. The PHY 3 uses CSK modulation method and it supports many kinds of emission sources with data rate from 12 Mbit/s to 96 Mbit/s [129].

2.6.3 Modulation techniques

Since the frequency of visible light is much higher than RF band, it is not practical to modulate the base band data to the visible light band using a frequency mixer. Hence a key difference between VLC and RF in the physical layer is the modulation part. The most likely front-end devices in VLC systems are the widely-used LEDs and photodiodes (PDs). At the receiver side, the intensity of the incoherent light from the LEDs can be detected by the PD. However, the modulation or detection of the phase and amplitude of the light wave by the PDs are not accurate. Hence, only intensity modulation and direct detection (IM/DD) is practical for VLC systems, which implies that the data signal is encoded into the intensity variation of light. Since the intensity of light cannot be negative, the digital signals should be real-valued and positive [123].

Real-valued modulation schemes should be used for VLC systems, such as OOK, PWM, PPM, CSK, OFDM, and so on. They utilize different principles and have different performances in terms of data rate.

A. On-Off keying (OOK)

OOK modulation is one of the simplest modulation for VLC, which modulates the data bits 1 or 0 by turning the LEDs on or off. When OOK is off, the intensity of the light is reduced to a level that the PD can distinguish clearly between “on” and “off”. It does not necessarily mean the LED is turned off completely. To prevent the LED generating flicker which is harmful to humans, Manchester RLL code is used in OOK modulation, see Figure 2.23[122]. The Manchester encoding, which provides a DC balanced code, generates an OOK symbol “01” to represent a logic zero and “10” as a logic one.

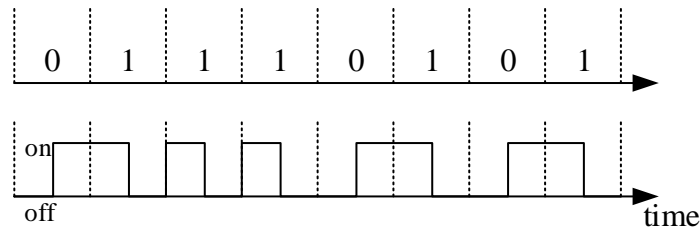


Figure 2.23 DC balanced OOK

The main advantage of OOK is that it is easy to implement. However, due to the slow response of the LEDs, the data rate of OOK is low. According to [124] below, the data rate of OOK can only reach up to 125 Mbps.

B. Pulse modulation techniques

The limitation of OOK motivates the research to find new modulation methods with higher data rates. Pulse Width Modulation (PWM) and Pulse Position Modulation (PPM) are two of the pulse modulation techniques. In PPM, the data pulse is modulated to different optical pulse positions according to the input data bits. PPM is more power efficient than OOK but requires more bandwidth. Variable PPM (VPPM), which was proposed in IEEE 802.15.7, can support dimming and prevent flickering. The dimming of VPPM is achieved by controlling the pulse width hence it can be

viewed as a combination of PPM and PWM. Other pulse modulation schemes, such as DPPM, MPPM, and PAPM, were also proposed for different reasons [125]-[127].

C. Colour-shift keying (CSK)

Colour-shift keying is another intensity modulation scheme for VLC systems, which is also outlined in IEEE 802.15.7. The data is transmitted by modulating the intensity of the visible light to multicolour LEDs, typically red, green, and blue LEDs. CSK is less likely to harm human eyes by fluctuations compared to colour-intensity modulation (CIM) because of the stable light intensity. The colour that is visible to the human eye is defined in CIE 1931 [128] that the chromaticity of a colour can be represented by values x and y , as shown in Figure 2.24.

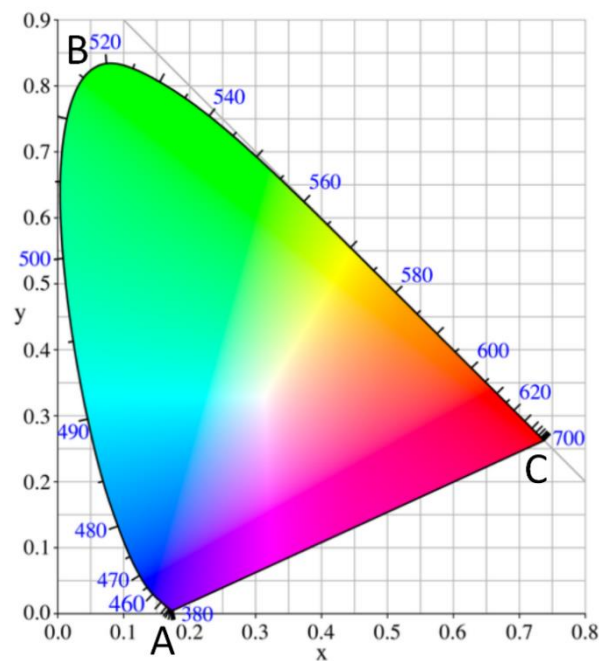


Figure 2.24 CIE 1931 colour space chromaticity diagram

IEEE 802.15.7 uses three colours LEDs (point A, B, and C shown in Figure 2.24) in the VLC system. The block diagram of a Trichromatic LEDs (TLED) VLC system is shown in Figure 2.25, where x_t , y_t represents the chromaticity, R_t , G_t and B_t

denotes the intensity of each LED. The intensity of the PDs of the receiver is represented by R_r , G_r and B_r .

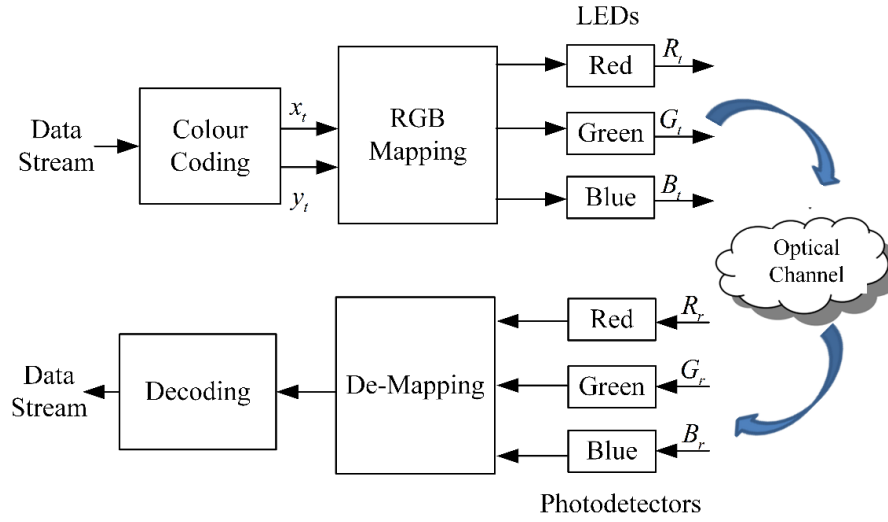


Figure 2.25 Block diagram of CSK modulation of VLC system

The input bit data is first mapped to x and y chromaticity and then convert to the intensity of each LED. The chromaticity satisfies that [129]

$$x = P_i x_i + P_j x_j + P_k x_k \quad (2.64)$$

and

$$y = P_i y_i + P_j y_j + P_k y_k, \quad (2.65)$$

where $P_i + P_j + P_k = 1$ and (x_i, y_i) , (x_j, y_j) , and (x_k, y_k) refer to the chromaticity values of each colour light and represent one CSK symbol.

D. OFDM

The above modulation schemes suffer from inter-symbol interference. OFDM is a more resilient modulation technique, which allows efficient equalization with single-tap equalizers. In RF systems, OFDM generates complex-valued bipolar signals which are not applicable for IM/DD VLC systems. The real part of an OFDM signal

can be got by reducing the bandwidth of the system and the negative part can be dealt with by some other approach to create a unipolar signal. Different OFDM forms for IM/DD systems can be used, DC biased optical OFDM (DCO-OFDM), asymmetrically clipped optical OFDM (ACO-OFDM), and others based on ACO-OFDM and DCO-OFDM [130].

An easy way to generate a unipolar OFDM signal is introducing a direct current (DC) bias to the bipolar OFDM signal. This way is known as DC-biased optical OFDM (DCO-OFDM). The block diagram of a DCO-OFDM transceiver is shown in Figure 2.26.

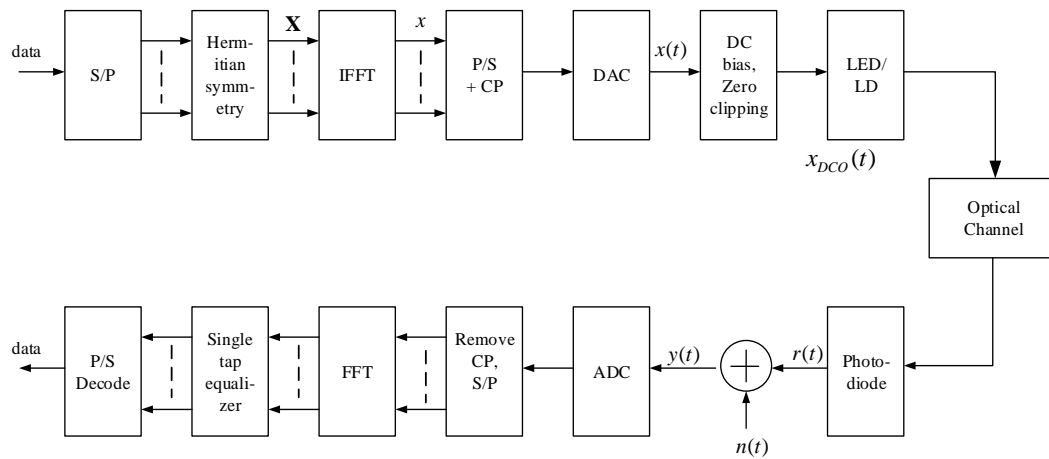


Figure 2.26 DCO-OFDM VLC system

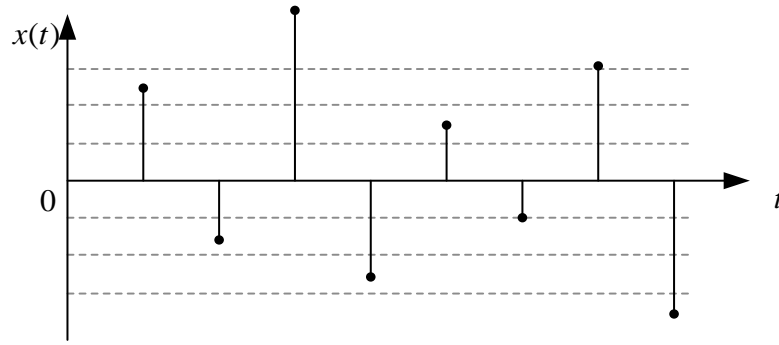
Let the baseband complex data signal before IFFT module be represented by $\mathbf{X} = [X_0, X_1, X_2, \dots, X_{N-1}]$. To make the output of IFFT be real value, \mathbf{X} needs to be Hermitian symmetry,

$$X_n = X_{N-n}^*, \quad 0 < n < N/2 \quad (2.66)$$

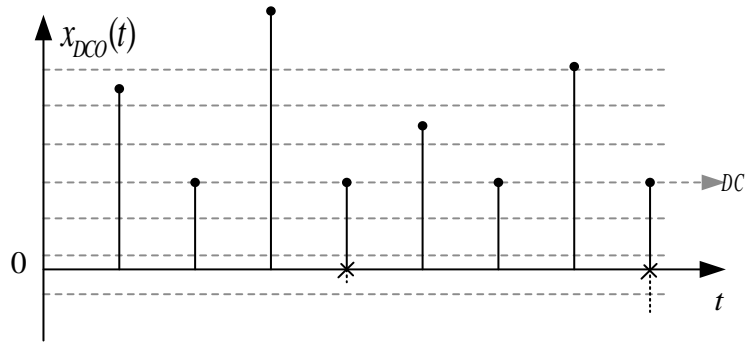
and $X_0 = X_{N/2} = 0$, where N is the point number of IFFT and n represents the n th subcarrier. If the input signal meets the above condition, the output signal of IFFT, x , is not a complex signal any more. After IFFT, the m th time-domain sample of x is

$$x_m = \frac{1}{N} \sum_{n=0}^{N-1} X_n \exp\left(\frac{j2\pi mn}{N}\right), \quad (2.67)$$

then the time domain signal x is converted to serial and a cyclic prefix (CP) is inserted as the guard band. After DAC, a DC bias is introduced to the analog signal $x(t)$ and the residual negative samples are clipped, as shown in Figure 2.27.



(a)



(b)

Figure 2.27 DCO-OFDM output: (a) Before adding DC bias; (b) After adding DC bias and zero clipping

At the receiver, a photodiode is utilized for the conversion from the optical signal to an electrical signal. The photodiode performs as the antenna in an RF system. Apart from that, the DCO-OFDM receiver is the same as that in an RF system. The electrical received signal $y(t)$ is impacted by additional white noise $n(t)$.

The addition of DC bias increases the transmit power and zero clipping will distort the information hence many other unipolar OFDM techniques without biasing have been proposed. ACO-OFDM is one of them, in which only the odd subcarriers are modulated with data. Hence the input of IFFT becomes $\mathbf{X}=[0, X_1, 0, X_3, \dots, X_{N-1}]$ with the same constraint as (2.30) and x is real and has the anti-symmetry property as:

$$x_m = -x_{m+N/2}, \quad 0 < m < N/2. \quad (2.68)$$

The other parts of the system are the same as DCO-OFDM systems. With ACO-OFDM, the output time domain signal is purely positive without any distortion.

Another modulation scheme named unipolar OFDM (U-OFDM) has been proposed in [131]. The positive part and negative part of the real bipolar OFDM time domain signal are separated into two different frames. It also generates a purely unipolar signal with no biasing. Both ACO-OFDM and U-OFDM utilize half bandwidth which sacrifices half system capacity. Hence, they have less spectral efficiency compared to DCO-OFDM.

3 CRS interference cancellation for Heterogeneous networks

This chapter introduces an interference cancellation method that takes CFO and TO into account. Firstly, the interfering model is presented with consideration of CFO and TO. Secondly, based on the interfering model, a cancellation solution is introduced. This cancellation solution relies on the interfering signal reconstruction subject to the interfering channel estimation and TO/CFO estimation. Finally, the interference is moved from the received signal and simulation results shows that this CRS interference cancellation method can improve the performance of the HetNet system. Note that since the 5G HetNet system is not standardized yet, the problem formulation of this chapter is based on LTE standard. Although the technique detail in future 5G may be different, the radio frame structure and the inter-cell interference model could be similar. Interference model. To the best of the author's knowledge, this is the first work that considers the CFO and TO when doing inter-cell interference cancellation for heterogeneous networks.

3.1 Interference model

As discussed in Section 2.3, enhanced inter-cell interference coordination was introduced in LTE Rel. 10 to handle the inter-cell interference. There are three categories of eICIC technique, time domain eICIC, frequency-domain eICIC, and power control eICIC.

3.1.1 Time domain eICIC

Time domain eICIC deals with the interference by restricting time domain resources. Almost blank subframe (ABS) and OFDM symbol shift are the two main time domain eICIC methods [113].

An almost blank subframe refers to a muted subframe, which has very low power, from the macrocell. By this way, the UE can be served by the picocell without the

interference from the macrocell during ABS time. This muting for the control and data channel is also defined as No-PDCCH case. And there is also a case that only the data channel is muted which is defined as Lightly-loaded PDCCH case, in which the macrocell transmits some signals, such as reference signal, physical downlink control channel (PDCCH), broadcasting channel (PBCH) and synchronization signals, to remain connection with the users.

OFDM symbol shift means the shifting of the OFDM symbol boundary from the macrocell by several OFDM symbols according to the subframe boundary of the UE. This technique can prevent the overlapping of the control channels from the picocell and the macrocell. However, the data channel from the macrocell may cause interference on the control channel of the UE, which can be avoided by muting the macrocell's data channel (PDSCH symbol muting) or consecutive subframe blanking. PDSCH symbol muting refers to the muting of the OFDM data symbols of the macrocell which causes high interference to the control channel from the picocell. While for consecutive subframe blanking, the ABS technique mentioned above is applied to the macrocell so that the overlapping of the macrocell's control channel and the picocell's control channel can be avoided.

Time domain techniques require strict time synchronization hence time offset is going to affect eICIC results severely. The interference cancellation method, which will be introduced later in this chapter, takes timing offset into consideration.

3.1.2 Frequency domain eICIC

A general way to reduce the inter-cell interference in the frequency domain is to allocate the resource blocks in a way that the macrocell RBs will not overlap with the picocell RBs, as shown in Figure 3.1. The synchronization and reference signals of the macrocell and picocell are orthogonal with each other.

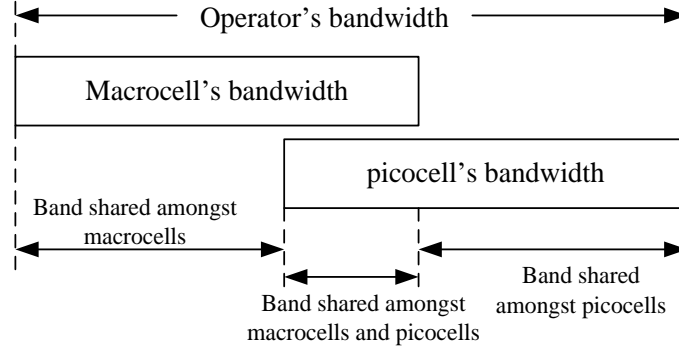


Figure 3.1 HetNet frequency sharing

This concept is presented in [115] and sharing the bandwidth amongst the macrocell and picocells is also studied in [116] as a commercial case.

3.1.3 Power control techniques

Power control techniques are proposed to alleviate the inter-cell interference problems in HetNet [114]. Power control techniques are applied to the picocells downlink transmission. This type of approach reduces the total throughput of the picocell users because of the loss of SNR but it increases the performance of the UEs affected by the macrocell. Different kinds of DL power control techniques can be used in picocells. They focus on different parameters such as strongest eNodeB received power, path loss between macrocell and the UE, the SINR of the picocell user and that of the macrocell.

The time domain scheme is considered in this dissertation that some almost blank subframes are transmitted by the macrocell to reduce the interference.

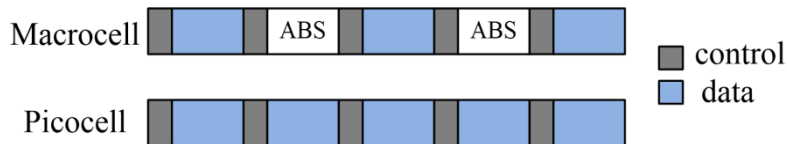


Figure 3.2 Illustration of almost blank subframe

As shown in Figure 3.2, there is a chance that the control signals from the macrocell and the picocell can conflict with each other. Hence, a further eICIC is introduced in

Rel. 11 to deal with the interference brought by CRS. CRS interference can be divided into two scenarios, colliding and non-colliding. The former one implies that CRS from the macrocell and picocell overlap with each other which cause severe performance loss. The latter one means that the CRS brings interferences to the data/control resource element, which will still reduce the system performance. Figure 3.3 below shows the CRS interference of both scenarios.

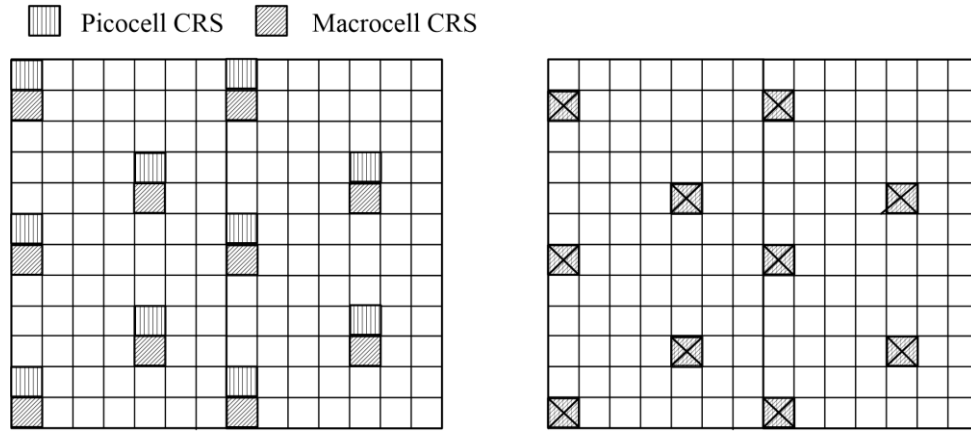


Figure 3.3 Non-colliding scenario (left) and colliding scenario (right)

The LTE downlink OFDM signal can be given as:

$$\begin{aligned}
 x_i(n) &= \frac{1}{\sqrt{K}} \sum_{k=0}^{K-1} X_{i,k} e^{j2\pi nk/K} \\
 &= \frac{1}{\sqrt{K}} \sum_{k=0}^{K-1} (d_{i,k} + p_{i,k}) e^{j2\pi nk/K},
 \end{aligned} \tag{3.1}$$

where the time domain signal $x_i(n)$ is produced from K-point Inverse Fast Fourier Transform (IFFT) of $X_{i,k}$, which is the frequency domain signal at the i^{th} symbol, n is the discrete time sequence index. $d_{i,k}$ and $p_{i,k}$ denotes the modulated data and pilot (cyclic prefix) respectively.

Taking the multipath wireless channel into account, the signal received by the user terminal can be modelled as:

$$y_i(n) = \sum_{l=0}^{L-1} h_l x_i(n - \tau_l) + w(n), \quad (3.2)$$

where the multipath channel has L paths and h_l , τ_l represents the gain and delay of each path; $w(n)$ is the discrete additional Gaussian white noise (AWGN). Consider the inter-cell interference occasion, the interfered signal at the receiver side is:

$$\begin{aligned} y_i(n) &= y_i^m(n) + y_i^p(n) + w(n) \\ &= \frac{1}{\sqrt{K}} \sum_{k=0}^{K-1} H_{i,k}^m X_{i,k}^m e^{j2\pi nk/K} + \frac{1}{\sqrt{K}} \sum_{k=0}^{K-1} H_{i,k}^p X_{i,k}^p e^{j2\pi nk/K} + w(n), \end{aligned} \quad (3.3)$$

where $y_i^m(n)$ and $y_i^p(n)$ represent the interfering signal from the macrocell and the serving signal from the picocell respectively; $H_{i,k}^m$ and $H_{i,k}^p$ denote the channel frequency response from the macrocell to the UE and from the picocell to the UE respectively.

Note that this thesis mainly focuses on the hardware impairments impact on the system. Timing offset (TO) and carrier frequency offset (CFO) are two main factors that are considered for inter-cell interference cancellation. Since the configuration consists the macrocell, picocell and the user terminal, the TO and CFO model are more complicated than the non-HetNet configuration. Although the TO/CFO between the picocell and the UE can be estimated via the synchronization signal, the interfering CRS contains the TO/CFO between the macrocell and the UE which will affect the system performance. In this situation, the non-colliding scenario with TO/CFO is shown in Figure 3.4.

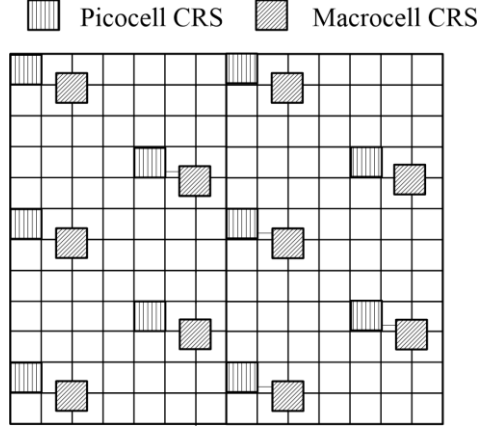


Figure 3.4 non-colliding CRS interference with TO/CFO

And (3.3) can be rewritten as:

$$\begin{aligned}
 y_i(n) &= y_i^m(n) + y_i^p(n) + w(n) \\
 &= \frac{1}{\sqrt{K}} \sum_{k=0}^{K-1} H_{i,k}^m X_{i,k}^m e^{j2\pi(n+\Delta d)(k+\Delta f)/K} + , \\
 &\quad \frac{1}{\sqrt{K}} \sum_{k=0}^{K-1} H_{i,k}^p X_{i,k}^p e^{j2\pi nk/K} + w(n)
 \end{aligned} \tag{3.4}$$

where Δd and Δf represent the relative timing offset and carrier frequency offset between the macrocell and the picocell. By the K-point FFT, the received signal can be transferred to the frequency domain OFDM symbols [40]:

$$\begin{aligned}
 Y_{i,k} &= Y_{i,k}^m + Y_{i,k}^p \\
 &= \sum_{n=-N/2}^{N/2} e^{\frac{j2\pi n \Delta d}{N}} H_{i,k}^m X_{i,k}^m \Phi_i + H_{i,k}^p X_{i,k}^p + W_{i,k} ,
 \end{aligned} \tag{3.5}$$

where $H_{i,k}^m$ and $H_{i,k}^p$ denotes the coefficients of macrocell channel and picocell at k^{th} subcarrier respectively; Φ_i is the inter-carrier interference (ICI) caused by the relative CFO Δf . During ABS duration, the macrocell only sends control and CRS signals, as shown in Figure 3.4, the data resource element signal model of the UE can be modelled as:

$$Y_{i,k} = \sum_{n=-N/2}^{N/2} e^{\frac{j2\pi n \Delta d}{N}} H_{i,k}^m p_{i,k}^m \Phi_i + H_{i,k}^p d_{i,k}^p + W_{i,k} . \tag{3.6}$$

According to (3.6) above, the received data is corrupted by

$$\sum_{n=-N/2}^{N/2} e^{j2\pi n\Delta d/N} H_{i,k}^m p_{i,k}^m \Phi_i .$$

3.2 CRS Interference cancellation

In this chapter, the relative timing and frequency offset between the macrocell and picocell are considered for non-colliding scenario inter-cell interference cancellation. The timing offset as well as the carrier frequency offset is estimated by making use of the synchronization signal. Then the interfering channel is estimated based on the character of channel statistics. Finally, the interfering signal is reconstructed and cancelled based on the TO/CFO estimation and the interfering channel estimation.

3.2.1 System model

The interference cancellation method proposed in this chapter is shown in Figure 3.5 below. The received signal $y_i(n)$ is first fed to an FFT module to convert to the frequency domain $Y_{i,k}$, where i is symbol index, n is timing index within one symbol and k is sub-carrier index. Then the timing and frequency offset are estimated based on the interfering CRS modelling, which provides PSS/SSS and CRS signal. Utilizing the interfering CRS signal, the interfering channel estimation is performed later to reconstruct the interference signal from the macrocell. At last, the CRS interference of the macrocell is rebuilt and cancelled which improves the overall system performance. $\tilde{y}^{(p)}(n)$ denotes the time domain signal after interference cancellation.

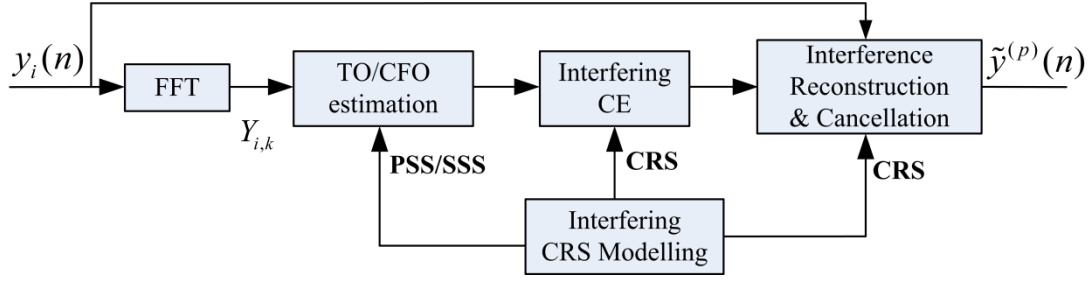


Figure 3.5 CRS interference cancellation scheme

3.2.2 TO/CFO estimation

Generally speaking, the timing and carrier frequency offset of an OFDM signal are estimated based on some redundant or known information inserted in the transmit signal, such as cyclic prefix [47] and the pilot data [48]. However, for this HetNet configuration, in the ABS duration, the CP has a very low power which makes it hard for the synchronization. For LTE, the primary and secondary synchronization signal can be used for TO/CFO estimation.

The PSS/SSS signal can be generated based on the specification which may differ in 5G network. But the estimation method can be the same, in this chapter, the cross-correlation in the time domain is used [49]:

$$\{\Delta\tilde{n}, \Delta\tilde{f}\} = \arg \max_{\Delta n, \Delta f} |C(\Delta n, \Delta f)|, \quad (3.7)$$

where

$$C(\Delta n, \Delta f) = \sum_{m=1}^M s_i^*(m) r(\Delta n + m) e^{\frac{-2\pi\Delta f n}{N}}; \quad (3.8)$$

$r(n)$ and $s_i(n)$ represent the received signal and the symbol which is generated at the receiver side; Δn and Δf denote the timing and frequency difference. Both $r(n)$ and $s_i(n)$ contain PSS/SSS so that the correlation results will show the timing and frequency difference. The correlation length M will decide the gain of the output

which needs to be multiple times of the synchronization symbols to get a high correlation gain especially when SNR is low.

With the TO and CFO estimated, the interfering signal can be reconstructed more precisely and the received signal after interference cancellation will be cleaner.

3.2.3 Interfering channel estimation

Channel estimation basically means the estimation of the channel state information (CSI). Normally there are two types of channel state, instantaneous channel state and statistical channel state. Instantaneous channel state refers to the case that the channel condition is known and remains unchanging during a very short period of time. This short-term channel state can be modelled by an impulse response of a filter. Utilizing this impulse response estimation, the received signal can be recovered. Statistical CSI represents the channel character and is viewed from the long-term statistical point of view. Different fading distribution, such as Rayleigh fading, is used to describe it. To estimate CSI, the short-term CSI is often the estimation target since the channel condition is varying. A popular method to achieve the goal is based on some training sequences, or pilot sequences, where some known signal is sent out from the transmitter. In this work, the training sequence is the cell-specific reference signal (CRS).

Although the original CRS from the macrocell can be generated according to the specification, it cannot be subtracted from the received signal directly because the received macrocell CRS signal has been changed by the channel. Therefore, it is important to get the channel response from the macrocell to the UE to estimate the CRS interference element inside the received signal.

There exist many channel estimation algorithms. Least-square (LS) estimation and Minimum Mean Square Error (MMSE) are two most popular ones among these

algorithms. Although MMSE estimation can achieve a higher gain (10-15dB) in SNR compared with LS estimation, the latter is much less complex in terms of computation complexity. Therefore, LS is more practically suitable especially at the UE side since normally the UE is a low-cost receiver. Hence the interfering channel estimation is accomplished by LS estimation according to the known training sequence as shown below:

$$p^m = \frac{1}{\sqrt{2}}(1 - 2c(2n)) + j \frac{1}{\sqrt{2}}(1 - 2c(2n+1)), \quad (3.9)$$

where $c(n)$ is a length-31 Gold sequence which is generated based on the system parameters such as cell ID and antenna port, m indicates the sequence from the macro cell. As discussed above, the PSS/SSS signal still exists and can be used for TO/CFO estimation. The cell ID is detected after the PSS/SSS synchronization step hence the CRS of the interference signal can be produced. Using LS channel estimation, the interfering channel can be estimated as follows:

$$\begin{aligned} \tilde{H}_{i,k}^m &= \frac{Y_{i,k}}{p_{i,k}^m} \\ &= \frac{H_{i,k}^m p_{i,k}^m + H_{i,k}^p d_{i,k}^p + W_{i,k}}{p_{i,k}^m} \\ &= H_{i,k}^m + H_{i,k}^p d_{i,k}^p / p_{i,k}^m + W_{i,k} / p_{i,k}^m \end{aligned} \quad (3.10)$$

The above equation shows that in the interfering channel estimation stage, the term $H_{i,k}^p d_{i,k}^p / p_{i,k}^m$ from the picocell signal acts as an interference, which is also known as a co-channel interference [50], will affect the accuracy of the estimation result. A cancellation method for this co-channel interference should be considered.

As the data resource elements are not certain, the co-channel interference term cannot be compensated directly. However, according to [50], the distribution of the co-channel interference can be represented by a Gaussian stochastic process if the resource block is large and non-Gaussian if the resource block is small. Since the

mean value of the stochastic converges to zero, therefore, from (3.10), the following equation can be obtained:

$$\begin{aligned}\mathbb{E}\{\tilde{H}_{i,k}^m\} &= \mathbb{E}\{H_{i,k}^m\} + \mathbb{E}\{H_{i,k}^p d_{i,k}^p / p_{i,k}^m\} + \mathbb{E}\{N_{i,k} / p_{i,k}^m\} \\ &\approx \mathbb{E}\{H_{i,k}^m\}\end{aligned}\quad (3.11)$$

where the expectation of the inter-cell channel response is derived. From the practical implementation point of view, the expectation can be simply realized by moving average. Assume that the interfering channel does not change within a coherent time duration Δt , the moving average window length can be as large as $f_s \Delta t$. Then the interfering channel can be approximated by its expectation $\mathbb{E}\{\tilde{H}_{i,k}^m\}$ especially in the slow fading scenario. The channel estimation algorithm only applies to the non-colliding scenario that the picocell data resource elements are collided by the macrocell CRS signal. The colliding scenario is beyond the scope of this work.

3.2.4 Interfering signal reconstruction and reduction

As we know, the interfering CRS signal is basically the known CRS signal propagate through a wireless channel to affect the UE received signal. Since the transmitted CRS signal is known, we can simply let it pass the interfering channel estimated above and then it can be reconstructed. At the same time, the timing and frequency offset need to be considered to improve the reconstruction accuracy.

Since the time duration of CP is normally smaller than timing offset, reconstructing CRS in the frequency domain may result in ISI and cause the cancellation to be inaccurate. In this proposed method, the reconstruction is carried out in the time domain and the cancellation is performed by subtracting the reconstructed signal from the received signal directly:

$$\tilde{y}^p(n) = y(n) - \tilde{x}(n + \Delta\tilde{n})e^{\frac{j2\pi\Delta\tilde{f}n}{N}} * \tilde{h}^m, \quad (3.12)$$

where ‘*’ means circular convolution and $\tilde{x}(n)$ is the reconstructed macrocell CRS interference and

$$\tilde{h}^m = IFFT \left\{ \tilde{H}_{i,k}^m \right\}. \quad (3.13)$$

After this step, the interfering signal is reconstructed and cancelled from the received signal. The performance after this cancellation is shown in the experiment results in next section.

3.3 Simulation results

The performance of the proposed method will be evaluated in this section via Monte Carlo simulation. By comparing the block error rate (BLER), which is a key performance parameter for LTE rather than bit error rate (BER), of the UE with or without the proposed method, the effectiveness of this method can be proved.

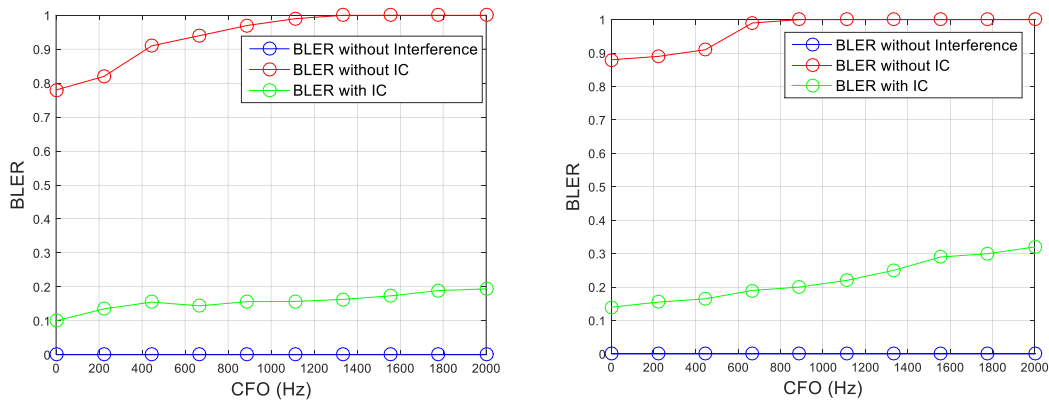
To test the robustness of the method, different modulation and coding schemes (MCS) are utilized in the simulation. The bandwidth of the picocell is set to be 10MHz while the bandwidth of the macrocell is 5MHz when it is transmitting ABS. Since the non-colliding scenario is considered, the macrocell CRS signal overlaps with the data REs of the picocell. In this simulation, the original picocell signal is interfered by the macrocell CRS signal and the mixed signal passes through a fading wireless channel where the propagation delay is smaller than CP duration. In addition, different TO and CFO are applied to the mixed interfered signal. The LTE Extended Pedestrian A (EPA) channel model is used as the fading channel with a different propagation delay. The cell search is assumed to be ideal as the main goal of this work is the interference cancellation. Under this assumption, the PSS/SSS/CRS can be constructed. The key simulation configuration is shown in Table 3-1:

Table 3-1 Simulation Configuration

Parameters	Picocell	Macrocell
------------	----------	-----------

Bandwidth (MHz)	10	5
Transmit power (dB)	0	0
MCS	5/20	N/A
Channel Model	EPA 0/30/70/90	EPA 0/30/70/90
SNR (dB)	15	15

Figure 3.6 (a)-(d) below shows the BLER performance of the IC algorithm under different EPA channel propagation delay σ (0/30/70/90 ns) and frequency offset. EPA channel means the Doppler frequency is 5Hz and the fading type is Rayleigh fading and the reason of choosing EPA channel is because the picocell is more likely serving slow moving objects, such as pedestrians etc. The red BLER curve of each figure shows the BLER simulation results under different CFO without any interference cancellation. The green curve shows the BLER with the proposed IC algorithm and the blue curve shows the BLER results without interference. It shows that, under different propagation delays, the proposed IC algorithm gives a significant improvement for the system BLER performance compared to the situation without interference cancellation. The reason why the BLER without interference is 0 is because of the limited simulation time within which there was no error appeared. However, finding out the real BLER of the situation without interference is not the target of this simulation, as long as we can see the BLER with IC is much lower than that without IC, the effectiveness of the proposed IC method is proved.



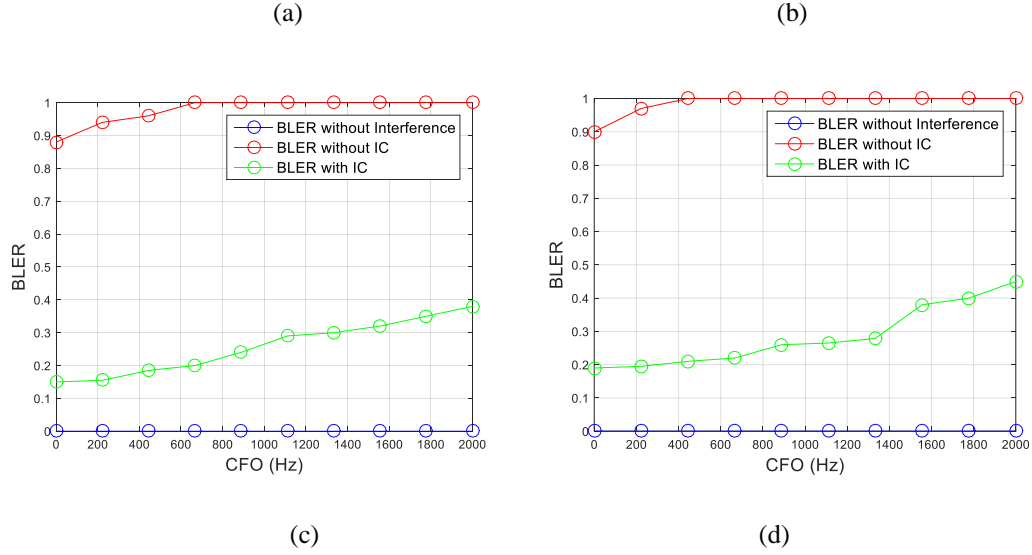
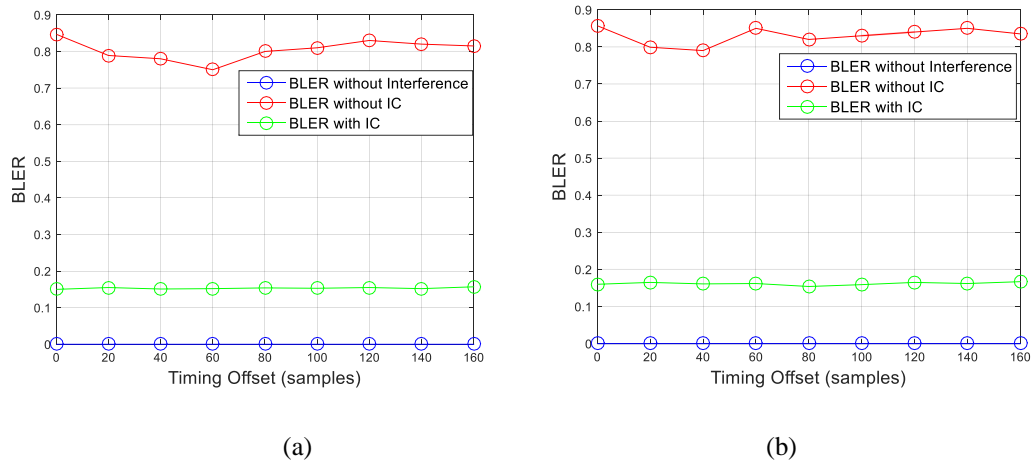
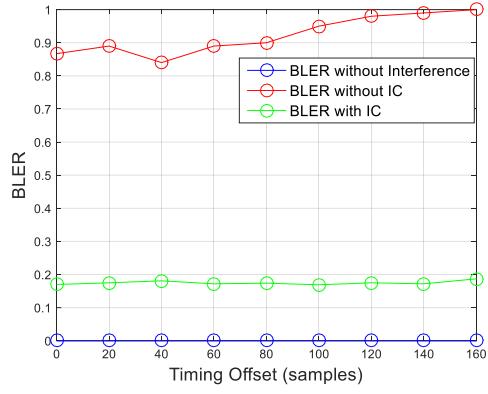


Figure 3.6 BLER performance of different CFO and propagation delay:

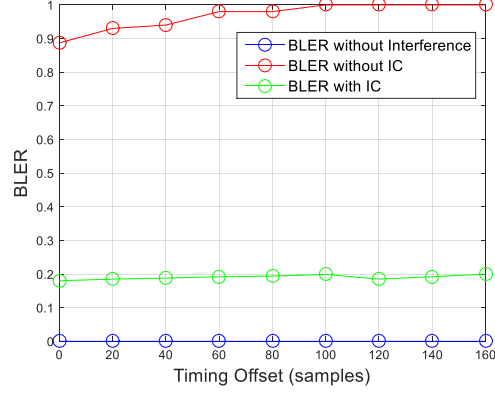
(a) $\sigma = 0$, (b) $\sigma = 30$, (c) $\sigma = 70$ (d) $\sigma = 90$

Figure 3.7(a)-(d) below show the BLER performance of the IC algorithm under different EPA channel propagation delay and timing offset. The red BLER curve of each figure shows the BLER simulation results under different TO without any interference cancellation. The green curve shows the BLER with the proposed IC algorithm and the blue curve shows the BLER results without interference. It shows that, under different propagation delays, the proposed IC algorithm gives a great improvement for the system BLER performance compared to the situation without interference cancellation.





(c)



(d)

Figure 3.7 BLER performance of different TO and propagation delay:

(a) $\sigma = 0$, (b) $\sigma = 30$,(c) $\sigma = 70$ (d) $\sigma = 90$

The BLER performance of the proposed algorithm under 5Hz Doppler frequency and different SNR is shown Figure 3.8 below. In this experiment, MCS is chosen to be 20 that the modulation order is 6 (64 QAM). When SNR is less than 9dB, the proposed IC algorithm cannot demodulate the received signal, but when SNR is larger than 9dB, the receiver with the proposed IC algorithm can achieve a much better BLER performance than the case without IC. In general, the system performance is significantly improved compared to the BLER without IC.

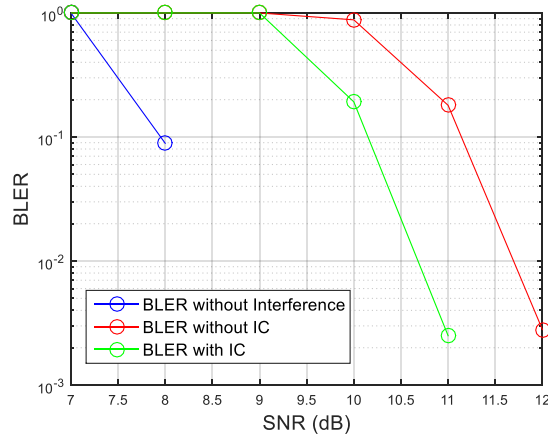


Figure 3.8 BLER performance of SNR

3.4 Summary

In this chapter, a novel inter-cell interference cancellation algorithm is proposed for HetNet receiver based on the interference CRS signal reconstruction. The interference model is first studied and analysed with CFO and TO being taken into consideration. The interference is generated by the residual control information from the macrocell during the ABS duration. Second, the interference cancellation scheme is discussed which can be divided into three steps. In the first step, the TO and CFO are estimated via the known PSS/SSS signal from the macrocell under the assumption of ideal cell search. Then TO and CFO are compensated, and the interfering channel estimation is applied for reconstructing the interfering CRS in the second step. Finally, the interference is cancelled in time domain after reconstruction and it is subtracted from the original received signal. The performance is shown in the simulation results which are obtained via Monte Carlo simulation where different TO/CFO/SNR and MCS are applied. The results indicate that the proposed IC algorithm can achieve very good performance in different channel conditions. In this work, only the non-colliding SISO scenario is considered. The MIMO and colliding scenario are more complex and need further research in the future.

4 Impairments Calibration for massive MIMO systems

This chapter studies the hardware impairments calibration for massive MIMO system. First, the impact of hardware impairments on precoding of a MIMO system is investigated, and then, the hardware impairments impact is expanded for a massive MIMO system because of its different configuration and usage scenario compared to a general MIMO system. In the third section, the hardware impairments calibration methods are introduced including the proposed closed-loop reciprocity calibration method for massive MIMO systems. In this proposed method, the first part is an optimized relative calibration which improves the overall calibration accuracy by optimizing the reference transmit/receive hardware chain. The second part is a test device-based calibration which can support online calibration and also improve the calibration accuracy compared to the existing calibration methods. This closed-loop calibration method can not only improve the reciprocity calibration accuracy but also support online calibration without any additional hardware circuit in the base station (BS). The performance of this method is evaluated by computer simulation and laboratory experiment.

4.1 Massive MIMO impairments

MIMO is a technique that has become an essential part of wireless communication and has been used in many standards such as IEEE 802.11n, IEEE 802.11a, HSPA+ (3G), LTE and WiMAX (4G). MIMO allows the base station to transmit more than one data stream over the same radio channel which improves the system capacity dramatically. According to [52], “the more antennas the transmitter/receiver is equipped with, the more degrees of freedom that the propagation channel can provide, the better the performance in terms of data rate or link reliability”. In LTE, it allows up to eight antennas at the base station, but it is not enough for the data throughput demand in the future. Massive MIMO is believed to be a promising candidate for the 5G network and beyond. Hundreds of antennas can be deployed at the base station

which can increase the capacity up to 100 times compared to the conventional system [53]. A MIMO system can be used as a point-to-point system or a multiuser system [54]. For a point-to-point MIMO channel that is equipped with N_t antennas at the transmitter side and N_r antennas at the receiver side, the achievable data rate scales as $\min(N_t, N_r) \log(1 + \text{SNR})$ where SNR represents the receiving signal-to-noise ratio. However, this kind of scenario requires expensive multiple-antenna user terminals. More impressive gains can be obtained for multiuser MIMO systems because multiple users can be served simultaneously. Multiuser MIMO is more flexible since the terminals in this scenario can be cheap and equipped with a single antenna.

The capacity improvement of MIMO is achieved by utilizing some signal processing categories such as beamforming, precoding, and spatial multiplexing. For the traditional beamforming or single-stream beamforming, the transmit antennas emit the same signal with appropriate weightings so that the antenna array pattern can point to the single-antenna receiver and the received signal power is maximized. One of the benefits is to increase the signal gain of the receiver by constructive summation of all the channels. The other one is to reduce interference to other receivers. These benefits can only improve the received signal SNR and the system capacity improvement is very limited due to Shannon theory. Multiple antenna receivers can improve the capacity by $\min(N_t, N_r)$ times ideally as discussed before, but the traditional single-stream beamforming will not apply. This can be solved by precoding that can support multiple streams transmission. In point-to-point systems, precoding refers to the scenario that multiple data streams, which are transmitted by the transmit antennas, are multiplied with different weightings. Precoding can maximize the receiver output throughput. However, since the point-to-point MIMO system is not very flexible and not very appropriate for cellular networks due to the multi-antenna receiver's physical size and cost, another scenario, that the base station is equipped with multiple

antennas, but the receiver has only one antenna, needs to be considered. This scenario can be simply referred to as multi-user MIMO.

Spatial multiplexing (SM) is a transmission technique that can be used in multi-user MIMO systems. By transmitting separately encoded and independent streams, the same space is multiplexed, and the overall channel capacity can be increased since the data streams are transmitted in parallel. Although for each receiver, the throughput is still limited by spectral bandwidth and coding scheme, the system capacity can be improved by ideally $\min(N_t, N_r)$ times as well. Spatial multiplexing can combine with precoding if channel state information is available. If CSI is not known, diversity coding techniques can be used which are beyond the scope of this work.

By utilizing spatial multiplexing, massive MIMO systems can achieve a great capacity improvement if the base station has a good enough knowledge of the downlink channel. In LTE standard, the base station sends out some training sequence to the UE. The UE performs channel estimation based on the training sequence and feeds the channel estimates back to the base station. However, it is not going to be applicable in a massive MIMO system because the training sequences of the transmit channels need to be orthogonal to each other and it is not practical to maintain the orthogonality for so many channels due to limited time and frequency resources. One solution to this problem is using indirect channel estimates which can be got from uplink channel estimation. Thanks to channel reciprocity, the downlink channel response is the same as uplink channel response ideally so that the massive MIMO downlink channel estimation can be replaced by uplink channel estimation when the system works in time division duplex (TDD) mode. For multi-user MIMO scenario, the user has only a single antenna hence the uplink channel training sequence has no orthogonal problem. However, the downlink channel and uplink channel are not always reciprocal because of hardware impairments especially when the base station

uses a direct conversion transceiver (DCT) and some cheap components with the aim of reducing the overall hardware cost of massive MIMO systems [55][56].

Although the impact of large impairments brought from the inexpensive hardware can vanish asymptotically and be mitigated by compensation algorithms [20], the residual impairments such as the power mismatch of each channel, will break the reciprocity between downlink and uplink, which makes reciprocity calibration one of the main challenges of massive MIMO systems' implementation [11].

4.2 MIMO systems Hardware impairments model

In Chapter 2, some basic types of hardware impairments of a transceiver are discussed. The impairments model of a MIMO system is more complicated as more channels with impairments are introduced. For a massive MIMO system in the future, the behaviour is going to change because of the impairments' randomness of each channel can be mitigated to some extent [20]. In massive MIMO system, some cheap transceiver structure such as direct conversion transceiver (DCT) and inexpensive components may be used so as to reduce the whole system cost, which will bring in more severe impairments.

4.2.1 Single antenna DCT impairments model

Despite the advantages of low-cost, simplicity and flexibility of DCT, some practical problems arise when implementing a direct-conversion transmitter because of the non-linearity of the radio frequency (RF) component, such as in-phase (I) and quadrature (Q) mismatches, clock jitter, power amplifier (PA) distortion and inter-modulation [64][65].

In an ideal DCT system, the RF signal can be modelled as:

$$\begin{aligned} y(t) &= \Re\left(x(t)e^{j2\pi f_c t}\right) \\ &= x(t)e^{j2\pi f_c t} + x^*(t)e^{-j2\pi f_c t}, \end{aligned} \tag{4.1}$$

where $x(t)$ is the complex baseband signal and $x(t) = x_I(t) + j \cdot x_Q(t)$. According to (4.1), $x(t)$ is shifted to RF frequency band f_c and its mirror band $-f_c$. Ideally there is no overlap between f_c and $-f_c$, however, due to the hardware impairments of real components, overlapping exists. As shown in Figure 2.15, the impairments of a DCT system include:

- 1). The amplitude and phase difference between the LO branches: g and Q .
 - 2). The DC offset of I and Q branch.
 - 3). The baseband gain and phase difference between I and Q branch.
 - 4). The mixer conversion gain and amplifier gain difference between I and Q branch.
- 1) and 2) are generally frequency-independent while 3) and 4) are frequency-dependent effects.

Apparently, by shifting the baseband frequency response to RF, the baseband impairments can be modelled in RF band. The frequency response of the equivalent RF model of FD impairments can be modelled by $h_I(t)$ and $h_Q(t)$. Therefore, the impaired RF signal can be written as:

$$\begin{aligned}
 y_{RF}(t) &= (x_I(t) + D_I) \cdot g \cdot \cos(2\pi f_c + \theta) * h_I(t) \\
 &\quad + (x_Q(t) + D_Q) \cdot \cos(2\pi f_c + \pi/2) * h_Q(t) \\
 &= x(t) e^{j2\pi f_c t} * g_1(t) + x^*(t) e^{j2\pi f_c t} * g_2(t) \quad , \\
 &\quad + x^*(t) e^{-j2\pi f_c t} * g_1^*(t) + x(t) e^{-j2\pi f_c t} * g_2^*(t) \\
 &\quad + D(t)
 \end{aligned} \tag{4.2}$$

where

$$\begin{aligned}
 g_1(t) &= \frac{g \cdot e^{j\theta} h_I(t) + h_Q(t)}{2} \\
 g_2(t) &= \frac{g \cdot e^{j\theta} h_I(t) - h_Q(t)}{2}
 \end{aligned}$$

and $D(t)$ is caused by DC offset. In (4.2), "*" stands for convolution and $()^*$ denotes conjugation.

Equation (4.2) shows that the ideal term $x(t)e^{j2\pi f_c t} + x^*(t)e^{-j2\pi f_c t}$ is distorted by $g_1(t)$ and $D(t)$. $D(t)$ can be easily cancelled by DC offset calibration [66]. Besides, the desired signal is also impaired by $x^*(t)e^{j2\pi f_c t} * g_2(t)$ and $x(t)e^{-j2\pi f_c t} * g_2^*(t)$ which can be viewed as a result of the image signal distorted by $g_2(t)$. The frequency dependent $g_1(t)$ results in the common gain of I and Q branch, while $g_2(t)$ gives rise to I/Q imbalance of the system.

To perform impairments calibration in the baseband, it is important to get a baseband equivalent model of the impairments. Let $H_I(f)$ and $H_Q(f)$ be the baseband equivalent of the frequency response of RF imbalanced filter, thus the baseband equivalent model after DC offset calibration can be given as [67][68]:

$$Y_{RF}(f) = X(f)G_1(f) + X^*(-f)G_2(f), \quad (4.3)$$

where

$$\begin{aligned} G_1(f) &= g \cdot e^{j\theta} H_I(f) + H_Q(f) \\ G_2(f) &= g \cdot e^{j\theta} H_I(f) - H_Q(f) \end{aligned}.$$

According to (4.3), both the amplitude and phase of the RF signal are distorted. For simplicity, let

$$\begin{aligned} Y_{RF}(f) &= M(f) \cdot Y(f) \\ &= |M(f)| \cdot e^{j\phi_M(f)} \cdot Y(f), \end{aligned} \quad (4.4)$$

where $M(f)$ stands for the transceiver hardware impairments which are frequency-dependent with the amplitude of $|M(f)|$ and phase of $\phi_M(f)$. Particularly, $M(f) = T(f)$ or $R(f)$, where $T(f)$ and $R(f)$ stand for the transmitter impairments and receiver impairments respectively. The demodulation model as shown in Figure

4.1 shows the conventional I/Q-demodulation-based model from which how the impairments impact the received signal can be seen. In massive MIMO systems with separate local oscillators (LO), each RF chain is independent, so the impairments are independent as well. Here, only one of the RF chains is analysed. In next section, the impact of hardware impairments on MIMO and massive MIMO system will be studied.

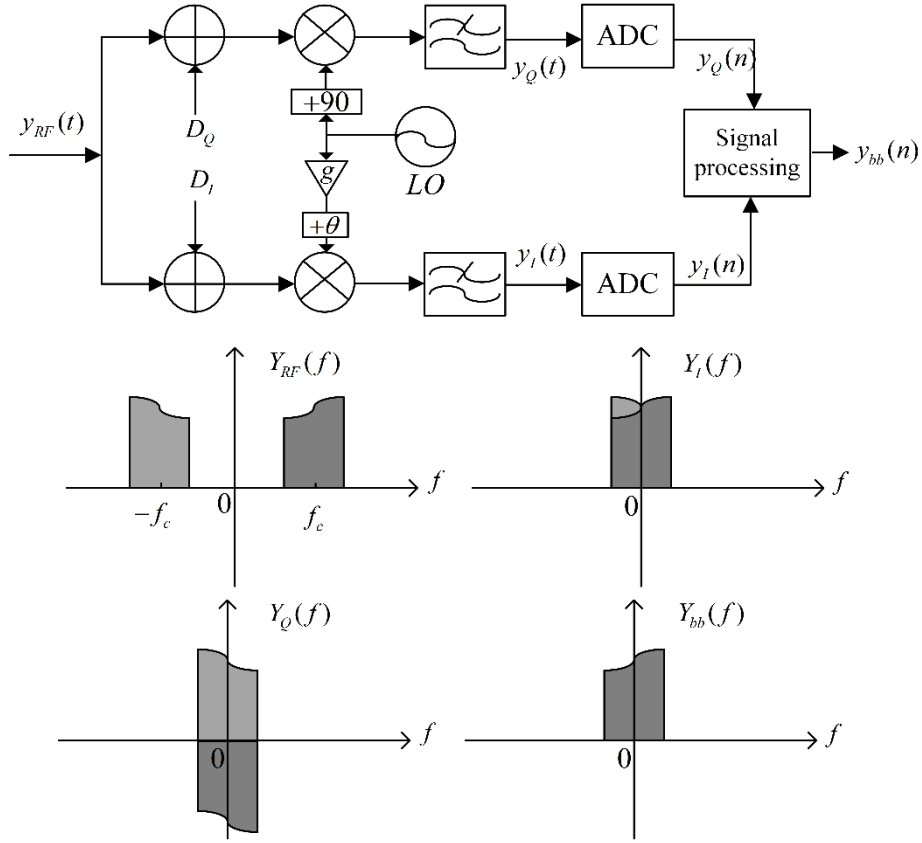


Figure 4.1 DCT demodulation model

4.2.2 MIMO and massive MIMO systems hardware impairments model

Assume that every hardware chain of a MIMO or massive MIMO system is using DCT structure, the hardware impairments cannot be mitigated completely by compensation algorithms. For example, the MIMO RF chains are using different power amplifiers, the output of different amplifiers may result in distorted beam pattern for beamforming and different receive gain at the UE side. Therefore, the

hardware impairments of MIMO systems need to be considered accordingly. As discussed before, the capacity performance of a multi-user MIMO system has a high relationship with the knowledge of channel state information. For a general MIMO system where the base station is only equipped with several antennas, the CSI can be obtained via downlink channel estimation. While for massive MIMO systems, it is not practical to get the CSI via downlink training, the solution is utilizing channel reciprocity and let the users transmit uplink training pilot using uplink channel estimates instead of downlink channel estimates. Hence, this subsection studies these two scenarios separately.

A. MIMO systems impairments model for beamforming

As mentioned before, the future MIMO system is very likely to support hybrid beamforming that includes traditional beamforming and multi-user precoding. The former focuses the beam angle to the target user and the latter serves multiple users simultaneously. Under the practical hardware impairments, the beam pattern after beamforming can be distorted which may results in a situation that the gain of the received signal is reduced dramatically or even the user cannot receive anything.

The BS transmitter adjusts the direction of the beam pattern by applying different weightings. However, due to the impact of impairments, the beamforming results can be different in terms of the beamforming gain and beam angle. The beamforming pattern with impairments can be given as:

$$F(\theta) = \sum_{n=0}^{N-1} T_n(f) P_n e^{j(\xi n d \sin \theta + n \alpha)}, \quad (4.5)$$

where $T_n(f)$ is the frequency domain impairments representation of the n th antenna. Figure 4.2 below shows the impaired transmit beamforming scheme.

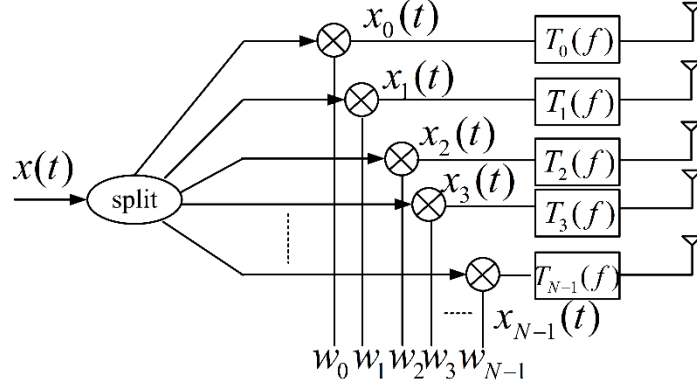


Figure 4.2 Impaired transmit beamforming

Each hardware chain is impacted by hardware impairments which are represented by a filter module whose frequency response is $T_n(f)$.

B. MIMO systems impairments model for precoding

The base station of a MIMO system sends pilots to the user terminals to get the downlink channel state information and then the UEs feed the channel estimates back to the base station. Using the same system configuration as described in Section 2.4.2 that the base station has N_t antennas and it serves K single-antenna user terminals simultaneously. Through precoding, the whole system capacity will ideally be improved by $\min(N_t, K)$ times that of a single antenna system. In practice, hardware impairments will be applied to the system and the improvement will be reduced. Consider that the impairments are frequency selective and the modulation scheme is OFDM. The system model of (2.13) will be changed to:

$$\mathbf{Y}_i = \mathbf{H}_{D,i} \mathbf{X}_i + \mathbf{N}_i, i = 1, 2, \dots, M, \quad (4.6)$$

where $\mathbf{H}_{D,i} = \mathbf{R}_{i,U} \mathbf{B}_i \mathbf{T}_{i,B}$ represents the impaired DL channel matrix at subcarrier i ; M is the OFDM subcarrier number. $\mathbf{T}_{B,i} = \text{diag}(t_{B,i,1}, \dots, t_{B,i,N_t})$ and $\mathbf{R}_{U,i} = \text{diag}(r_{U,i,1}, \dots, r_{U,i,K})$ denote the BS transmitter frequency-dependent impairments and the terminal receiver impairments respectively; the propagation channel state information is represented by $\mathbf{B}_i \in \mathbf{C}^{K \times N_t}$. For DCT transceivers, $\mathbf{T}_{B,i}$

and $\mathbf{R}_{U,i}$ can be represented by $\mathbf{T}(f_i)$ and $\mathbf{R}(f_i)$ as described in 4.2.1., where f_i is the frequency of i th subcarrier. The impairment model is given in Figure 4.3 below:

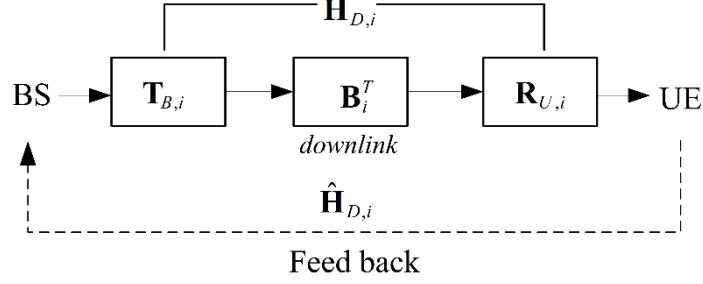


Figure 4.3 MIMO precoding impairments model

The downlink channel estimates $\hat{\mathbf{H}}_{D,i}$ are fed back to the BS after channel estimation and the estimates contain not only propagation state information, but also the impairments of the BS transmitter and the UE receiver. Note that the objective of this chapter is to calibrate the BS's hardware impairments. Therefore, only the BS's hardware impairments are considered in our system model, so the terminals are assumed to be ideal and they can be treated as perfect transceivers. Hence $\mathbf{R}_{U,i} = \mathbf{I}$ and (4.6) can be rewritten as:

$$\mathbf{Y}_i = \mathbf{T}_{B,i} \mathbf{B}_i \hat{\mathbf{H}}_{D,i}^H (\hat{\mathbf{H}}_{D,i} \hat{\mathbf{H}}_{D,i}^H)^{-1} \mathbf{D}_i + \mathbf{N}_i \quad (4.7)$$

C. Massive MIMO systems impairments model for precoding

For massive MIMO systems, it is not practical to send pilots from the base station because the orthogonality of pilots requires many OFDM resource elements which will reduce the throughput. Hence the downlink channel estimates of massive MIMO systems can be obtained via uplink training sequence according to the reciprocity of the propagation channel. The reciprocity and impairments model are shown in Figure 4.4 below.

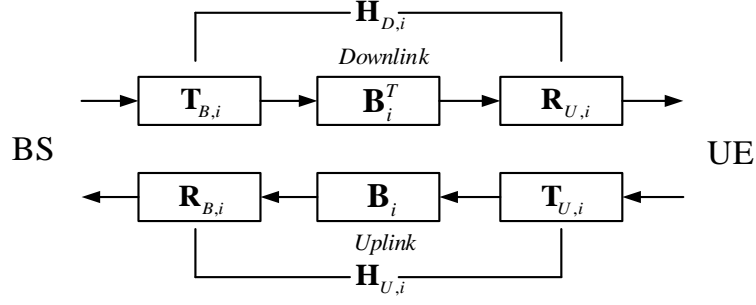


Figure 4.4 Massive MIMO reciprocity and impairments model

Under the condition of perfect UE transceivers, $\mathbf{R}_{U,i} = \mathbf{I}$ and $\mathbf{T}_{U,i} = \mathbf{I}$. Utilizing some simple algebra, (4.6) can be rewritten as

$$\begin{aligned} \mathbf{Y}_i &= \mathbf{H}_{D,i} \mathbf{H}_{U,i}^H (\mathbf{H}_{U,i} \mathbf{H}_{U,i}^H)^{-1} \mathbf{D}_i + \mathbf{N}_i \\ &= \mathbf{B}_i \mathbf{T}_{i,B} (\mathbf{B}_i \mathbf{R}_{i,B})^H (\mathbf{B}_i \mathbf{R}_{i,B} (\mathbf{B}_i \mathbf{R}_{i,B})^H)^{-1} \mathbf{D}_i + \mathbf{N}_i \end{aligned} \quad (4.8)$$

As shown in (4.8), it can be easily seen that the symbols received by the terminals are distorted by $\mathbf{T}_{i,B}$ and $\mathbf{R}_{i,B}$. The system performance is degraded by the hardware impairments.

For one channel of a MIMO or massive MIMO system, the impairments may result in I/Q imbalance, TO, and CFO. From the whole system's perspective, additional impairments can be brought to the users because the phases of the hardware chains are not aligned, and the amplitudes are not at the same level. Therefore, calibration of hardware impairments for MIMO systems is essential for the implementation.

4.3 Impairments calibration for MIMO systems

Despite that there exists many calibration and compensation algorithms for hardware impairments, they mainly focus on the specific impairments such as TO/CFO. For traditional beamforming, a very important aspect is the phase and amplitude alignment between each two of the hardware chains. Thus, for the precoding of a general MIMO system, as long as the impairments remain stable within the coherent time, no calibration is needed because the phase and amplitude difference of the

hardware chains are included in the precoding matrix and can be mitigated automatically by precoding. However, for a massive MIMO system, due to the non-reciprocity between the downlink and uplink, a reciprocity calibration is necessary. This section studies the hardware impairments calibration for MIMO systems.

4.3.1 Transmit beamforming calibration

Beamforming calibration is a very old topic that has been studied for decades and different calibration methods have been proposed. In [77], the mutual coupling effect between the antenna elements is utilized for self-calibration. According to [78], the phase and amplitude difference of the antenna array can be estimated and calibrated by a reference transmit antenna. A very common way to perform beamforming calibration is via sending the orthogonal training pilots to a receiver and the receiver feeds the impairments measurements back to the transmitter [79][80].

This work uses the existing method in [79] for concept proving and the analysis of MIMO hardware impairments features. The calibration scheme is as shown in Figure 4.5 below:

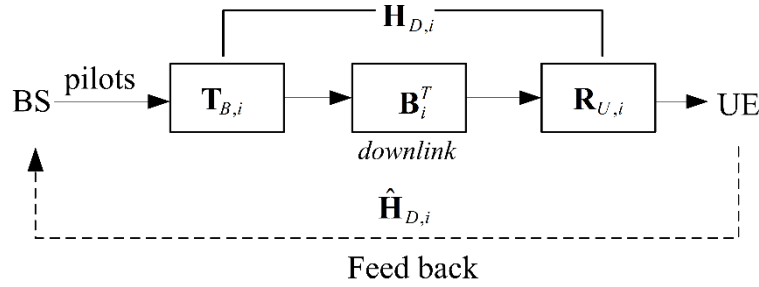


Figure 4.5 Transmit beamforming calibration

This diagram is very similar to Figure 4.3 hence beamforming and precoding can be combined in some way. When beamforming is applied to a wideband signal, the phase and amplitude non-alignment introduced by hardware impairments are going to affect the beamforming result. Generally speaking, beamforming only cares about the

non-alignment at the centre frequency which is not appropriate for a wideband OFDM signal.

Therefore, the calibration should consider a frequency selective impairments and frequency selective channel. Therefore, the pilots and the impairments estimates (or channel estimates) are full band. The signal model of calibration is given as:

$$\mathbf{Y}_i = \mathbf{H}_{D,i} \mathbf{P}_i + \mathbf{N}_i, i = 1, 2, \dots, M, \quad (4.9)$$

where \mathbf{P}_i is the orthogonal pilots at the i th subcarrier. The channel estimates $\hat{\mathbf{H}}_{D,i}$ include both the hardware impairments and the channel model so that beamforming is more appropriate for a slow-fading scenario. Based on the channel estimates, the calibrated signal can be modelled as:

$$y(t) = \mathbf{H}_D \hat{\mathbf{H}}_D^{-1} \mathbf{w}x(t) + n. \quad (4.10)$$

4.3.2 Reciprocity calibration for massive MIMO systems

Many articles have discussed reciprocity calibration of massive MIMO systems. A relative calibration method was first proposed and implemented in [56]. The hardware impairments are generally equivalent to the frequency response of the hardware. R. Rogalin et al. in [55] studied an over-the-air calibration protocol which can scale well with the network size for distributed massive MIMO. A least square (LS) problem was generated for reciprocity calibration which took the noise effect into account. In [69], a hardware-based calibration setup, which can support online calibration, was studied. However, additional hardware-based circuitries are needed which will increase the cost of BS. All the above methods are all based on the "relatively accurate estimation", in which one of the antennas is used as a reference for calibration. Although this will not result in beam pattern distortion as mentioned in [56], it will degrade the performance (such as EVM) of the base station. A calibration

method based on parameter estimation was carried out in [70] which can be performed without any aid of users or external devices. For each pair of transmit (TX) and receive (RX) hardware chains, a parameter was estimated as a function of analog transmit and receive frontends. However, the calibration accuracy is limited because the impairments' parameters of the TX and RX are not estimated separately.

In this work, a closed-loop reciprocity calibration method is proposed which can increase the system EVM performance for massive MIMO systems. Zero-forcing (ZF) is used as the precoding method. Our major contribution is a reciprocity calibration method with higher calibration accuracy compared to the existing methods. Meanwhile, it can support online calibration without any additional hardware circuit in the BS. The calibration method includes two parts. In the first part, offline-over-cable calibration is performed. The calibration algorithm in this step is the same as the traditional relative calibration but the reference TX/RX hardware chain is specified to be of high accuracy. In the second part, a test device (TD) based over-the-air calibration is introduced to support online calibration and also improve the calibration accuracy compared to the existing calibration methods. Furthermore, with our proposed calibration, the BS does not need any additional hardware circuit which can reduce the cost of the BS compared to that in [69] and [70]. Some other references about over-the-air calibration for the antenna system can be found in [71]-[73]. Based on the analysis of hardware impairments in Section 4.2, the signal model and system structure of the proposed reciprocity calibration is first studied.

A. Optimized relative calibration

In relative calibration, one of the base station antennas is used as a reference antenna. The pilots designed for impairments measurement are transmitted between the reference antenna and another antenna. The impairments measurement is performed by channel estimation utilizing the known pilots sent out and the received signal by

the BS antennas. The channel estimation results from the reference antenna and the other antennas can be written as:

$$\hat{\mathbf{H}}_{i,1 \rightarrow j} = t_{B,i,1} \mathbf{B}_{i,1 \rightarrow j} \mathbf{R}_{B,i,j} \quad (4.11)$$

where BS antenna 1 is set to be the reference antenna and the other antenna j is between 2 and N_t . Meanwhile, the reverse procedure is carried out and the channel estimates can be given as:

$$\hat{\mathbf{H}}_{i,j \rightarrow 1} = \mathbf{T}_{B,i,j} \mathbf{B}_{i,j \rightarrow 1} r_{B,i,1} \quad (4.12)$$

In particular, the calibration procedure is as follows:

- 1) Sequentially transmit pilots to and from each base station antenna and reference antenna.
- 2) Perform channel estimation for the received signal of each transmission.
- 3) Calculate the calibration coefficients.

After reciprocity calibration, the calibration coefficients are calculated:

$$\begin{aligned} \mathbf{C}_{rel,i} &= \hat{\mathbf{H}}_{i,j \rightarrow 1} (\hat{\mathbf{H}}_{i,1 \rightarrow j})^{-1} \\ &= \frac{r_{B,i,1}}{t_{B,i,1}} \mathbf{T}_{B,i} (\mathbf{R}_{B,i})^{-1} \end{aligned} \quad (4.13)$$

Using pre-precoding calibration in [74], the UL channel estimation results used for precoding will be

$$\hat{\mathbf{H}}_{U,i} = \mathbf{C}_{rel,i} \mathbf{B}_i \mathbf{R}_{B,i} \quad (4.14)$$

By combining (4.8) and (4.14), the signal received by the UEs can be rewritten as:

$$\begin{aligned}
\mathbf{Y}_i &= \mathbf{B}_i \mathbf{T}_{B,i} (\hat{\mathbf{H}}_{U,i})^H (\hat{\mathbf{H}}_{U,i} (\hat{\mathbf{H}}_{U,i})^H)^{-1} \mathbf{D}_i + \mathbf{N}_i \\
&= \frac{t_{B,i,1}}{r_{B,i,1}} \mathbf{D}_i + \mathbf{N}_i
\end{aligned} \tag{4.15}$$

As shown in (4.15), the quality of the received signal is affected by the impairments of the reference antenna.

Based on relative calibration, we propose an optimized relative calibration which can improve the calibration accuracy. Simply, some high accuracy components, as well as transceiver structure, are used in the reference antenna chain. Particularly, the LO, mixer, ADC and filter in this RF chain have high accuracy and the transceiver can be replaced with an intermediate frequency (IF) transceiver. In this case, the impairments of the reference antenna are very small and the calculated reciprocity calibration coefficients are more accurate compared to the traditional relative calibration coefficients. The UEs received signal will be:

$$\mathbf{Y}_i = \mathbf{D}_i + \frac{a_{ref,T} - a_{ref,R}}{1 + a_{ref,R}} \mathbf{D}_i + \mathbf{N}_i \tag{4.16}$$

where $a_{ref,T}$ and $a_{ref,R}$ stand for the transmitter and receiver impairments of the reference antenna respectively.

B. TD based closed-loop calibration

Although the BS system is calibrated before release, the accuracy of electronic components naturally drifts over time and it is necessary to calibrate the BS hardware at regular intervals to maintain the intended level of accuracy. By taking advantage of less expensive maintenance of the BS, it is very important that the system can be calibrated online rather than shut down the BS and calibrate it offline.

The TD based calibration that we propose is shown in Figure 4.6 which utilizes both UL and DL channel estimation as a closed-loop method. This calibration procedure

includes four steps. Firstly, the UL channel estimation is performed by sending pilots from the TD. Secondly, the BS gives a response to the TD and sends pilots back to the TD. Thirdly, the TD transmits the DL channel estimation results \mathbf{H}_{DL} back to the BS. Finally, the calibration coefficient is calculated according to the DL and UL channel estimation results. The calibration is done by multiplying the calibration coefficient with the UL channel estimation results between the BS and UEs.

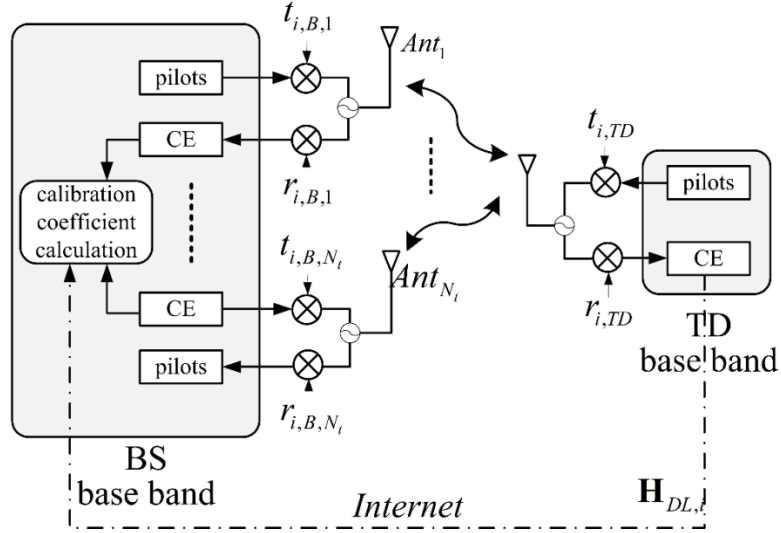


Figure 4.6 TD based closed-loop reciprocity calibration diagram

We understand that the BS may not transmit pilots to the UEs during normal working mode in this scenario. Assume that in this case, the BS can switch to a so-called calibration mode that can respond to the TD pilots and send pilots back to TD which makes online calibration possible. As the BS only sends pilots to the TD in calibration mode, the time-frequency resource will not be a problem to orthogonalize the pilots of each antenna element. To feed the DL channel estimation results back to BS, the BS can be typically connected to the wired Internet [55], [75]. Since the wired Internet is used for uploading DL CSI and the BS only responds to the TD uplink pilots for sending calibration pilots via downlink, the Internet latency between the BS and TD will not affect the calibration results.

After DL and UL channel estimation, the calibration coefficient can be calculated as

$$\begin{aligned}\mathbf{C}_{TD,i} &= \mathbf{H}_{DL,i}(\mathbf{H}_{UL,i})^{-1} \\ &= \frac{r_{TD,i}}{t_{TD,i}} \mathbf{T}_{B,i}(\mathbf{R}_{B,i})^{-1}\end{aligned}\quad (4.17)$$

where $\mathbf{H}_{DL,i}$ and $\mathbf{H}_{UL,i}$ are the DL and UL channel matrix between the BS and TD at subcarrier i respectively; $t_{TD,i}$ and $r_{TD,i}$ are the transmitter and receiver impairments of TD which can be very small if the TD is built with accurate components and sufficiently calibrated in advance. Hence, the impairments of TD are assumed to be approximate to 0 with remaining calibration error a_T and a_R respectively. Using pre-coding calibration in [74], the UL channel estimation results used for precoding will be

$$\hat{\mathbf{H}}_{U,i} = \mathbf{C}_{TD,i} \mathbf{B}_i \mathbf{R}_{B,i} \quad (4.18)$$

Combining (4.8) and (4.18), the signal received by the UEs can be rewritten as:

$$\begin{aligned}\mathbf{Y}_i &= \mathbf{B}_i \mathbf{T}_{B,i} (\hat{\mathbf{H}}_{U,i})^H (\hat{\mathbf{H}}_{U,i} (\hat{\mathbf{H}}_{U,i})^H)^{-1} \mathbf{D}_i + \mathbf{N}_i \\ &= \frac{t_{TD,i}}{r_{TD,i}} \mathbf{D}_i + \mathbf{N}_i \\ &= \mathbf{D}_i + \frac{a_T - a_R}{1 + a_R} \mathbf{D}_i + \mathbf{N}_i\end{aligned}\quad (4.19)$$

which means that the impairments caused by the transceivers of the BS are calibrated with the residual error of $(a_T - a_R)/(1 + a_R)$ remains. To maintain the accuracy of calibration as much as possible, the TD is designed to exhibit low hardware impairments and the moving average algorithm can be used for the channel estimates, which can help to reduce the impact of noise. Moreover, some appropriately accurate channel estimation method can be introduced as well [76].

With the proposed reciprocity calibration method, the non-reciprocity caused by hardware impairments of the BS transceiver can be calibrated. In next section, a

MIMO experimental system is built up for measurement of the hardware impairments as well as the implementation of the proposed calibration method.

4.4 Simulation and experiment results

This section gives the results of simulation and lab experiment for beamforming calibration and reciprocity calibration for MIMO systems. The simulation is performed via MATLAB on PC. A test bed is built up in the lab and the calibration methods are implemented on it. The experiment results captured in the experiments are analysed in MATLAB and the analysis shows that the proposed reciprocity method can improve the MIMO system capacity.

4.4.1 Computer simulation

The computer simulation is accomplished by MATLAB. The signal model of a single antenna is based on LTE downlink standard. As the system model of beamforming and precoding are different, the simulation is separated into three parts, beamforming, MIMO precoding and massive MIMO precoding. They utilize the same OFDM signal parameter which is shown in Table 4-1 below:

Table 4-1 OFDM parameters

Parameter	Value
Bandwidth	20MHz
Carrier frequency	2.6 GHz
f_s : sampling frequency	30.72MHz
Number of symbols	100
Number of subcarriers	1200
Subcarrier spacing	15KHz
FFT size	2048
Length of CP	512
Length of guard band	425

Table 4-1 lists the detail of the OFDM signal that is used in the simulation.

A. Beamforming calibration

The simulation of beamforming uses only 8 antennas as a transmit antenna array. The antenna array is set to be a uniform linear array with a spacing of $\lambda/2$ which is 7.5 cm. To estimate the phase and amplitude non-alignments of the MIMO system in the whole bandwidth, some pilots are inserted in the OFDM signal. The pilot structure is designed as shown in Figure 4.7 below:

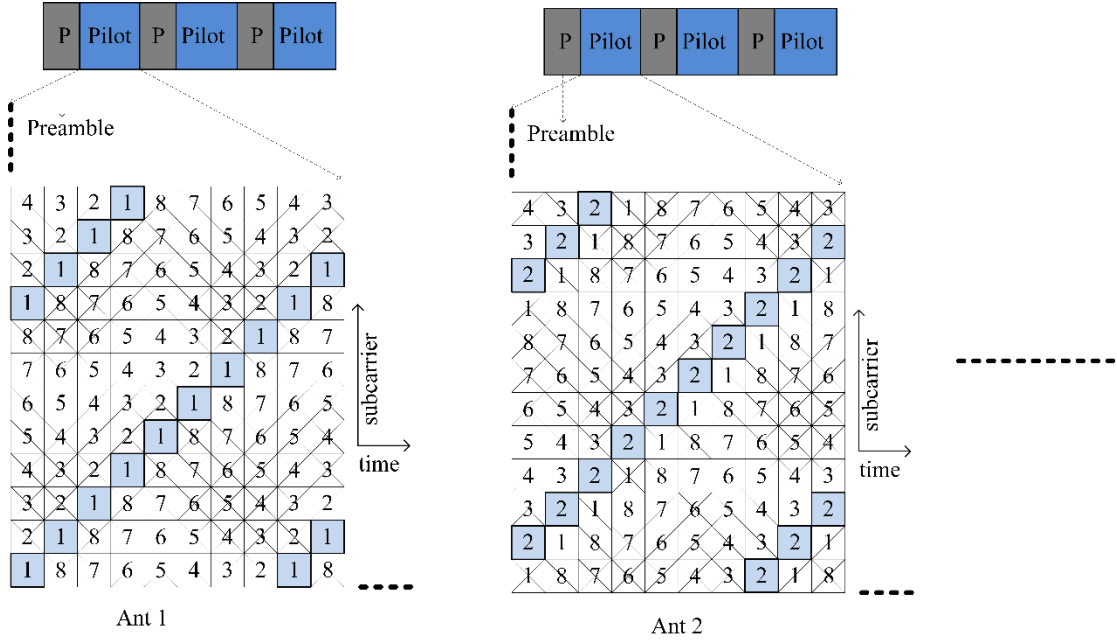


Figure 4.7 OFDM frame and pilot structure

The blue grids represent the pilot data while the white grids denote empty grids which are the positions of other antennas. An OFDM frame consists of a preamble and pilot data. The pilot data includes 100 symbols with 1200 subcarriers of each symbol. The pilot of each BS channel is designed by using the comb-type in [10] and is orthogonal in the frequency domain. In each symbol, one subcarrier is only used by one antenna. The pilot of the next symbols is shifted to the next subcarrier so that each antenna uses the full bandwidth through several symbols. As in this scenario, the UE is slow moving so the coherent time is long enough for channel estimation.

As discussed in Section 4.2, the transmitter impairments of a single channel can be modelled as a lowpass filter $T(f)$. The impairment filter is designed as i.i.d. complex variables with uniformly distributed phase over $[-\phi, \phi]$ and uniformly distributed normalized magnitude in $[0-\sigma, 0+\sigma]$ respectively [55] where $\sigma = 1.5\text{dB}$ and $\phi = 90$ degree. Figure 4.8 and Figure 4.9 show the impaired transmit beamforming beam pattern and the calibrated beamforming beam pattern in both the Cartesian coordinate and Polar coordinate. Two different test cases are applied to the simulation with the corresponding amplitude and phase parameter shown in Table 4-2. In the simulation, antenna 1 is used as a reference so the phase and amplitude of the other antennas given are the relative phases and amplitudes.

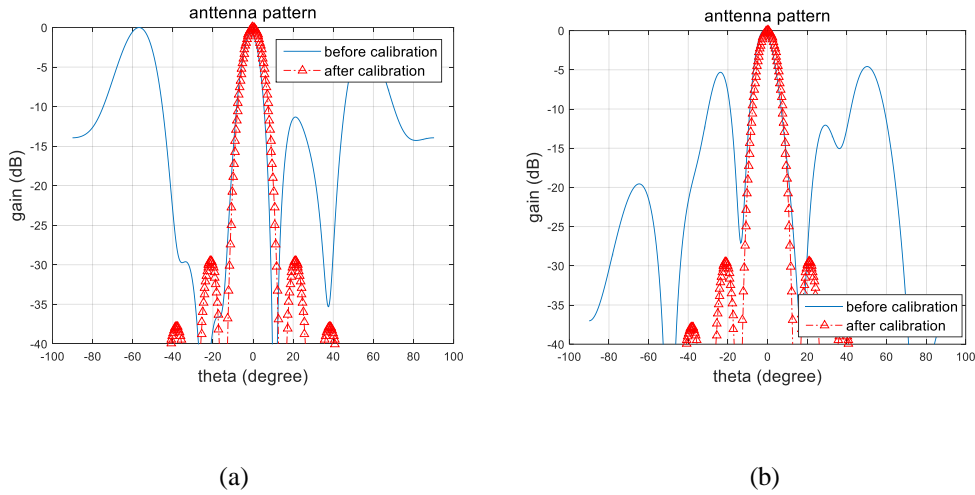


Figure 4.8 Antenna pattern comparison after beamforming in Cartesian coordinate: (a) Test case 1; (b) Test case 2

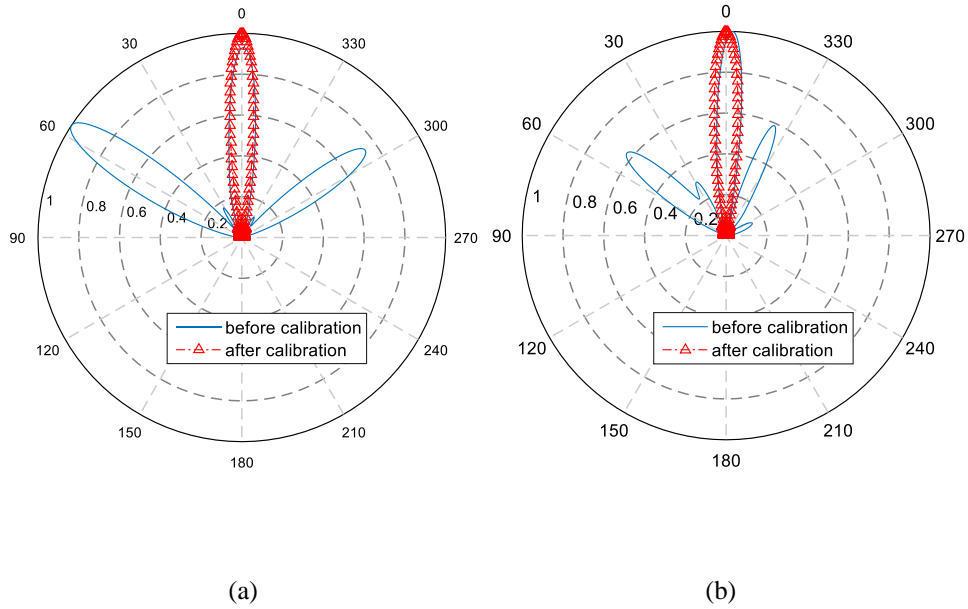


Figure 4.9 Antenna pattern comparison of beamforming in Polar coordinate: (a) Test case 1; (b) Test case 2

Table 4-2 Beamforming phase and amplitude parameters

Test Case	Antenna No.	1	2	3	4	5	6	7	8
1	Amplitude (dB)	0	-0.49	-0.15	-1.79	-0.85	-0.77	-2.36	-0.83
	Phase (degree)	0	115.82	6.30	141.17	-6.90	-0.08	87.86	-6.42
2	Amplitude (dB)	0	-1.57	-1.90	-1.58	-1.61	-0.79	-1.22	-1.30
	Phase (degree)	0	72.65	-58.22	-7.57	71.80	30.54	11.45	-92.26

The simulation results show that the main beam direction of the impaired beam pattern is not 0 and the impaired beamforming generates interferences to other directions. In test case 1, the main beam direction is -56.8. While in test case 2, the main beam direction is -1.6. The side lobe level is an important parameter to characterize the beam pattern. In test case 1, the side lobe is -0.7dB which is very large and not effective for beamforming. In test case 2, the side lobe is -5dB that is still much larger than the ideal value. In both cases, the side lobe is going to interfere the receivers at other directions. Note that under the random impairments, the beam

pattern can be random. Only two cases are chosen to show in the figures and there are many other beam pattern results if there is no calibration. The red curves in the above two figures show the beam pattern after calibration and it is obvious that the calibrated beamforming can achieve a larger gain (30dB) and more accurate directionality.

B. Reciprocity calibration

In this reciprocity calibration, the BS is deployed with 40 antennas which serve 8 single-antenna users. The simulation scheme is shown in Figure 4.10 and Figure 4.11 below.

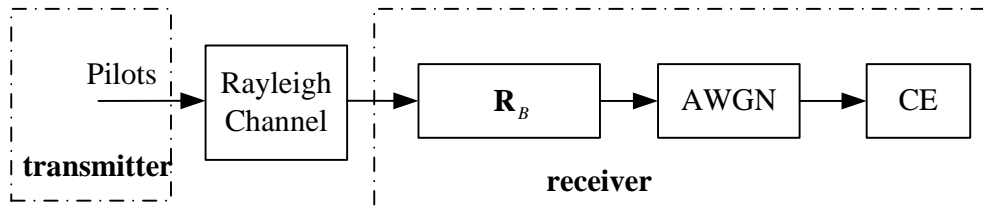


Figure 4.10 Uplink simulation diagram

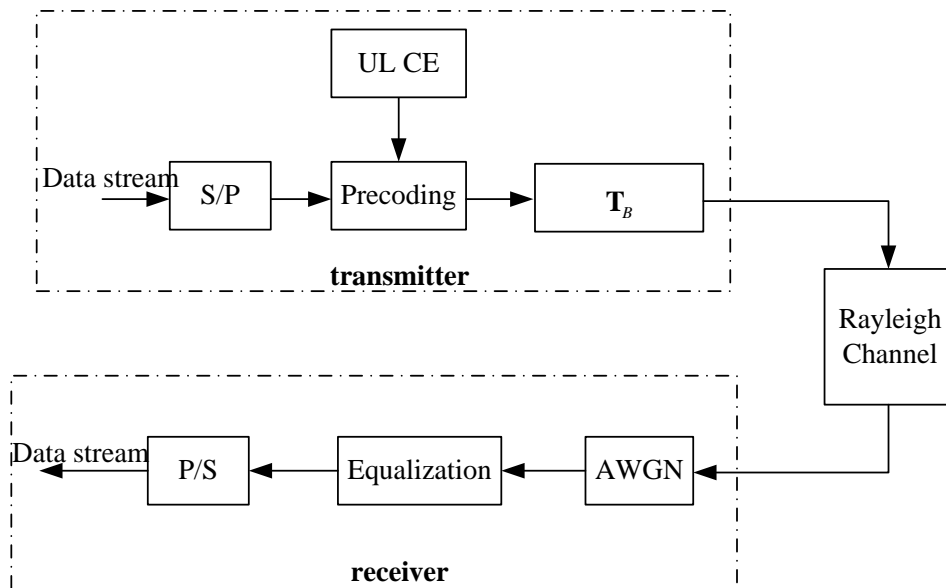


Figure 4.11 Downlink simulation diagram

In the uplink, the UEs only send pilots up to the BS because only downlink performance is considered in this work; hence the uplink frame structure is the same as the reciprocity calibration frame for channel estimation, which is shown in Figure 4.7.

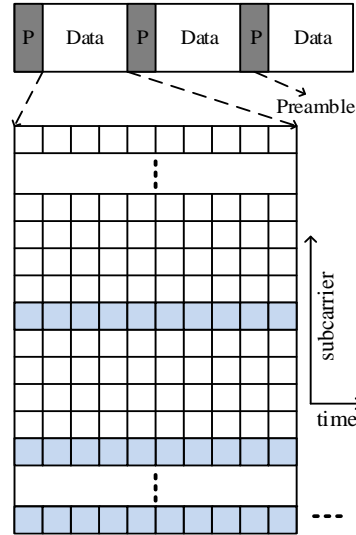
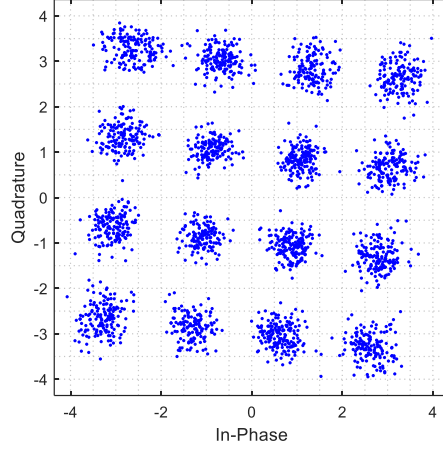
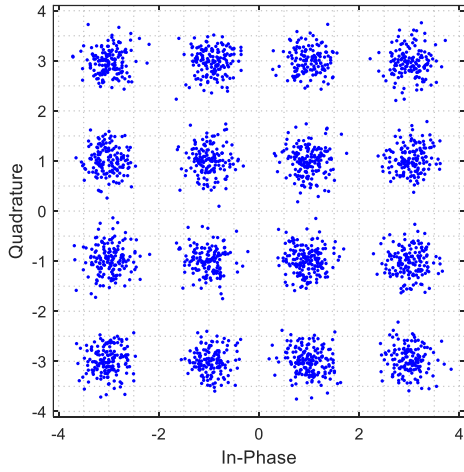


Figure 4.12 Downlink frame structure for data transmission

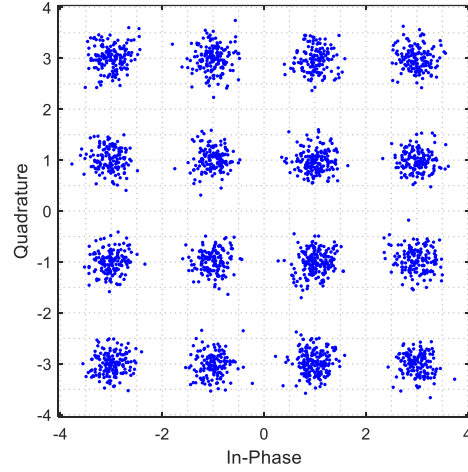
The downlink data frame contains a preamble and some precoded data (the blue grids) as shown in Figure 4.12. Some pilots are inserted into the data symbols to estimate CFO and SFO. These pilots are precoded using the same precoding matrix as the data. Although inserting pilots into data symbols will cause data rate loss, it will not affect the system EVM performance. The preamble is used for time synchronization and the amplitude of preamble is normalized to 1 as well as the precoded data and the pilots. The downlink data stream, modulated in 16-QAM, is firstly converted from serial to parallel (S/P) according to the user number. Rayleigh channel is used in this simulation. Transmitter impairments \mathbf{T}_B and receiver impairments \mathbf{R}_B of the BS are applied to the downlink and uplink respectively. Additional Gaussian noise is added at the receiver side with 20dB SNR. The EVM of the received signal is measured to evaluate the system performance. Figure 4.13 below shows the constellation of three cases, without calibration, with relative calibration, and with TD based calibration.



(a)



(b)



(c)

Figure 4.13 Constellation comparison: (a) without calibration; (b) Relative calibration; (c) TD based calibration

Figure 4.13(a) shows the constellation of the UE received signal without reciprocity calibration, the EVM of the received signal is 17.12% which is 15.33dB. Figure 4.13(b) shows the constellation after relative calibration with the EVM performance of 11.84% which is 18.53dB. The relative calibration can improve the EVM performance of the system and the EVM is improved around 3dB. Figure 4.13(c) gives the constellation of the proposed TD based calibration. The EVM of the received signal with the proposed method is 10.09% which is 19.93dB and the

performance is improved by 1.4dB compared to the relative calibration in [69] and is close to the situation that is without hardware non-reciprocity. This simulation only shows that the proposed method can improve the calibration accuracy of reciprocity calibration. Regarding the performance of on-line calibration, it is not applicable.

Figure 4.14 gives the EVM performance over SNR. Different SNRs are set in this simulation to test the robustness of the proposed method. Two impairments cases that contain different levels of impairments are introduced. The results show that the proposed method can achieve a better EVM under different signal quality than the methods in [69] and [70]. The reference curve refers to the case without hardware impairments. The reason choosing [69] and [70] as benchmarks is because the similarity of system structure between these two methods and the proposed method.

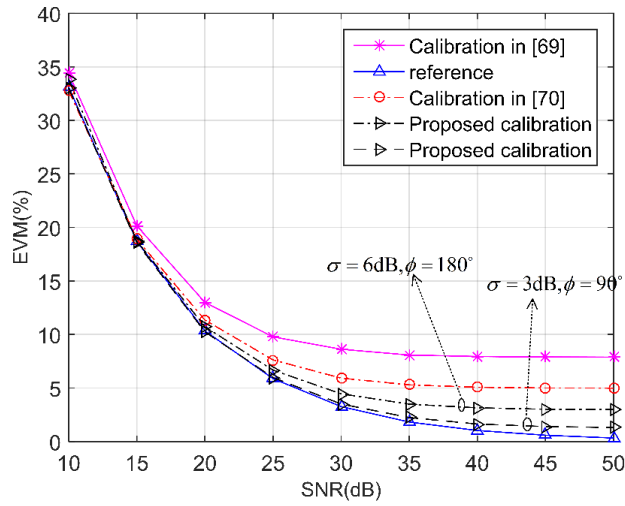


Figure 4.14 EVM performance over SNR

Figure 4.15 shows the EVM performance over different BS antenna numbers. The proposed method is robust to different massive MIMO configurations and the impact of BS hardware impairments cannot be cancelled out by scaling law [83] from reciprocity calibration point of view.

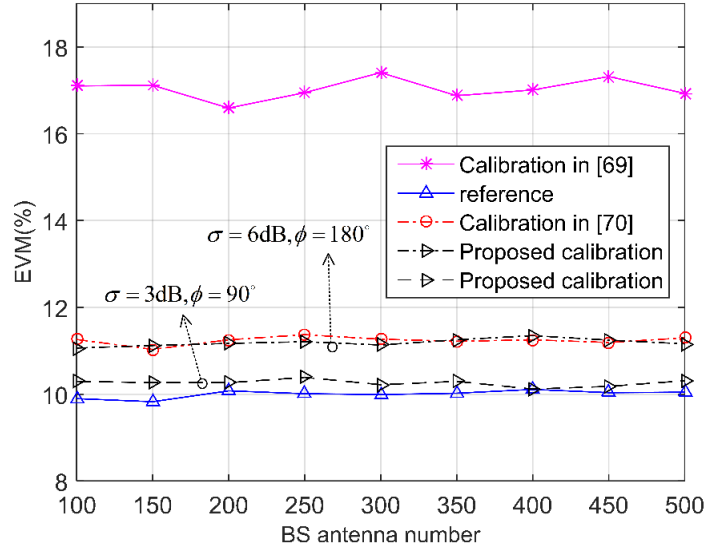


Figure 4.15 EVM performance over different antenna number

4.4.2 Lab experiment

The lab experiment is based on a Linux Server communicating with each National Instrument (NI) Universal Software Radio Peripheral (USRP) (DCT) via Internet and the system is controlled to work in TDD mode. C language is utilized to communicate with USRP driver and Python is used for system control, data capture control and graphic user interface etc. As the USRPs are physically separated, time and frequency synchronization need to be carefully designed. All the RF chains are synchronized by sharing a 10MHz reference signal and a pulse-per-second (PPS) signal from a signal generator between each BS USRP [81]. In this experiment, the RF signal is transmitted over-the-air via antennas hence all the measurements and calibrations are based on over-the-air data.

The objective of this experiment is to prove the effectiveness and robustness of the proposed close-loop reciprocity calibration method. By over-the-air experiment, the experiment result in terms of the EVM of the received signal can prove that the proposed method is working effectively.

A. Beamforming calibration

The beamforming calibration experiment is carried out as shown in Figure 4.16. The BS is made of four USRP X310 as an 8-antenna transceiver system. An 8-antennas uniform linear antenna array is allocated as the BS antennas. Each antenna element is omnidirectional and is separated by a half wavelength. Each USRP can support 2x2 MIMO and is configured to work at 2.6GHz and the bandwidth of USRP is set to be 20MHz. A well calibrated USRP, which has lower DC offset and IQ imbalance compared to the USRPs that are not calibrated, is used as a receiver. The receiver shares the downlink channel estimates, which include the transmitter impairments of the BS, via the Internet which is controlled by Python and C.

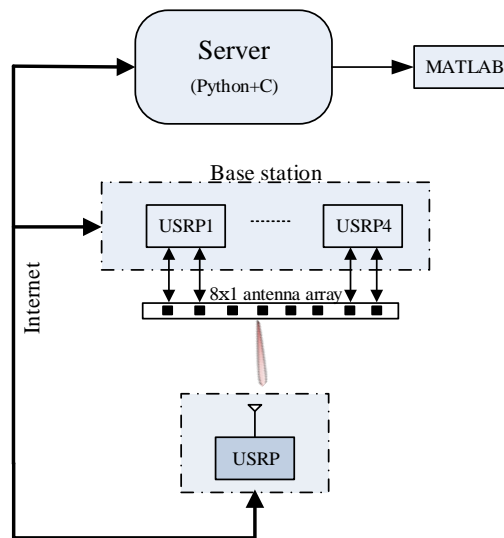
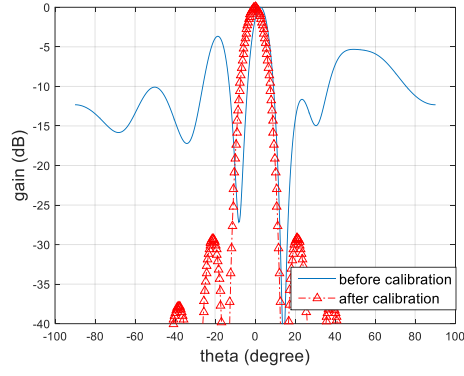
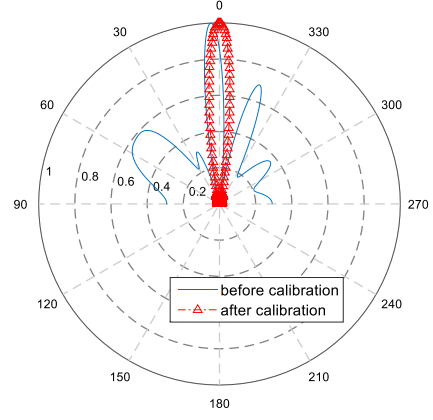


Figure 4.16 Beamforming implementation setup

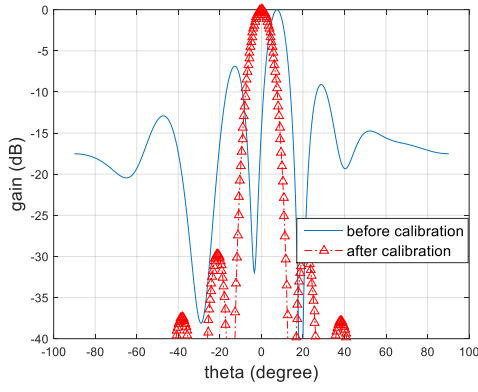
What's more, as the receiver can calculate the frequency-selective phase and amplitude impairments of the transmit hardware chains, the real digital beam pattern can be simulated. Figure 4.17 shows the beam pattern before calibration and after calibration that the calibrated beam pattern points to angle 0 and the side lobe level is very low (-30dB).



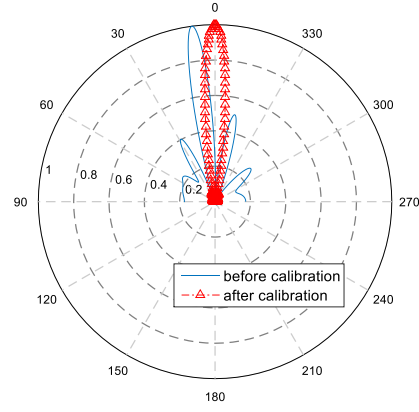
(a)



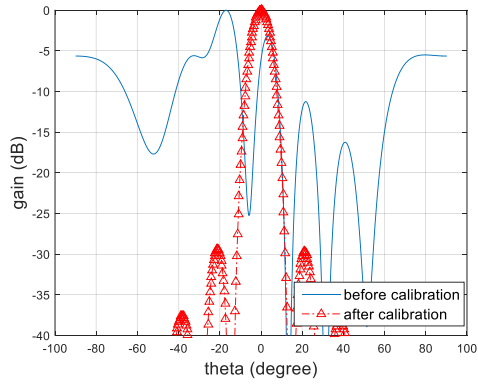
(b)



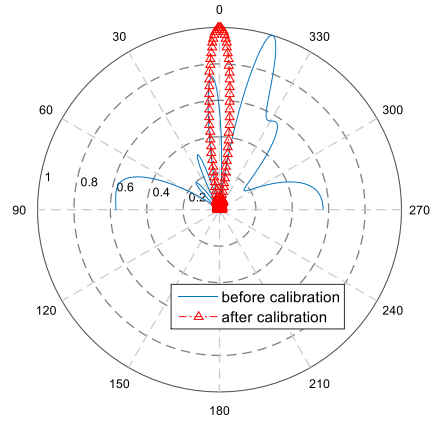
(c)



(d)



(e)



(f)

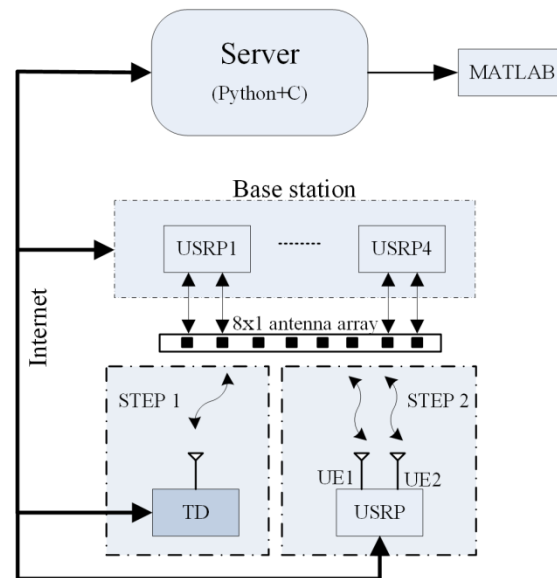
Figure 4.17 Beamforming calibration implementation: (a)(b) first subcarrier; (c)(d) middle subcarrier; (e)(f) last subcarrier

As can be seen in these figures, each different subcarrier has a different impairment which results in the different beamform pattern. For the first subcarrier (Figure 4.17(a) and (b)), the main lobe direction is 2.7 degrees and the side lobe is -3.6dB, while for the middle subcarrier (centre frequency), they are 7.5 degrees and -7dB respectively, and for the last subcarrier, the main lobe direction is -17.5 degrees and the side lobe is -3.2dB. The different behaviours of the different subcarriers are caused by the frequency-selective impairments. However, they can all be calibrated and each subcarrier beam pattern can achieve an improved performance.

The beamforming calibration experiment gives a brief idea of how the transmitter impairments can cause phase and amplitude non-alignment between one antenna and another. This will further result in the beam pattern being distorted. The experiment show the same results as the simulation which in turn proves the method's validity.

B. Reciprocity calibration

The reciprocity calibration uses similar system setup as beamforming, as shown in Figure 4.18(a). The only difference is the receiver in beamforming calibration plays the role as a TD in this part. To ensure high SNR, the TD antenna is placed within 4 meters' distance to the BS in the line of sight, as shown in Figure 4.18(b).



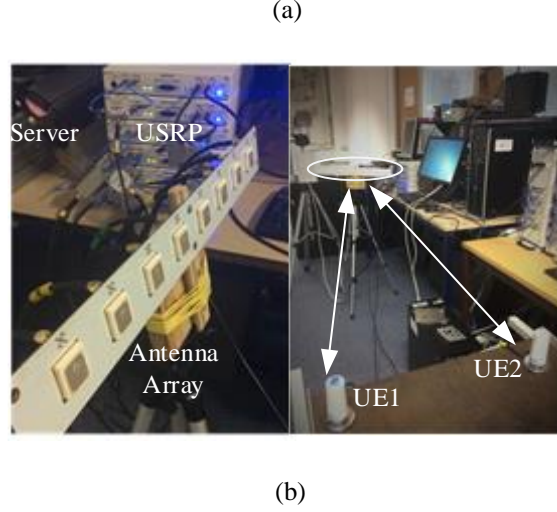
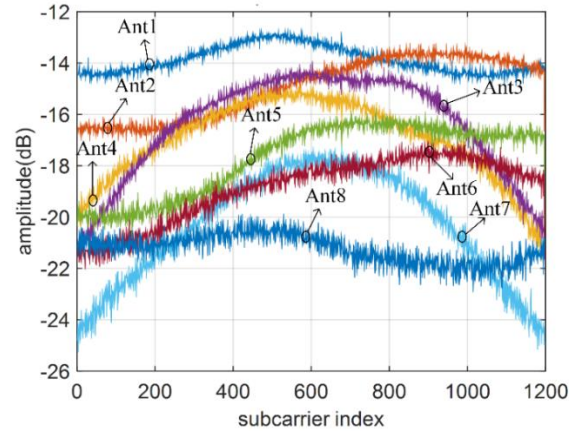


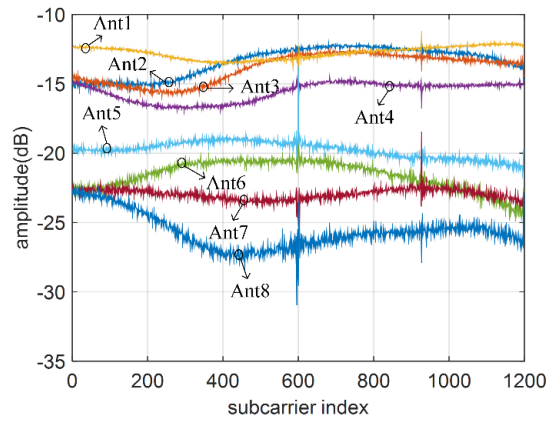
Figure 4.18 Experiment setup: (a) The experiment setup scheme; (b) The experiment implementation setup.

The first step of this experiment is the reciprocity calibration in which the TD is used as a test instrument of the BS. The BS sends pilots to the TD for downlink channel estimation while the TD sends pilots back to the BS for uplink channel estimation. These two processes are undertaken over-the-air. Moving average is introduced to the channel estimation results to reduce the noise effect. After this experiment, the channels between the BS and TD (\mathbf{H}_{DL} and \mathbf{H}_{UL}) are measured. As the BS and TD are placed in a very short distance and line-of-sight, the channel estimation results are treated as the BS transceivers' impairments. As CFO and sampling frequency offset (SFO) are the inherent drawbacks of OFDM-based systems [82], we inserted some block-type pilots in the modulated data for CFO/SFO estimation and cancellation to reduce the impact on the performance of the proposed calibration method.

The channel estimates are shown in Figure 4.19 and Figure 4.20 and they show the non-reciprocity between the BS's transmitters and receivers and the impairments are frequency-dependent. Figure 4.19 shows the amplitude impairments of the BS's transmitters and receivers. Different antennas have different gains and the difference can be up to 10dB.



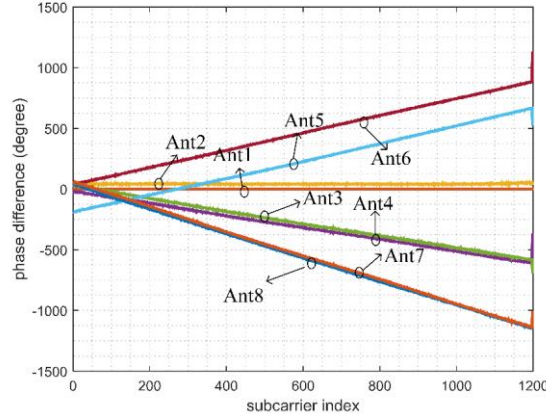
(a)



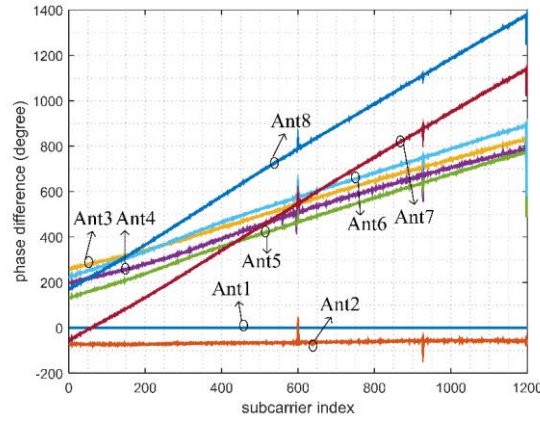
(b)

Figure 4.19 BS hardware impairments' amplitude: (a) transmitters' impairments' amplitude (b) receivers' impairments' amplitude

Figure 4.20 shows the phase impairments of the BS's transmitters and receivers. Antenna 1 is used as the reference in the figure for the sake of comparison. As each USRP supports a 2x2 MIMO, the phase difference between the two antennas of the same device remains constant throughout the subcarriers. The phase difference between two devices changes linearly according to the subcarrier index which is due to different sampling offset.



(a)



(b)

Figure 4.20 Phase impairments: (a) transmitters phase impairments; (b) receivers phase impairments

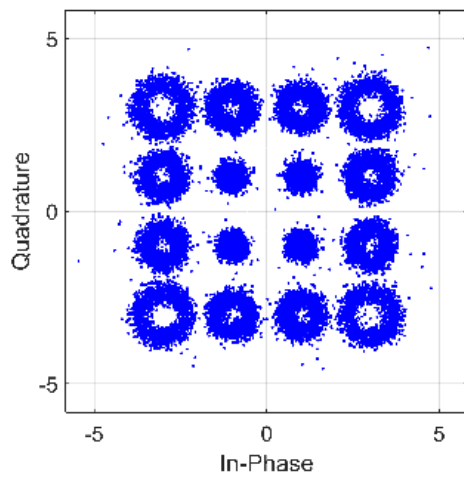
Although both phase and amplitude impairments exist between one antenna and another and the transmitters impairments are different with the receivers, which results in non-reciprocity, the BS can acquire the downlink channel estimates via the Internet which makes the reciprocity calibration closed-loop. After this step, the reciprocity calibration coefficients \mathbf{C}_{TD} can be obtained.

In the second step, a USRP with two different single antennas is used as two UEs as shown in Figure 4.18(b). These two UE antennas are separated by multiple wavelengths to ensure independence. Precoding at the BS is based on uplink channel

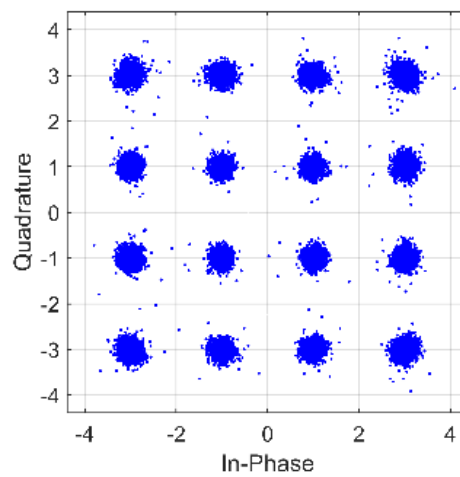
estimation results and the original modulated data is generated offline. The working procedure of this step is shown as below:

- 1) UEs send pilots to the BS via uplink.
- 2) Channel estimation is performed at BS based on the received signal.
- 3) Calibrate the uplink channel estimates using the reciprocity calibration coefficients in the first step.
- 4) Precode the modulated data with calibrated uplink channel estimates.
- 5) BS sends out precoded data.
- 6) UEs capture data and evaluate the system performance.

The system performance is evaluated via MATLAB by calculating the EVM of the captured data of the UEs. The constellation of the captured data with the relative reciprocity calibration in [69] is shown in Figure 4.21(a) while the calibration result based on the proposed TD method is shown in Figure 4.21(b). The EVM of the former is 16.34% (15.73dB), the latter EVM is 5.76% (24.84dB). The proposed method improves the EVM performance by nearly 9dB which proves the validity of this method.



(a)



(b)

Figure 4.21 Constellation of captured data over radio: (a) The constellation using traditional relative calibration; (b) the constellation using proposed TD based calibration.

The reciprocity calibration experiment evaluated the performance of the proposed reciprocity calibration method. The TD is used to measure the BS's transmitters' impairments and receivers' impairments. The measurement results show the non-reciprocity between uplink and downlink. For massive MIMO systems, reciprocity is utilized for precoding to serve multiple users simultaneously. Hence it is important to calibrate the non-reciprocity. The data captured from the UEs received signal shows the effectiveness of this calibration method as the EVM of that is improved compared to the existing methods.

4.5 Summary

Impairments calibration is one of the major challenges for MIMO systems which can be a candidate for 5G network and beyond. This chapter proposed a novel and innovative closed-loop reciprocity calibration method for massive MIMO systems with a better performance compared to the existing methods. The beamforming calibration is undertaken to support the idea that the hardware impairments of a MIMO system will impact the system performance and need to be calibrated. The proposed method consists of two parts, optimized relative calibration and TD based calibration. The former is performed offline over cable using one of the BS antennas as a reference, which is designed to be high quality and low impairments. The latter is accomplished over the air using an accurate specially designed test device. By transmitting the DL CSI from the TD to the BS via the Internet, the calibration coefficient is calculated. In this scenario, the BS can respond to the TD and switch to calibration mode, which makes online calibration possible. The reciprocity calibration is performed by multiplying the calibration coefficient with UL channel estimation results between the BS and UEs.

A lab measurement setup is built for BS beamforming and the implementation of TD based calibration. The beamforming results show that the beam pattern after calibration can achieve a better directionality and lower side lobe level. The TD based calibration is implemented on an 8x2 MIMO system and the EVM of received signal confirms the validity of our proposed method, which can enhance the capacity of massive MIMO systems for 5G network and beyond.

5 Low latency Parallel Turbo decoding and its implementation

The decoding of Turbo code is timing consuming hence parallel decoding methods are needed to reduce the decoding latency and improve both the system's latency and throughput performance. This chapter gives the study and implementation of an improved parallel decoding method for Turbo code. A low latency interleaving address generator is proposed to reduce the decoding latency which in turn will improve the decoding throughput. Via the implementation of the parallel decoder on FPGA, the trade-off between latency and throughput with a limited hardware resource is also been addressed.

5.1 Parallel turbo decoding method

Note that a number of parallel turbo decoders have been proposed previously, and most of them mainly tried to improve the level of parallelism to get a higher throughput and lower latency. In [101], a fully-parallel turbo decoder was implemented using analog decoder, but only short message lengths are supported. According to [102], a parallel turbo decoder algorithm that operates based on stochastic bit sequences was proposed which requires more processing time than Log-BCJR algorithm. A high performance parallel turbo decoder was introduced in [99] with configurable interleaving network which is implemented on very-large-scale integration (VLSI). A fully-parallel turbo decoding algorithm was studied in [103] which can support all LTE and WiMAX standards. However, the computing complexity is too high and is not practical for hardware platform like Field-programmable gate array (FPGA).

For the sake of concept proving for future generation terrestrial systems, it is important that the parallel turbo decoding can be implemented on a platform like FPGA due to the high cost of VLSI or Application Specific Integrated Circuits (ASIC). Besides, FPGA is believed to be a keystone for the Centralized/Cloud Radio

Access Network (C-RAN), which is one of the promising evolution paths for future mobile network architecture [104]. In this chapter, a parallel turbo decoder is implemented on a Testbed which is designed to support multi-Gbps throughput and deployed with several FPGA processors. A reasonable level of parallelism is chosen to meet the demand of throughput, latency and acceptable computing complexity as well. A reverse address generator is proposed to reduce the interleaver complexity and reduce the iteration time at the same time. Modulo operation is an essential part of interleave index generation. A practical method of modulo operation is designed which helps to reduce the complexity in FPGA especially when the parallelism level is high. The contribution of this work is that a feasible solution is provided for parallel turbo decoder implementation on FPGA with latency reduced and throughput improved.

5.1.1 Parallel decoding structure

To reduce the decoding latency of the RSC decoder, the RSC decoder can be designed to perform in a parallel decoding way. A parallel decoder is performing in parallel by separating the whole block into P sub-blocks, where P is the level of parallelism. In this way, the decoding time is reduced because the length of sub-block $K = N / P$ is much smaller than the whole block. Generally speaking, the higher the level of parallelism, the less decoding time is needed. According to the parallel decoding algorithm proposed in [96], as shown in Figure 5.1, the last forward variable $\alpha_{(p-1)K}(S_{(p-1)K})$ and backward variable $\beta_{pK+1}(S_{pK+1})$ from the previous iteration of the neighbour sub-blocks $(p-1)$ th and $(p+1)$ th are utilized as the initial value of computation for this p th sub-block. Here, $\alpha_i(S_i)$ and $\beta_i(S_i)$ represent the forward and backward information at m th trellis stage of the i th decoding bit with state S . α_0 and β_0 denote the forward and backward initial variable values of each sub-block in the first iteration and also represents the first and last sub-blocks in the later iterations respectively.

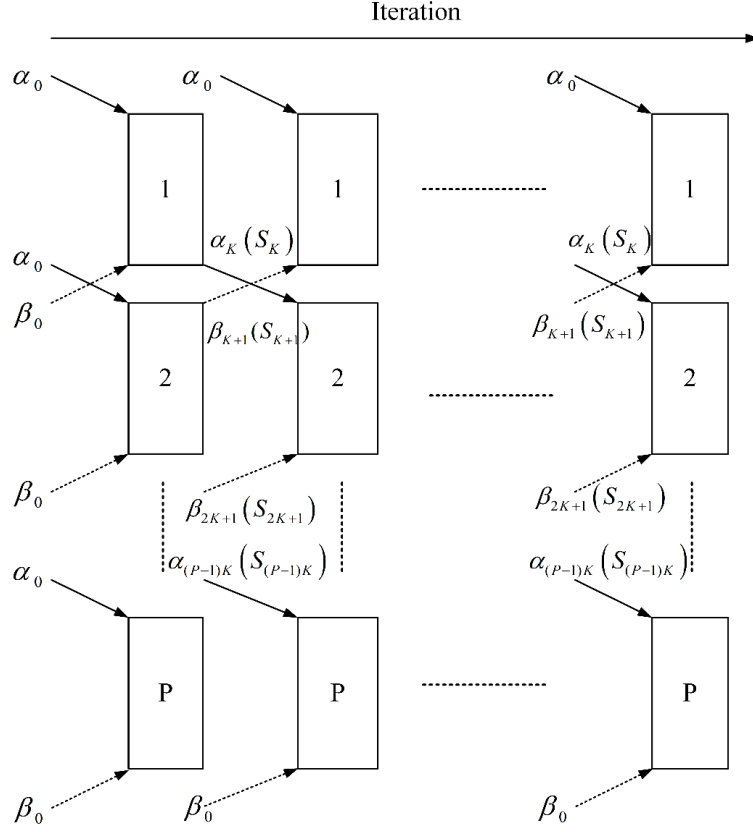


Figure 5.1 Parallel RSC decoding structure

5.1.2 Parallel interleave/deinterleave

All the sub-decoders of each sub-block are performed in parallel and simultaneously so that the parallel decoder can reduce the decoding time to $1/P$ of the sequential decoding time. Note that between each two iterations, the output LLR of the previous iteration will be processed by interleaving/deinterleaving. Hence interleaving and deinterleaving also need to perform in a parallel way. Since the interleaving index mapping is determined by the block size, parallel interleaver needs to be designed to support all block sizes, as well as parallel deinterleaver. A memory access contention may occur during the interleaving of extrinsic information. Therefore, based on some algebraic constructions, contention-free interleavers have been proposed in [97], [98] and references therein. The interleaver used in this chapter is also known as quadratic

polynomial permutation (QPP) interleaver. It is a contention-free interleaver and suitable for parallel Turbo decoding.

Equation (2.32) can be modified as

$$\begin{cases} \pi(i) = A(i) \bmod N \\ A(i) = f_1 i + f_2 i^2 \end{cases}, i = 0, 1, 2, \dots, N-1, \quad (5.1)$$

where f_1 is an odd number and f_2 is even, i is the index number of input data \mathbf{y}_i and $\pi(i)$ is the index number after interleaving. For parallel interleaver, if the parallel level P can divide block size N , then this interleaver is contention-free [99]. For example, if $N=6144$, the decoder can be designed as a parallel decoder with maximum parallel level of 2048 ideally.

To generate the target interleaving address according to (5.1), the compute complexity is quite high if using realtime multiply operation to calculate $A(i)$ because the index i increases progressively till $N-1$. Therefore, an optimized address generator is proposed in [99] which has low complexity. The address generation is accomplished by recursion and the derivation is as follows. According to (5.1),

$$\begin{aligned} A(0) &= 0, A(1) = f_1 + f_2 \\ A(i+1) - A(i) &= f_1 + f_2 + 2i \cdot f_2 \\ A(i+2) - A(i+1) &= f_1 + 3f_2 + 2i \cdot f_2 \end{aligned} \quad (5.2)$$

then

$$A(i+2) - 2A(i+1) + A(i) = 2f_2. \quad (5.3)$$

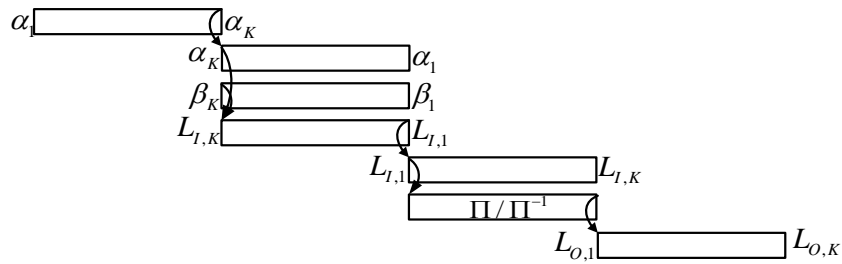
Since $A(0)$ and $A(1)$ are known initial factors, by recursion, the following interleaving index can be generated from (5.3). In this way, no multiplication is needed, which helps reduce complexity dramatically. This address generator cannot

be used in a parallel interleaver directly because all the sub-blocks are processing simultaneously hence a parallel address generator is needed.

The memory of parallel interleaver is divided into P banks corresponding to P sub-blocks. The i th extrinsic information will be stored in the $\pi(i)/K$ th bank at the address of $\pi(i) \bmod K$ after interleaving. In addition, deinterleaving is the inverse operation of interleaving for which the principle of address generation is the same as interleaving.

5.1.3 Proposed reverse address generator

Based on the forward and backward computation structure of turbo codes, the sequence of backward variables $\beta_i(s)$ should be reversed to calculate the extrinsic information. This process adds processing time by at least N clock cycles for sequential decoding or K clock cycles for parallel decoding as shown in Figure 5.2(a). Utilizing the characteristics of interleaving, the sequence of extrinsic information can remain reversed and does not affect interleaving while the processing time can be reduced to 3/4 of the original sequence interleaving (see Figure 5.2 (b)). $L_{I,k}, k = 1, 2, \dots, K$ and $L_{O,k}$ represent the LLR of a sub-block before interleaving/deinterleaving and after interleaving/deinterleaving respectively.



(a)

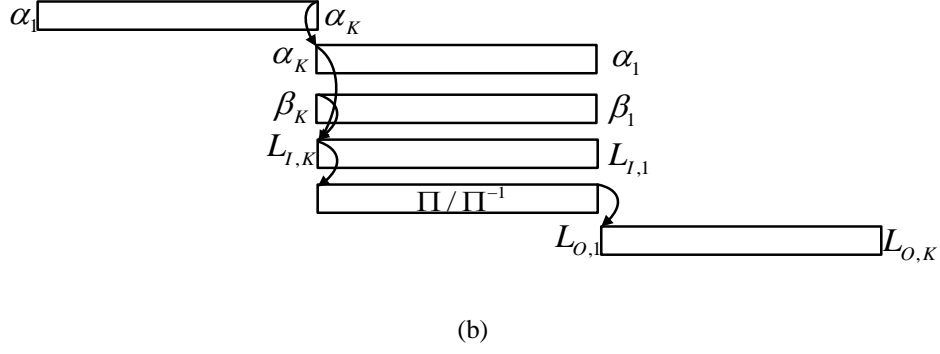


Figure 5.2 Timing diagram of each iteration; (a) original sequence interleaving; (b) reversed sequence interleaving

Since the sequence of interleaver input is reversed, the address generator should be changed accordingly. Therefore, we proposed a reversed address generator for parallel turbo decoding to reduce the computation complexity and processing latency as well.

The address of target memory bank $\pi(i) \bmod K$ can be modified as

$$\begin{aligned}\Lambda_{p,k} &= \pi(pk) \bmod K \\ &= A(pk) \bmod K\end{aligned}\quad (5.4)$$

where

$$A(pk) = f_1 pk + f_2 (pk)^2, k = K-1, K-2, \dots, 0, p = 1, 2, \dots, P$$

Note that the first two addresses of each sub-block that need to be generated are $\Lambda_{p,K-1}$ and $\Lambda_{p,K-2}$. According to (5.4),

$$\begin{aligned}\Lambda_{p,K-1} &= f_1 p(K-1) + f_2 (p(K-1))^2 \bmod K \\ \Lambda_{p,K-2} &= f_1 p(K-2) + f_2 (p(K-2))^2 \bmod K\end{aligned}\quad (5.5)$$

since f_1 , f_2 and p are all integers, (5.5) can be simply modified to

$$\begin{aligned}\Lambda_{p,K-1} &= -f_1 + 2f_2 \bmod K \\ \Lambda_{p,K-2} &= -2f_1 + 4f_2 \bmod K\end{aligned}\quad (5.6)$$

Using similar derivation as (5.2), the following address of each filter bank can be generated by recursion

$$\Lambda_{p,k+2} - 2\Lambda_{p,k+1} + \Lambda_{p,k} = 2f_2 \bmod K \quad (5.7)$$

From (5.7), we can find that the recursion process and initial values have nothing to do with p hence the addresses of all these sub-blocks are the same and only one channel of address generator is necessary for this parallel interleaver:

$$\Lambda_{k+2} - 2\Lambda_{k+1} + \Lambda_k = 2f_2 \bmod K \quad (5.8)$$

The destination bank that the LLR of a sub-block should be mapped into is decided by the value of $\pi(i)/K$. The division operation here is costly therefore recursive computation is needed for this reverse address generator. Let $\pi(i)/K$ be redefined as

$$\begin{aligned} \Gamma_{p,k} &= \pi(pk)/K \\ &= [A(pk) \bmod N]/K \end{aligned} \quad (5.9)$$

The recursion has two dimensions. First, the recursion direction is from $k = K - 1$ to $k = 0$. $\Gamma_{p,K-1}$ and $\Gamma_{p,K-2}$ are the initial values. Second, another recursion is performed from $p = 1$ to $p = P$ where $\Gamma_{1,k}$ and $\Gamma_{2,k}$ are the initial values. To accomplish this two dimensional recursion, $\Gamma_{1,K-1}$, $\Gamma_{1,K-2}$, $\Gamma_{2,K-1}$ and $\Gamma_{2,K-2}$ must be known before the computation.

$$\begin{cases} \Gamma_{p,k+2} - 2\Gamma_{p,k+1} + \Gamma_{p,k} = [2p^2 f_2 \bmod N]/K \\ \Gamma_{p+2,k} - 2\Gamma_{p+1,k} + \Gamma_{p,k} = [2K^2 f_2 \bmod N]/K \end{cases} \quad (5.10)$$

Since the interleaver/deinterleaver is placed after the whole computation of $\alpha_{i,m}(s)$, $\Gamma_{1,K-1}$, $\Gamma_{1,K-2}$, $\Gamma_{2,K-1}$ and $\Gamma_{2,K-2}$ can be calculated via pipeline of the multiplication cell before the address generator, as well as $[2p^2 f_2 \bmod N]/K$ and $[2K^2 f_2 \bmod N]/K$. By this recursion, no realtime multiplication is needed during the address generation.

With the reverse address generator mentioned above, the parallel interleaver/deinterleaver can reduce the processing time compared to the method in [99] even though it may cost a little more computation resource.

5.2 Parallel turbo decoder implementation

Due to its low cost and short development cycle, FPGA is one of best hardware platform choices for a real-time proof of concept system. In this work, the parallel LTE turbo decoder including the proposed interleaving address generator is implemented on Xilinx Virtex VII. In this section, the detail of decoding implementation is introduced. This decoder can support all the block sizes of the LTE standard. The different parallel level P can be configured according to the specific block size. Considering that the higher the parallel level is, the more complex the decoder will be and the more computing resources will be used, P is set to be 64 when the block size N ranges from 2048 to 6144 and $P=8$ when $256 \leq N < 2048$, otherwise $P=1$.

5.2.1 Extraction

The LTE received data before turbo decoding has a certain format, with all the systematic bits and two frames of parity bits included. Therefore, before the calculation of LLR, the extraction of the received data frame is needed. An interleaver is located here as well to generate sys2 to match par2 for sub-decoder 2, as shown in Figure 5.3. A FIFO is placed after the interleaver to synchronize with par2.

Since during the iteration of decoding, all these systematic and parity frames that are going to be reused, sys1, par1 and par2, are stored in block RAMs and the read of RAMs is controlled by the request signals (req_1 and req_2) from the LLR calculation module. And because the extraction will generate a group of parallel input data for LLR calculation, a configurable parameter is used here to make this decoder compatible with different parallelism levels. Moreover, to produce the same number

of block RAMs according to the parallel level, the method of source code generation is utilized. For example, the generate operation was created in Verilog HDL.

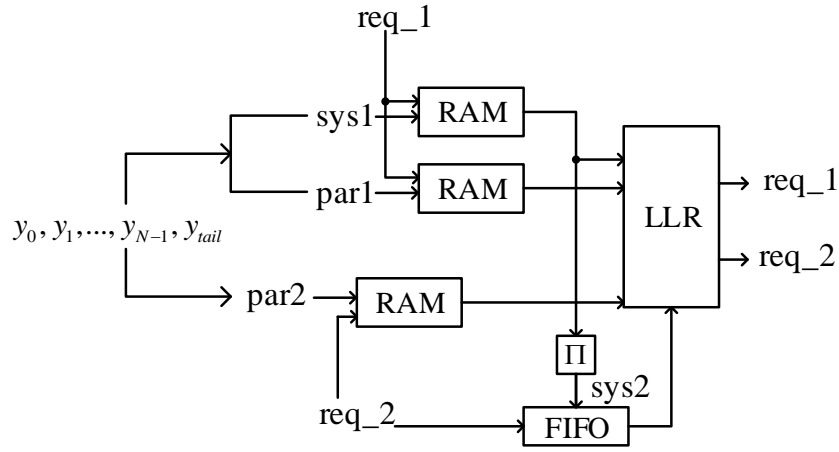


Figure 5.3 Extraction of received data.

5.2.2 Extrinsic information

Extrinsic information calculation includes forward variable α and backward variable β calculation. As shown in Figure 5.4, a block RAM is placed after α module to reverse the sequence as mentioned in Section 5.1. Source code generation is used here as well to produce P groups of α modules, β modules and LLR modules.

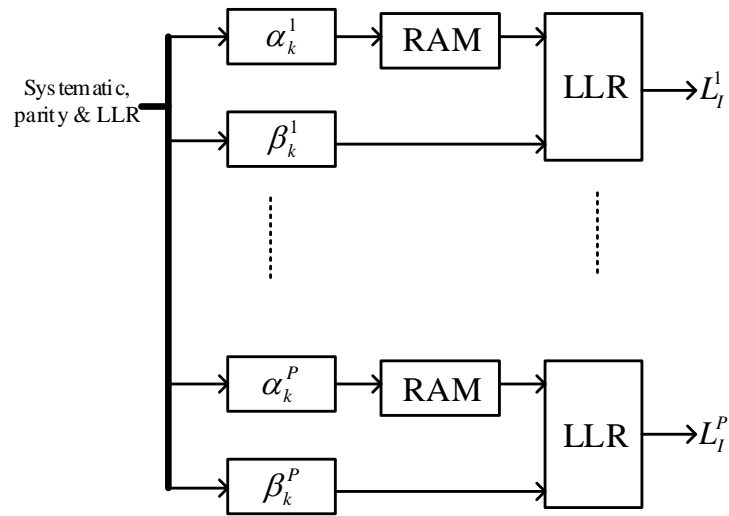


Figure 5.4 Parallel extrinsic information calculation.

Note that in the theoretical calculation of α and β , the initial value should meet the following condition:

$$\begin{aligned}\alpha_i(s) &= -\infty, s \neq 1 \\ \beta_i(s) &= -\infty, s \neq 1\end{aligned}\quad (5.11)$$

However, minus infinity does not exist in the practical fixed-point calculation. A logical comparator is utilized because if $\alpha = -\infty$, then $\alpha + x = -\infty$ where x can be any value except infinity. Hence $-\infty$ in FPGA is replaced by a least signed value. More specifically, in our case, a 16-bit hexadecimal 2's complement value 8000H is used. By comparison, if α equals 8000H, then $\alpha + x$ is still 8000H. The same method is used to deal with β .

5.2.3 Interleaver/Deinterleaver

As mentioned in Section 5.1, the interleaver is contention-free as long as the parallel level P can divide the block size N . Memory contention would not happen in our study because all the block sizes can be divided by P . The interleaver/deinterleaver is a memory dynamic mapping process. The target address and memory bank are generated by the proposed reverse address generator. As shown in Figure 5.5, for the interleaving process, using multiplexer in FPGA, the realtime LLR can be written into the related block RAMs and read from them sequentially after writing has been completed.

As mentioned in Section 5.1, the RAM writes address $\Lambda_{p,k}$ of each sub-block is independent of p so only $\Lambda_{1,k}$ is produced from the address generator. The LLR results of LLR modules are mapped to different RAMs according to $\Gamma_{p,k}$. On the other hand, the writing process is sequential for deinterleaver while read address $\Lambda(i)$ and bank number $\Gamma(i)$ are generated by the address generator.

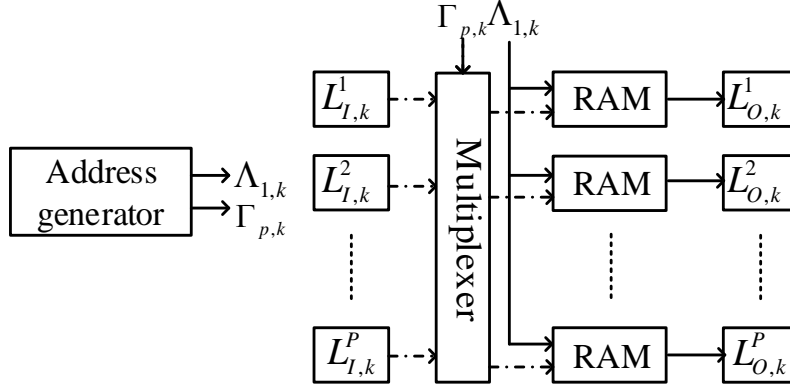


Figure 5.5 Parallel Interleaving.

5.2.4 Modulo operation

Modulo operation $C \% D$ is a costly part of the address generator. The result of modulo is the remainder of a division operation. For Xilinx FPGA, the only existing function for modulo operation is the division intellectual property (IP) core which takes many logic units. Some other faster methods like the bitwise operation also exist but they assume D as a constant or the number of powers of 2 [100]. In our study, D is not a constant or a power of 2. A modulo function based on Verilog HDL should be designed with less computing resource and fast speed as well.

Inspired by the bitwise operation, a modulo function is designed that it uses a shifter and comparator to get the remainder of the division but not the quotient. Let E be the maximum bit width of C , F be the maximum bit width of D . The procedure of the proposed modulo function is shown in Figure 5.6 below.

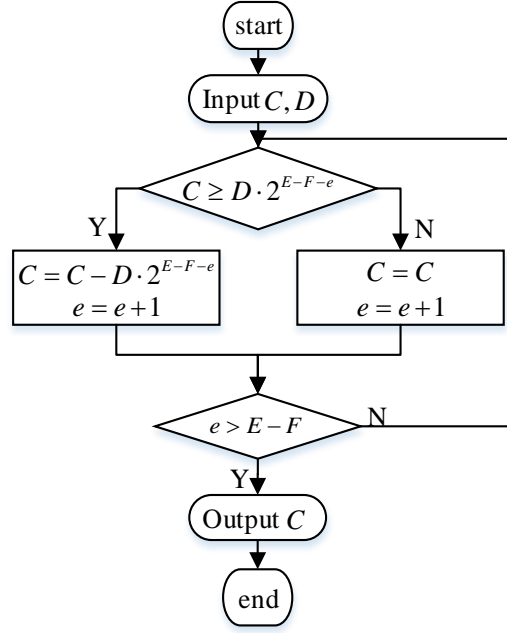


Figure 5.6 Modulo function flowchart.

5.2.5 Double buffering

To maximize the throughput of the turbo decoder, double buffering is utilized in this design. Since the calculation of α , β and LLR is sequential, the previous module is idle when the latter module is working. For instance, as shown in Figure 5.2, β module works after the whole sub-block calculation of α . Obviously, another sub-block of α , denoted by $\tilde{\alpha}$, can keep calculating during that period, as shown in Figure 5.7 below. In this way, the whole decoder can decode two frames simultaneously which can nearly double the throughput. Double buffering is very functional as only double storage space is used, however, the logic and compute resources, e.g. lookup tables (LUTs), Flip-Flop, Multipliers, are reused so the utilization of FPGA resource is more efficient.

communication system, this Testbed is designed to support multi-Gbps decoding throughput. The structure of it can be found in Figure 5.8.

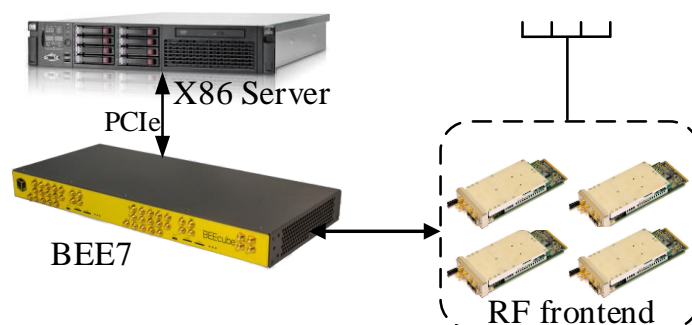
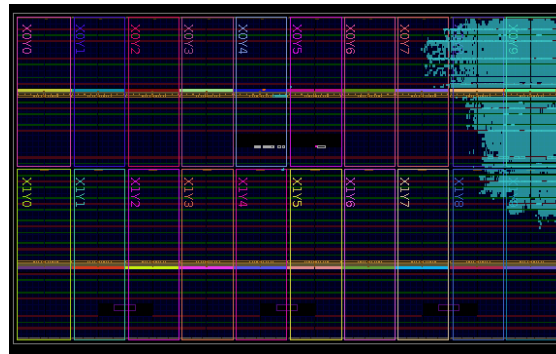


Figure 5.8 Testbed system structure.

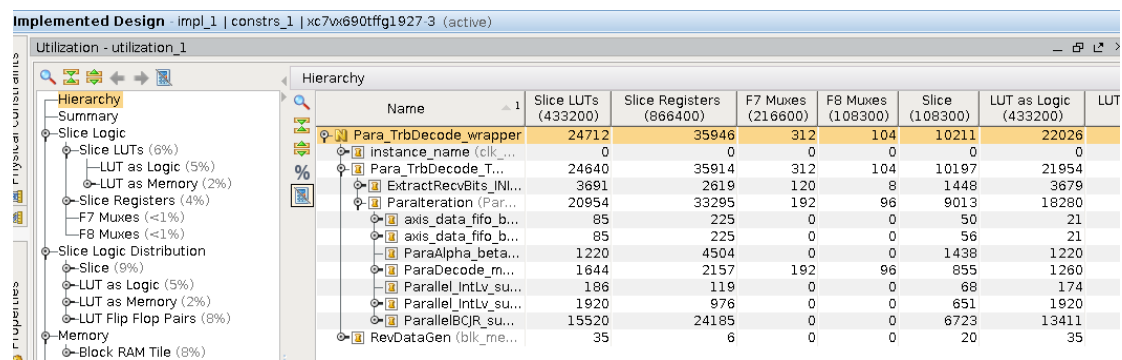
The X86 Server is the control center of this Testbed, which is connected via Peripheral Component Interconnect Express (PCIe) with BEE7. BEE7 is a programmable hardware platform used for algorithm exploration, research, prototyping and so on. Four Xilinx Virtex-7 FPGAs are allocated on this platform. With one FPGA processor, the throughput requirement cannot be met. BEE7 can link with several RF frontends to build a MIMO transceiver. Vivado Design Suit from Xilinx is utilized as the FPGA development tool.

Note that for a particular FPGA device, the hardware resource is fixed. As we know that higher parallelism means lower latency, it also takes more computing resources especially logic resources such as Look Up Tables (LUTs) and Flip-Flops. For one FPGA in BEE7, the parallel level can reach to $P = 64$ with a latency of 23.2 us at 250 MHz clock rate where the iteration number is 8. Although the latency is quite low compared to lower parallelism, the throughput of this system is only 2.12 Gbps, because only one decoder can put on the FPGA, which is not enough. While when $P = 8$, although the latency is larger than $P = 64$, 8 decoders can be put on the FPGA and perform decoding simultaneously and the overall throughput can be higher than $P = 64$.

Figure 5.10 below shows the results, in terms of resource utilization and FPGA area utilization, when the parallel level is 8 after implementation.



(a)

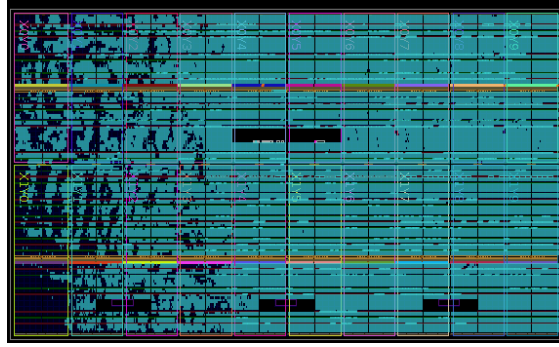


(b)

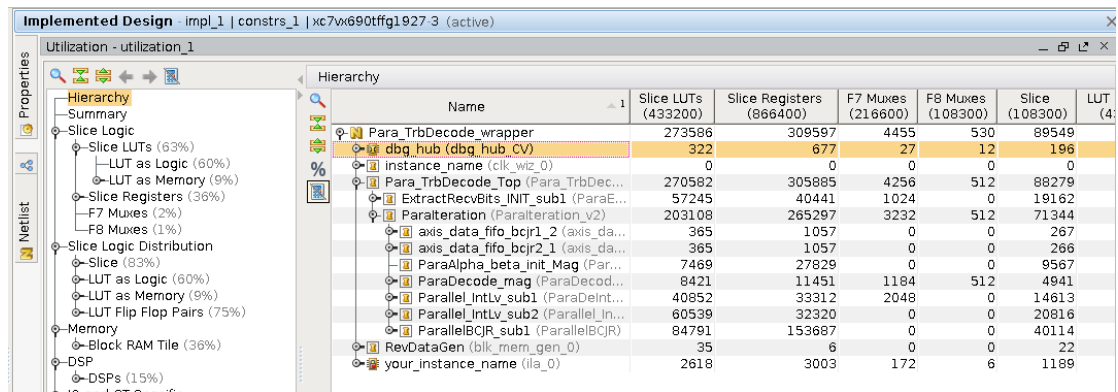
Figure 5.9 FPGA utilization, $P = 8$, (a) Area utilization; (b) Resource utilization.

Figure 5.9(a) gives the area utilization of the FPGA. In the case that $P = 8$, the area utilization of this FPGA is around 10% so more decoders can be put on the FPGA which can improve the system decoding throughput because all these decoders can perform channel decoding simultaneously. The resource utilization shown in Figure 5.9(b) tells that the slice used in logic distribution is 9%.

Figure 5.10 shows the results of the decoder after implementation when the parallel level is 64. Figure 5.10(a) gives the area utilization of the FPGA. The area of this FPGA is almost used up so it is obvious that only one decoder can be put on the FPGA. The resource utilization shown in Figure 5.10(b) tells that the slice used in logic distribution is 83%. When the parallel level is high, each module increases the usage of resource especially the interleaving and deinterleaving module.



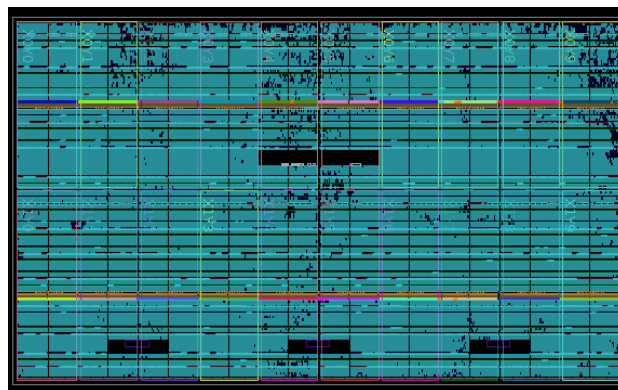
(a)



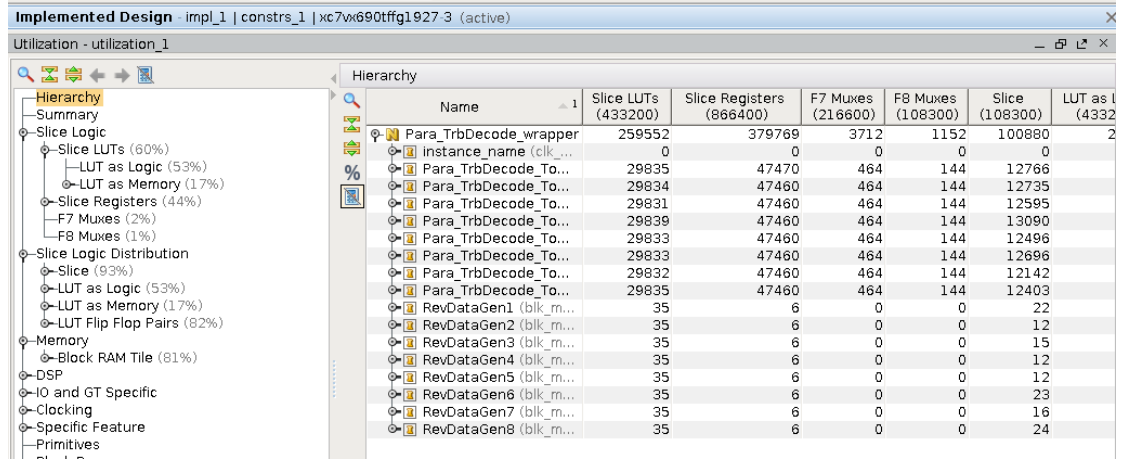
(b)

Figure 5.10 FPGA utilization, $P = 64$, (a) Area utilization; (b) Resource utilization.

Figure 5.11 gives the implementation results that 8 decoders with $P = 8$ are put on the FPGA at the same time. The area of the FPGA is almost filled up and the Slice usage is 93%.



(a)



(b)

Figure 5.11 FPGA utilization, $P = 8$, with 8 decoders, (a) Area utilization; (b) Resource utilization.

The FPGA resource usage when $P = 64$ is slightly less than the case that 8 $P = 8$ decoders work simultaneously. But the decoding throughput and latency are different. The throughput and latency comparison of different parallel levels are listed in Table 5-1.

Table 5-1 Throughput and latency comparison

	This work		[99]	
Iteration	8			
Block size	6144			
P	8	64	8	64
Clock(MHz)	250			
Extraction time (us)	18.4			
Iteration time (us)	38.4	4.8	51.2	6.4
Latency (us)	56.8	23.2	69.6	24.8
Throughput (Gbps)	6.92	2.12	2.82	0.99

The results above are obtained via ModelSim simulation after placement and routing. This simulation can measure how many clock cycles are needed for a whole decoding process. By some simple calculation shown in (5.12) and (5.13), the latency and throughput can be calculated. When the parallel level is 8, which is 8 times lower than 64, but the latency is not 8 times larger. This is because the extraction of the received

data takes a fixed amount of time. Moreover, since it takes much fewer resources when $P=8$, 8 parallel turbo decoders can be put on a single FPGA at the same time, which makes the throughput reach to 6.92Gbps. Even though the throughput is low when $P=64$, its good latency performance can still be used for the case of a strict latency requirement.

The implementation validity is evaluated by Integrated Logic Analyzer (ILA) of Xilinx. A fixed test block is stored in a block RAM. By capturing the output of the turbo decoder, the decoding results can be examined. As shown in Figure 5.12, the original frame before turbo encoding is a square wave, and we can see that the output is a square wave that matches the original frame.

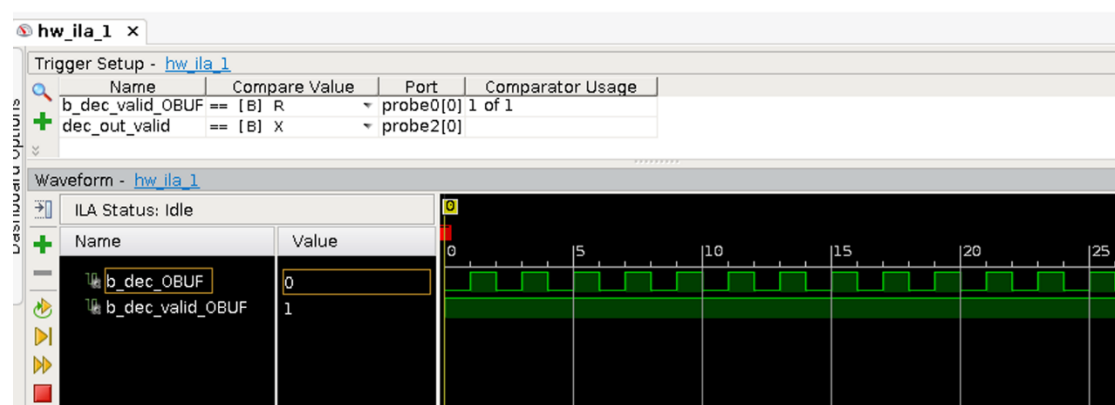


Figure 5.12 Decoding results capture

For LTE standard, the maximum number of C is $A(6143)$ in (5.1), the bit width E cannot be larger than 35 and F cannot be greater than 13. Therefore, it only takes 22 clock cycles to finish modulo operation. The latency and complexity comparison between this function and the division IP can be found in Table II.

Table 5-2 Modulo operation comparison

	Modulo function		IP	
Latency (clock)	22		38	
Complexity (1 module)	Slice	Slice	Slice	Slice
	LUT	register	LUT	register
	754	539	627	2407

Complexity (64 modules)	48256	34496	40128	154048
----------------------------	-------	-------	-------	--------

Table 5-2 shows that modulo function we designed can save nearly 3/4 slice registers compared to the IP from Xilinx although it takes a little more slice LUTs. It is significant that modulo function can save much more slice registers when the parallel level is high, e.g. $P = 64$, and it uses fewer clock cycles to complete the computation.

For LTE standard, block error rate (BLER) is used to test the decoder performance. The BLER of this decoder is evaluated via MATLAB simulation and ModelSim simulation. MATLAB simulation results are used as the reference of decoding BLER. The encoded frames with Gaussian white noise are first written into a test file and then read by Verilog test file. By Monte Carlo simulation, 1000 random frames for each signal-to-noise ratio (SNR) value, the BLER of different SNR can be obtained as shown in Figure 5.13. Since the simulation is performed without rate mapping, it works well even at very low SNR. The purpose of this simulation is to make sure that the decoder implementation is working as expected. The BLER results of this FPGA parallel decoder are similar with its MATLAB theoretical simulation. The BLER is slightly higher because of the fixed-point quantification error. Moreover, although the parallel decoder can increase system throughput and reduce latency, the BLER performance will be degraded.

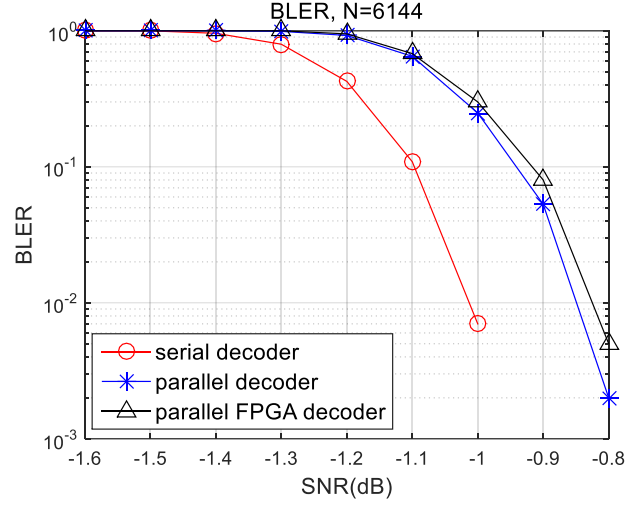


Figure 5.13 BLER simulation results

5.4 Summary

Parallel turbo decoding is a practical way to increase the system throughput and reduce the latency to meet the requirements of 5G network and beyond. In this paper, the implementation on FPGA of the parallel turbo decoder is introduced. A reverse address generator of interleaver/deinterleaver is proposed to reduce the processing time of each iteration and decrease latency further. The address generator uses recursion to generate the realtime address needed by the interleaver, which saves computing resources. A modulo function, that uses fewer clock cycles and logic resources compared to the Xilinx division IP, is designed to perform modulo operation. Moreover, to utilize the limited FPGA resources more efficiently, a double buffering technique is used to double the throughput in this parallel turbo decoder, which needs more storage space but reuses the logic resources.

The implementation of this decoder is accomplished on a Testbed system with 4 FPGA processors. On the one hand, by capturing the decoding results via Xilinx ILA, the validity of the parallel decoder is evaluated. On the other hand, the latency and BLER are tested by ModelSim simulation after placement and routing. The system decoding throughput is calculated based on the latency measured. Although the

throughput of this Testbed is less than 10Gbps, which is the demanded requirement of next generation wireless communication systems, this research gives a clue that parallel turbo decoder can be implemented on FPGA and meet multi-Gbps throughput requirement at the same time and that the throughput can be further improved by using more hardware resources.

6 Channel state estimation for indoor VLC MIMO Systems

The VLC MIMO channel matrix can be easily ill conditioned when the receivers are deployed at the centre of the LEDs, along the axes or when the users are carrying out unconstrained movement. This condition of channel makes it hard to maximize the performance of multi-user precoding. A channel state estimation method is proposed in this chapter to provide a solution for ill-conditioned indoor MIMO VLC. The proposed method can improve the system capacity via scheduling an appropriate precoding algorithm according to the condition of the channel.

6.1 VLC channel model

6.1.1 Indoor SISO Channel model

This chapter mainly focuses on indoor scenarios of VLC systems. Different with RF signal, the indoor VLC channel is more likely to consist with a line-of-sight (LOS) path and a diffuse path. In the case that there are no reflections at the receiver, a pure LOS link can represent the channel. When both LOS and diffuse signals are received, the VLC channel link is hybrid. There are also situations that the LOS is blocked by an object such as a person, the channel could be a pure diffuse link that the receiver detects the signal reflected by the other objects such as walls. Figure 6.1 gives the diagram of these three different channel scenarios.

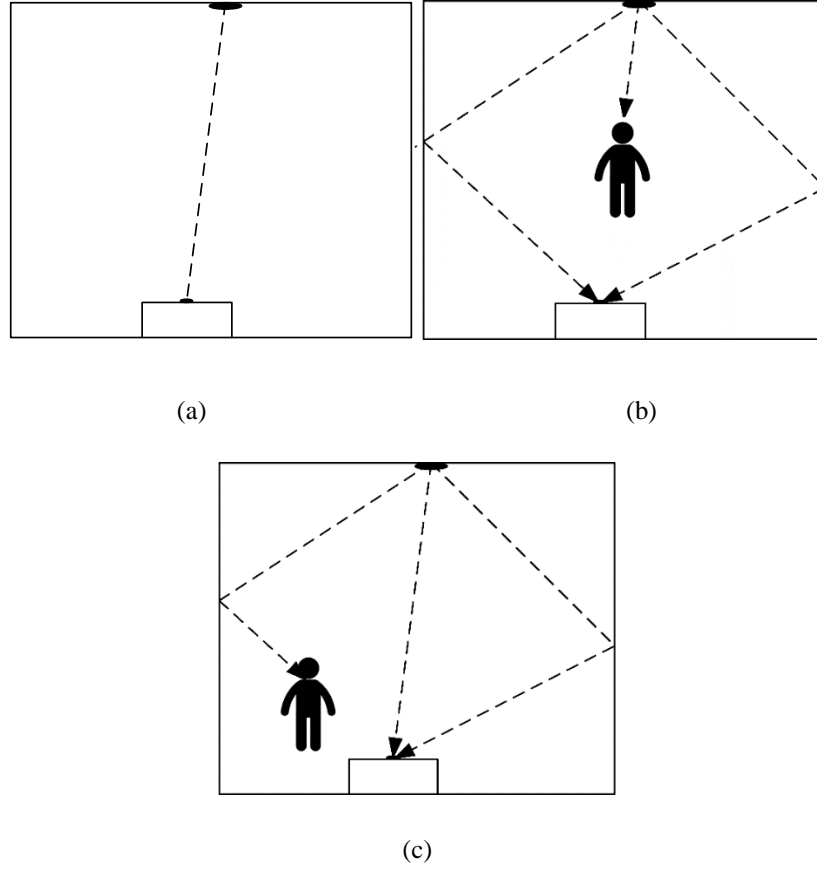


Figure 6.1 VLC channel model: (a) LOS; (b) Diffuse; (c) Hybrid VLC link

Figure 6.2 below shows the configuration of an indoor VLC system, where l , w , and h represent the height, width, and height of the room, respectively.

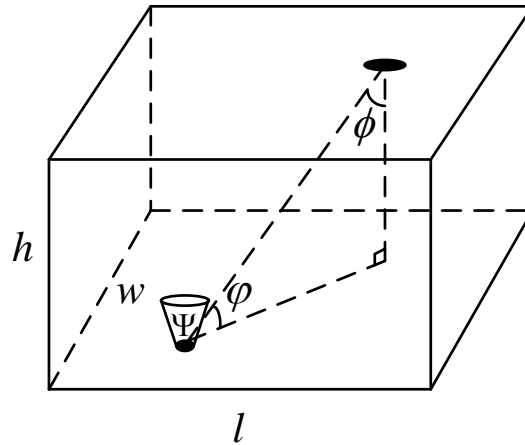


Figure 6.2 Indoor VLC configuration

In Figure 6.2, the transmitter is located on the ceiling and the receiver is situated on a desk. The distance between them is d and the angle of the irradiance and incidence

are ϕ and φ respectively. The LOS link does not suffer from large multipath interference because the diffuse signals are relatively small [133]. In an LOS link, the receive range is decided by the field-of-view (FOV) of photodetector. The gain of the LOS path η is [134][135]

$$\eta = \begin{cases} \frac{(m+1)A}{2\pi d^2} \cos^m(\phi) \cos \varphi g(\varphi), & \varphi < \Psi \\ 0, & \varphi > \Psi \end{cases}, \quad (6.1)$$

where m is the Lambertian emission order; Ψ is the FOV of the PD that has a physical surface area of A ; $g(\varphi)$ represents the optical concentrator gain which is given as [134]:

$$g(\varphi) = \frac{n_{oc}^2}{\sin^2(\Psi)}, \quad (6.2)$$

where n_{oc}^2 is the optical concentrator's refractive index.

A diffuse link, on the other hand, is more flexible due to the mobility of terminals and it suffers from the multipath dispersion. The indoor diffuse channel can be modelled as a low-pass filter that can be given as [135]:

$$h(\tau) = \frac{\varepsilon}{\tau_c} \exp\left(-\frac{\tau - \Delta\tau_{dif}}{\tau_c}\right) u(\tau - \Delta\tau_{dif}), \quad (6.3)$$

where τ_c is the time constant of the channel decay and $\Delta\tau_{dif}$ is the signal delay of the diffuse path; $u(\tau - \Delta\tau_{dif})$ denotes the unit impulse and ε is written as:

$$\varepsilon = \frac{A}{A_{room}} \frac{\rho}{1 - \rho}, \quad (6.4)$$

where ρ is the reflectivity of the indoor objects and A_{room} represents the surface area of the room which is

$$A_{room} = 2(l \cdot w + l \cdot h + w \cdot h). \quad (6.5)$$

A hybrid VLC link includes the LOS path and the reflections. It also can be modelled as a filter as [136]:

$$h(\tau) = \eta \delta(\tau - \Delta\tau_{LOS}) + \frac{\varepsilon}{\tau_c} \exp\left(-\frac{\tau - \Delta\tau_{dif}}{\tau_c}\right) u(\tau - \Delta\tau_{dif}), \quad (6.6)$$

where $\delta(\tau - \Delta\tau_{LOS})$ denotes the Dirac delta pulse and $\Delta\tau_{LOS}$ is the delay of the LOS path.

6.1.2 Indoor MIMO channel model

Although the bandwidth available for VLC is around 400 THz, white LEDs can only use several MHz of the electrical bandwidth [137]. Hence, to get a higher data rate and meet the demand of future communication systems, highly spectral-efficient techniques such as OFDM and MIMO can be employed. Besides, for indoor scenarios, it is natural that several LEDs are illuminating hence MIMO technique can be utilized to increase the system throughput.

In recent years, multiuser MIMO (MU-MIMO) has been studied for indoor VLC systems, where precoding schemes can be applied. The MIMO channel can be described by a channel matrix \mathbf{H} , which is:

$$\mathbf{H} = \begin{bmatrix} h_{11} & \cdots & h_{1M} \\ \vdots & h_{ij} & \vdots \\ h_{N1} & \cdots & h_{NM} \end{bmatrix}, \quad (6.7)$$

where M and N denotes the number of transmitters and receivers, respectively; h_{ij} represents the channel gain between transmitter i and receiver j . Figure 6.3 gives an example of VLC MU-MIMO configuration.

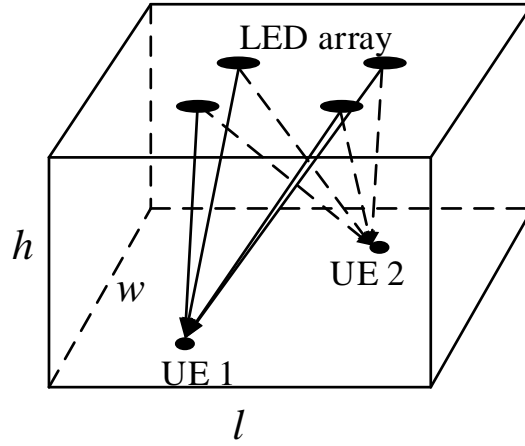


Figure 6.3 VLC MU-MIMO configuration

The VLC MU-MIMO is different with traditional RF MU-MIMO because of the real nonnegative-valued visible light signal. Besides, indoor VLC channels are highly correlated as LOS channels are most common which will degrade the MIMO performance. According to [138], OFDM can be applied in MIMO VLC systems by utilizing the channel gain and phase difference in the frequency domain for the MU-MIMO indoor scenario. And precoding techniques, such as conjugate precoding, zero forcing, and linear MMSE precoding, can be introduced to OFDM-MIMO systems as the multiuser precoding techniques [139][140].

6.2 Channel state estimation

Many literatures have discussed multiuser precoding for MIMO VLC systems. In [139], linear ZF and ZF dirty paper coding (ZF-DPC) are studied for MU-MIMO broadcasting for indoor VLC system. The performances of these two techniques are compared. According to [140], a robust MMSE linear precoding is introduced with imperfect channel estimation taken into account. A multiuser MIMO-OFDM for VLC with ZF and MMSE precoding is studied in [141]. All these methods require the knowledge of the channel and the performances of these precoding methods will be affected by the condition of the channel. As mentioned in [141], zero forcing is a good approach for high power or low noise scenarios, but its performance will be reduced if the channel matrix is ill-conditioned. Furthermore, according to [142], the

channel matrix can be ill conditioned when the receivers are deployed at the centre of the LEDs, along the axes or when the users are carrying out unconstrained movement. Hence, the channel of indoor MIMO VLC system can easily be ill-conditioned. Therefore, in this chapter, a channel state estimation (CSE) method is proposed to estimate the realtime channel state between the transmitters and the receivers. The method helps to decide which precoding method to choose to obtain a better system performance. To the best of the author's knowledge, there is no literature concerned with the channel state estimation for MIMO VLC systems yet.

6.2.1 System model

VLC systems are different with RF wireless systems. The uplink cannot be assumed to be realized via visual light. Therefore, VLC can only be used for a fast downlink and CSI can only be estimated at the terminals and fed back to the BS through other links such as WiFi [143]. The MIMO VLC system in this chapter is configured with M transmit LEDs at the BS and K user terminals with a single photodetector. The downlink received signal can be given as:

$$\mathbf{Y} = \mathbf{H}\mathbf{T}\mathbf{X} + \mathbf{N}, \quad (6.8)$$

where \mathbf{X} is the $K \times 1$ data symbols; \mathbf{H} and \mathbf{T} denote the $K \times M$ channel response matrix and $M \times K$ precoding matrix respectively; \mathbf{N} stands for white Gaussian noise and $E(\mathbf{N}\mathbf{N}^H) = \sigma^2 \mathbf{I}$. Here the total transmit power is constrained to normalized power.

Let the downlink channel estimation result be $\tilde{\mathbf{H}}$, the precoding matrix \mathbf{T} with different precoding methods can be expressed as:

$$\begin{aligned} \mathbf{T} &= \tilde{\mathbf{H}}^H, & MF \\ \mathbf{T} &= \tilde{\mathbf{H}}^H (\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H)^{-1}, & ZF \\ \mathbf{T} &= \tilde{\mathbf{H}}^H (\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H + \alpha \mathbf{I})^{-1}, & MMSE \end{aligned} \quad (6.9)$$

where $(\cdot)^H$ means matrix Hermitian transpose; and α is the loading factor vector of MMSE precoding, which is used to maximize the SINR at the receiver.

According to (6.9), we can see that CJ precoding does not need matrix inversion so it has the lowest computation complexity and does not rely on the channel to be good condition. ZF precoding can achieve linear capacity growth with $\min(M, K)$ but the SNR at the receivers will be dramatically reduced when the channel matrix is ill-conditioned. Moreover, the channel inversion will be problematic if the channel is rank-deficient. MMSE precoding can solve the ill-conditioned channel problem but its performance needs to be evaluated. Therefore, how to choose the precoding method becomes a key practical problem. A CSE is introduced as shown in Figure 6.4 to estimate the realtime condition of the channel and help to choose the precoding method according to the corresponding channel capacity of the system.

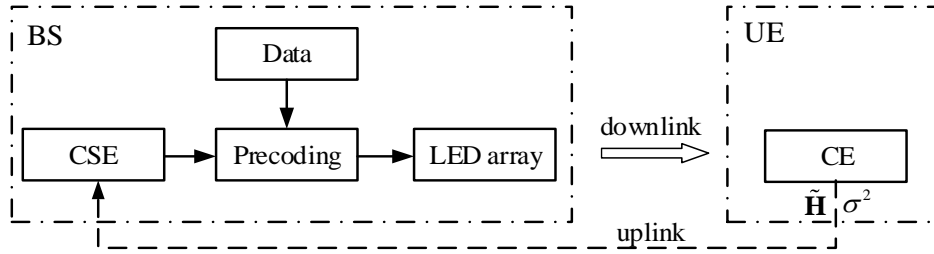


Figure 6.4 Block diagram of precoding with CSE

6.2.2 Proposed CSE

Like in the RF MIMO systems, the downlink receiver can estimate the downlink channel and the power of noise σ^2 and feed the results back to the BS as shown in Figure 6.4. To obtain the downlink capacity of a multiuser MIMO system with CJ precoding, the effective SINR at each terminal is evaluated and given as [144]

$$\text{SINR}_{CJ} = \frac{\frac{1}{K+1} \left[\left(\sum_{l=1}^K \lambda_l \right)^2 + \sum_{l=1}^K (\lambda_l)^2 \right]}{\sum_{l=1}^K (\lambda_l)^2 - \frac{1}{K+1} \left[\left(\sum_{l=1}^K \lambda_l \right)^2 + \sum_{l=1}^K (\lambda_l)^2 \right] + K\sigma^2}, \quad (6.10)$$

where λ_i denotes the i th eigenvalue of $\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H$. So the total capacity will be:

$$C_{CJ} \approx K \log_2 [1 + \text{SINR}_{CJ}]. \quad (6.11)$$

MMSE precoding, which is also called regularized ZF precoding, can improve the performance of ZF precoding by pseudo-inverse regularization for rank-deficient channels. The SINR and achievable capacity are [144]:

$$\text{SINR}_{MMSE} = \frac{\frac{1}{K+1} \left[\left(\sum_{i=1}^K \frac{\lambda_i}{\lambda_i + \alpha} \right)^2 + \sum_{i=1}^K \left(\frac{\lambda_i}{\lambda_i + \alpha} \right)^2 \right]}{\sum_{i=1}^K \left(\frac{\lambda_i}{\lambda_i + \alpha} \right)^2 - \frac{1}{K+1} \left[\left(\sum_{i=1}^K \frac{\lambda_i}{\lambda_i + \alpha} \right)^2 + \sum_{i=1}^K \left(\frac{\lambda_i}{\lambda_i + \alpha} \right)^2 \right] + K\sigma^2}, \quad (6.12)$$

and

$$C_{MMSE} \approx K \log_2 [1 + \text{SINR}_{MMSE}] \quad (6.13)$$

respectively. When $\alpha=0$, MMSE precoding becomes ZF precoding. If the channel is ill-conditioned, some of the singular values of $\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H$ are very large, which results in very low SNR at some receivers, so the overall capacity of the system is reduced. When $\alpha = K\sigma^2$, the SINR of MMSE precoding is optimized. And the optimized SINR of MMSE precoding will be:

$$\text{SINR}_{MMSE} = \frac{\frac{1}{K+1} \left[\left(\sum_{i=1}^K \frac{\lambda_i}{\lambda_i + K\sigma^2} \right)^2 + \sum_{i=1}^K \left(\frac{\lambda_i}{\lambda_i + K\sigma^2} \right)^2 \right]}{\sum_{i=1}^K \left(\frac{\lambda_i}{\lambda_i + K\sigma^2} \right)^2 - \frac{1}{K+1} \left[\left(\sum_{i=1}^K \frac{\lambda_i}{\lambda_i + K\sigma^2} \right)^2 + \sum_{i=1}^K \left(\frac{\lambda_i}{\lambda_i + K\sigma^2} \right)^2 \right] + K\sigma^2}. \quad (6.14)$$

To choose a precoding method with better performance, we simply compare the SINR of these methods. As we know that MMSE can perform better than ZF, only CJ and MMSE precoding are compared here.

Apart from the channel matrix $\tilde{\mathbf{H}}$, the noise power σ^2 of received signal is also fed back to the BS for the CSE to evaluate SINR. As SINR is variable, due to the channel not always being constant, it is hard to obtain an equation of the comparison results between SINR_{CJ} and SINR_{MMSE} . Therefore, the CSE uses the following two steps to compare SINR_{CJ} and SINR_{MMSE} :

- A. Calculate the eigenvalues of $\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H$;
- B. Calculate the SINR_{CJ} and SINR_{MMSE} .

If $\text{SINR}_{CJ} > \text{SINR}_{MMSE}$, conjugate precoding is chosen as the precoding method. Otherwise, MMSE precoding is used. As a result, by using CSE, the system can automatically choose the precoding method with larger capacity.

6.3 Simulation results

In this section, the evaluation of the proposed method is given via MATLAB numerical simulation. The channel model of an indoor VLC system is firstly simulated. Table 6-1 gives the simulation parameters for one transmitter and one receiver scenario. The transmitter is placed on the roof of the room while the receiver is on the ground floor. Figure 6.5 shows the result of the channel gain in this simulation according to (6.1).

Table 6-1 Simulation parameters

Half power angle (degree)	70
Transmit power of the LED (dBm)	20
Transmitter position axis (x, y)	0,0
Detector physical area of a PD (m ²)	1e-4
Gain of an optical filter	1
Distance between source and receiver plane (m)	3
Refractive index of a lens at a PD	1.5
Field of View of a receiver (degree)	45
Room dimension (m)	5x5x3

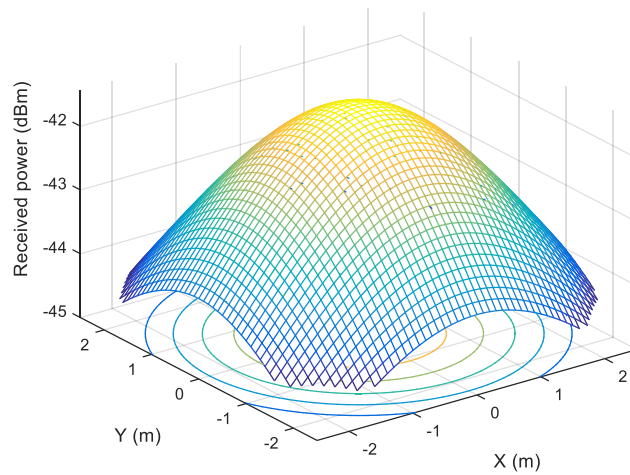


Figure 6.5 Channel gain of a single transmitter

For VLC MIMO systems, the channel gain distribution is different with one transmitter scenario. Table 6-2 gives the simulation parameters that each transmitter has the same specification as Table 6-1 but 4 LEDs are placed on the roof.

Table 6-2 Simulation parameters of MIMO VLC

Theta Half Power (degree)	70
Transmit power of a LED (dBm)	20
Detector physical area of a PD (m ²)	1e-4
Gain of an optical filter	1
Distance between source and receiver plane (m)	3
Refractive index of a lens at a PD	1.5
Field of View of a receiver (degree)	45
Room dimension (m)	5x5x3
LED 1 Position axis	-1, -1

LED 2 Position axis	-1, 1
LED 3 Position axis	1, -1
LED 4 Position axis	1, 1

Considering the situation of multiple receivers, the channel matrix can be ill-conditioned when the channel matrix is not full rank. For example, suppose there are 2 receivers. The axis of the first one is (-1.9993, -1.7656) and the second one is (-1.7107, -1.9870). The channel matrix is

$$\mathbf{H} = \begin{bmatrix} 0.4136 & 0.4201 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \times e^{-4},$$

and the rank of this channel matrix is the only one which is ill-conditioned. The channel gain of this VLC MIMO configuration is shown in Figure 6.6.

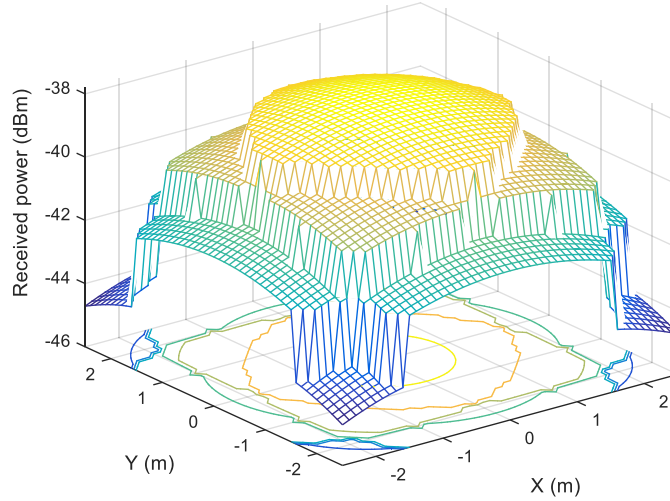


Figure 6.6 Channel gain of 4 transmitters

The capacity performance of the MU-MIMO VLC system with CSE is evaluated via Monte Carlo simulation. The MIMO configuration is set to be $M = 10$ and $K = 4$ with different signal to noise ratios (SNRs). The position of the transmitter array is $(x = [-1, -0.5, 0, 0.5, 1], y = [-0.5, 0.5])$ and the receivers are randomly located in this room. Here SNR equals $1/\sigma^2$ as the transmit power is normalized. The system

capacity with CJ or MMSE precoding only are used as comparisons. Figure 6.7 shows that the CSE can help choose the precoding method with larger capacity automatically and switch between them according to the corresponding SINRs.

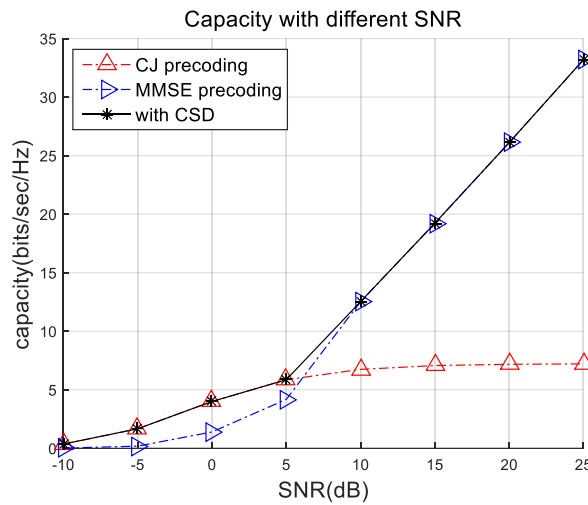


Figure 6.7 System capacity with different SNR

When the SNR is 0dB, the system with CSE will choose CJ precoding and the capacity is larger than that with MMSE precoding. When the SNR is 10dB, choosing MMSE precoding can achieve a larger capacity.

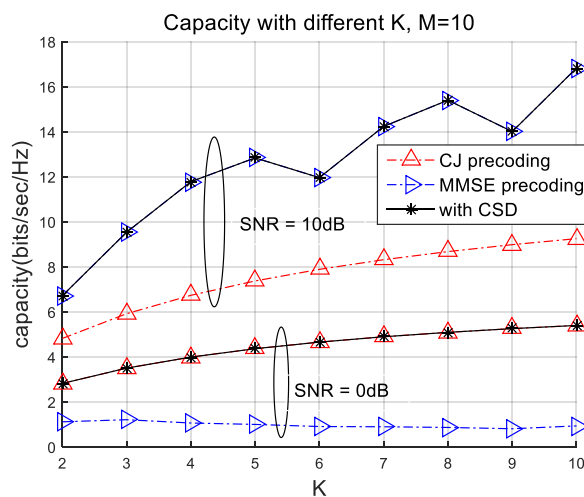


Figure 6.8 System capacity with different terminal number

6.4 Summary

In this chapter, a channel state estimation method is proposed for MU-MIMO indoor VLC systems. The architecture of VLC systems is first introduced including the physical layer, modulation techniques, and the channel models. The channel state estimation method utilizes the channel condition as well as the corresponding precoding methods. First, the SINRs with different precoding methods are evaluated by utilizing the downlink SNR and CSI fed back from the user terminals. Second, the precoding method with larger SINR is chosen to obtain larger system capacity. From the simulation results, we can conclude that with CSE, the system can achieve larger capacity than that with CJ or MMSE precoding only.

7 Conclusions

7.1 Discussions

In this thesis, some practical techniques have been studied to enhance the system capacity of 5G network and beyond. As a part of 5G network, HetNet is suffering from inter-cell interference. A CRS interference cancellation method is proposed to handle the inter-cell interference caused by the CRS from the macrocell. This method takes the CFO and TO of the users' received signal into account. By estimating the interfering channel and reconstructing the interfering signal, the interference can be cancelled. The system BLER performance is improved according to the simulation results.

For massive MIMO systems, multi-user precoding is used to improve the system capacity. However, due to RF impairment, the performance of precoding is degraded. To tackle the impairment problem, a closed-loop reciprocity calibration method is proposed. A TD is introduced in this method to assist the calculation of calibration coefficient. For downlink, the TD estimates the downlink CSI, that includes the hardware impairments, and feeds the results back to the BS. For uplink, the TD sends pilots and the BS will estimate the uplink channel. This closed-loop calibration can improve the system capacity compared to the existing methods and also support online calibration.

In 5G network, apart from the requirement of capacity, low latency is another critical demand. Channel decoding is normally a time-consuming part of the receiver especially for serial decoding algorithm like BCJR of turbo codes. Implementing the decoding algorithm in parallel is a practical way to improve the latency as well as the throughput performance. A reverse interleave address generator is proposed in thesis for a parallel turbo decoder which is implemented on a FPGA platform that can meet multi-Gbps throughput with a latency as little as 56.8us. The implementation results

also show that a lower latency can be achieved if increasing the decoding parallelism with more hardware resource used.

In addition to the above techniques, this thesis also investigated multi-user precoding in indoor MIMO VLC systems. As a green and secure technology, VLC could become a part of 5G network and beyond. The study in this part shows that it is important that the system can acquire the information of the channel state because the indoor MIMO VLC channel can easily be ill-conditioned. A channel state estimation method is proposed based on the SINR of the received signal. Simulation results show that the proposed method can improve the system capacity.

7.2 Further work

To enhance the system capacity for 5G network and beyond, the above techniques have been studied. Although these methods can enhance the system capacity for 5G and beyond network to some extent, some improvements can be made and achieved in the future.

- 1) The current method used in CRS interference cancellation is only applicable to the non-colliding scenario. For colliding scenario, the interference model should be different and needs more study.
- 2) The reciprocity calibration studied in this thesis gives a closed-loop calibration under the condition that the impairments estimation is accurate. The impact of impairments estimation error on the calibration results should be considered in the future.
- 3) A low latency parallel turbo decoder was implemented with some latency improvement because of the reverse interleave address generator. The address generator brings limited latency improvement hence some low latency decoding algorithm or implementation structure can be studied in the future. Moreover, a more

efficient way of the decoder implementation can be investigated to reduce the cost of hardware resource.

4) Multi-user precoding of MIMO VLC systems is still at the theory study stage. Some experiment result is important to support the research in the future. A testbed can be built, and a measurement of the MIMO VLC channel will be helpful.

Reference

- [1]. <https://www.statista.com/statistics/274774/forecast-of-mobile-phone-users-worldwide/>
- [2]. T. S. Rappaport, G. R. MacCartney, M. K. Samimi and S. Sun, "Wideband Millimeter-Wave Propagation Measurements and Channel Models for Future Wireless Communication System Design," *IEEE Trans. Comm.*, Vol. 63, No. 9, Sept. 2015, pp. 3029-3056.
- [3]. G. R. Maccartney, T. S. Rappaport, S. Sun and S. Deng, "Indoor Office Wideband Millimeter-Wave Propagation Measurements and Channel Models at 28 and 73 GHz for Ultra-Dense 5G Wireless Networks," *IEEE Access*, Vol. 3, 2388-2424, October 2015.
- [4]. Ozgur Ergul, Ergin Dinc , Ozgur B. Akan, "Communicate to illuminate: State-of-the-art and research challenges for visible light communications." *Physical Communication* , vol.17, 72-85, 2015
- [5]. S. Wu, H. Wang and C. H. Youn, "Visible light communications for 5G wireless networking systems: from fixed to mobile communications," *IEEE Network*, vol. 28, no. 6, pp. 41-45, Nov.-Dec. 2014.
- [6]. T. E. Bogale and L. B. Le, "Massive MIMO and mmWave for 5G Wireless HetNet: Potential Benefits and Challenges," *IEEE Vehicular Technology Magazine*, vol. 11, no. 1, pp. 64-75, March 2016.
- [7]. <http://www.5gtf.org/>
- [8]. TS V5G.211 V1.7 : Verizon 5G TF; Air Interface Working Group; Verizon 5th Generation Radio Access; Physical channels and modulation (Release 1), Oct. 2016.
- [9]. H. Luo, W. Li, Y. Zhang, L. K. Huang, J. Cosmas and Q. Ni, "CRS interference cancellation algorithm for heterogeneous network," in *Electronics Letters*, vol. 52, no. 1, pp. 77-79, 1 8 2016.
- [10]. S. Coleri, M. Ergen, A. Puri, and A. Bahai, "Channel estimation techniques based on pilot arrangement in OFDM systems", *IEEE Trans. Broadcast.*, vol. 48, no. 3, pp. 223–229, Sep. 2002.
- [11]. H. Luo, Y. Zhang, L. K. Huang, J. Cosmas and A. Aggoun, "A Closed-Loop Reciprocity Calibration Method for Massive MIMO in Terrestrial Broadcasting Systems," in *IEEE Transactions on Broadcasting*, vol. 63, no. 1, pp. 11-19, March 2017.
- [12]. P. Robertson, E. Villebrun, and P. Hoeher, "A comparison of optimal and sub-optimal MAP decoding algorithms operating in the log domain," *IEEE Int. Conf. on Communications*, vol. 2, Seattle, WA, USA, June 1995, pp. 1009–1013.
- [13]. Aleksandar J., Junyi L, Tom R., Qualcomm Research: 'Visible light communication: Opportunities, challenges and the path to market', *IEEE Commun. Mag.*, 51, pp. 26–32, Dec. 2013.
- [14]. J. Vućić, Christoph K, Stefan N., et al.: '230 Mbit/s via a wireless visible-light link based on OOK modulation of phosphorescent white LEDs', in *Proc. OFC/NFOEC*, San Deigo, CA, USA, pp. 1-3, Mar. 2010.
- [15]. Dominic C.O'Brien: 'Visible light communications: Challenges and potential', in *Proc. IEEE photonics Conf.*, pp. 365-366, 2011
- [16]. Dobroslav T., Sinan S., Harald H.: 'Practical MIMO capacity for indoor optical wireless communication with white LEDs', in *IEEE VTC Spring*, pp. 1–5, 2013

- [17]. B. E. Priyanto, S. Kant, F. Rusek, S. Hu, J. Chen and C. Wugengshi, "Robust UE Receiver with Interference Cancellation in LTE Advanced Heterogeneous Network," *2013 IEEE 78th Vehicular Technology Conference (VTC Fall)*, Las Vegas, pp. 1-7, Nov. 2013.
- [18]. Fundamentals of Wireless Communication, Cambridge University Press; 1 edition (July 11, 2005)
- [19]. H. Hashemi, "Impulse response modeling of indoor radio propagation channels," in *IEEE Journal on Selected Areas in Communications*, vol. 11, no. 7, pp. 967-978, Sep 1993.
- [20]. E. Björnson, J. Hoydis, M. Kountouris and M. Debbah, "Massive MIMO Systems With Non-Ideal Hardware: Energy Efficiency, Estimation, and Capacity Limits," in *IEEE Transactions on Information Theory*, vol. 60, no. 11, pp. 7112-7139, Nov. 2014.
- [21]. Jian, Q. and S. Aissa, "On the Power Amplifier Nonlinearity in MIMO Transmit Beamforming Systems." *Communications, IEEE Transactions on* 60(3): 876-887, 2012.
- [22]. Cheng-Po Liang, Je-Hong Jong, W. E. Stark and J. R. East, "Nonlinear amplifier effects in communications systems," in *IEEE Transactions on Microwave Theory and Techniques*, vol. 47, no. 8, pp. 1461-1466, Aug 1999.
- [23]. M. T. Le and L. Thibault, "Performance evaluation of COFDM for digital audio broadcasting part II: Effects of HPA nonlinearities", *IEEE Trans. Broadcast.*, vol. 44, pp. 165–171, June 1998.
- [24]. J. Boccuzzi, "Performance evaluation of nonlinear transmit power amplifiers for North American digital cellular portables", *IEEE Trans. Veh. Technol.*, vol. 44, pp. 220–228, May 1995
- [25]. N. Zimmermann, "Design and implementation of a broadband RF-DAC transmitter for wireless communications: Universitätsbibliothek, 2011
- [26]. A. A. Abidi, "Direct-conversion radio transceivers for digital communications," *Solid-State Circuits, IEEE Journal of*, vol. 30, pp. 1399-1410, 1995.
- [27]. W. Li, Y. Zhang, L. K. Huang, J. Cosmas, C. Maple and J. Xiong, "Self-IQ-Demodulation Based Compensation Scheme of Frequency-Dependent IQ Imbalance for Wideband Direct-Conversion Transmitters," in *IEEE Transactions on Broadcasting*, vol. 61, no. 4, pp. 666-673, Dec. 2015.
- [28]. R. A. Casas, S. L. Biracree, and A. E. Youtz, "Time domain phase noise correction for OFDM signals," *IEEE Trans. Broadcast.*, vol. 48, no. 3, pp. 230–236, Sept. 2002.
- [29]. M. R. Gholami, S. Nader-Esfahani, and A. A. Eftekhar, "A new method of phase noise compensation in OFDM," in *Proc. IEEE Int. Conf. on Communications (ICC)*, May 2003, pp. 3443–3446.
- [30]. D. D. Lin, R. A. Pacheco, T. J. Lim, and D. Hatzinakos, "Joint estimation of channel response, frequency offset, and phase noise in OFDM", *IEEE Trans. Signal Processing*, vol. 54, no. 9, pp. 3542–3554, Sept. 2006.
- [31]. Q. Zou, A. Tarighat, and A. H. Sayed, "Compensation of phase noise in OFDM wireless systems," *IEEE Trans. Signal Processing*, vol. 55, no. 11, pp. 5407–5424, Nov. 2007.
- [32]. Hamming, R. W., ["Error Detecting and Error Correcting Codes"](#) . Bell System Tech. J. USA: AT&T. 29 (2): 147–160, April 1950
- [33]. Khan, L. U. "Visible light communication: Applications, architecture, standardization and research challenges." *Digital Communications and Networks*

- [34]. Damnjanovic A., Wei, Yongbin, "A SURVEY ON 3GPP HETEROGENEOUS, Wireless Communications," vol. 18, no. 3, June 2011.
- [35]. Heli, Z., Shanzhi, C., Xi, L., Hong, J., and Xiaojiang, D.: 'Interference management for heterogeneous networks with spectral efficiency improvement', *IEEE Wirel. Commun. Mag.*, 2015, 22, pp. 101–107
- [36]. Deb, S., Monogioudis, P., Miernik, J., and Seymour, J.P., "Algorithms for enhanced inter-cell interference coordination (eICIC) in LTE HetNets", *IEEE/ACM Trans. Netw.*, 2014, 22, pp. 137–150
- [37]. S. N. S. Kshatriya, S. Kaimalettu, S. R. Yerrapareddy, K. Milleth, and N. Akhtar, "On interference management based on subframe blanking in Heterogeneous LTE networks," in *Communication Systems and Networks (COMSNETS)*, 2013 Fifth International Conference on, 2013, pp. 1-7
- [38]. "Overview of the 3GPP Long Term Evolution Physical Layer", Jim Zyren, *white paper*, Freescale Semiconductor
- [39]. 3GPP TS 36.101, "Evolved universal terrestrial radio access (E-UTRA), user equipment (UE) radio transmission and reception," v. 11.0 Mar. 2012.
- [40]. N.-L. Hung, T. Le-Ngoc, and K. Chi-Chung, "RLS-Based Joint Estimation and Tracking of Channel Response, Sampling, and Carrier Frequency Offsets for OFDM," *Broadcasting, IEEE Transactions on*, vol. 55, pp. 84-94, 2009.
- [41]. Tyrrell, A., and Schoeneich, H.: 'Advanced receiver signal processing techniques: Evaluation and characterization', *ARTIST4G D2.2 Deliverable*, June 2011
- [42]. Yoo, T., Luo, T. Mallik, S. Damnjanovic, A. Montojo, J. 'Dynamic switching between common reference signal interference cancellation & resource element puncturing in a co-channel heterogeneous network'. *Patent US2012/0087261 A1*, 2013
- [43]. Qualcomm Inc.: 'Enabling communication in harsh interference scenarios', R4-102350, 3GPP-RAN WG4 AH#10-03, Bratislava, 2010
- [44]. Qualcomm Inc.: 'Link level simulations for FeICIC with 9 dB cell range expansion', R4-123313, 3GPP-RAN WG4 #63, Prague, May 2012
- [45]. Huang, M., and Xu, W.: 'Macro-femto inter-cell interference mitigation for 3GPP LTE-A downlink'. *IEEE Wireless Communication and Network Conf*, April 2012, pp. 75–80
- [46]. Priyanto, B.E., Kant, S., Rusek, F., Hu, S., Chen, J., and Wugenshi, C.: 'Robust UE receiver with interference cancellation in LTE advanced heterogeneous network'. *IEEE Vehicular Technology Conf.*, Las Vegas, September 2013, pp. 1–7
- [47]. J. J. van de Beek, M. Sandell, and P. O. Borjesson, "ML estimation of time and frequency offset in OFDM systems," *Signal Processing, IEEE Transactions on*, vol. 45, pp. 1800-1805, 1997.
- [48]. B. Jong-Seob and S. Jong-Soo, "Effective Symbol Timing Recovery Based on Pilot-Aided Channel Estimation for MISO Transmission Mode of DVB-T2 System," *Broadcasting, IEEE Transactions on*, vol. 56, pp. 193-200, 2010.
- [49]. Tsai, Y.-H., and Sang, T.-H.: 'A new timing synchronization and cell search procedure resistant to carrier frequency offsets for 3GPP-LTE downlink'. *IEEE Int. Conf. on Communication in China (ICCC)*, Beijing, August 2012, pp. 334–338
- [50]. F. Cheng, C. Hongyu, M. Meng, and J. Bingli, "On Statistical Properties of Co-channel Interference in OFDM Systems," *Communications Letters, IEEE*, vol. 17, pp. 2328-2331, 2013.

- [51]. Van Veen, B.D.; Buckley, K.M, "[Beamforming: A versatile approach to spatial filtering](#)" (PDF). *IEEE ASSP Magazine*. 5 (2): 4, 1988.
- [52]. F. Rusek et al., "Scaling Up MIMO: Opportunities and Challenges with Very Large Arrays," in *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40-60, Jan. 2013.
- [53]. E. G. Larsson, O. Edfors, F. Tufvesson and T. L. Marzetta, "Massive MIMO for next generation wireless systems," in *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186-195, February 2014.
- [54]. T. L. Marzetta, "Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas," in *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590-3600, November 2010.
- [55]. R. Rogalin et al., "Scalable synchronization and reciprocity calibration for distributed multiuser MIMO," *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, pp. 1815–1831, Apr. 2014.
- [56]. C. Shepard et al., "Argos: Practical many-antenna base stations," in *Proc. 18th Annu. Int. Conf. Mobile Comput. Netw.*, Istanbul, Turkey, pp. 53–64, 2012.
- [57]. Huberman, S. and T. Le-Ngoc, "MIMO Full-Duplex Precoding: A Joint Beamforming and Self-Interference Cancellation Structure." *IEEE Transactions on Wireless Communications* 14(4): 2205-2217, 2015.
- [58]. Van Veen, B. D. and K. M. Buckley, "Beamforming: a versatile approach to spatial filtering." *ASSP Magazine, IEEE* 5(2): 4-24, 1988.
- [59]. Nasser, Mona, "Iterative Channel Estimation Algorithm in Multiple Input Multiple Output Orthogonal Frequency Division Multiplexing Systems". *Journal of Computer Science*. 6 (2): 224–228, Feb. 2010
- [60]. Q. H. Spencer, C. B. Peel, A. L. Swindlehurst, and M. Haardt, "An introduction to the multi-user MIMO downlink," *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 60–67, Oct. 2004.
- [61]. E. Björnson, R. Zakhour, D. Gesbert, B. Ottersten, [Cooperative Multicell Precoding: Rate Region Characterization and Distributed Strategies with Instantaneous and Statistical CSI](#), *IEEE Transactions on Signal Processing*, vol. 58, no. 8, pp. 4298-4310, 2010
- [62]. A. Wiesel, Y. C. Eldar and S. Shamai, "Zero-Forcing Precoding and Generalized Inverses," in *IEEE Transactions on Signal Processing*, vol. 56, no. 9, pp. 4409-4418, Sept. 2008.
- [63]. C. B. Peel, B. M. Hochwald and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication-part I: channel inversion and regularization," in *IEEE Transactions on Communications*, vol. 53, no. 1, pp. 195-202, Jan. 2005.
- [64]. D. Tandur and M. Moonen, "Joint adaptive compensation of transmitter and receiver IQ imbalance under carrier frequency offset in OFDM-based systems," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5246–5252, Nov. 2007.
- [65]. Bostamam, Y. Sanada, and H. Minami, "IQ imbalance compensation scheme in the presence of frequency offset and dynamic DC offset for a direct conversion receiver," *IEEE Trans. Wireless Commun.*, vol. 8, no. 5, pp. 2214–2220, May 2009.
- [66]. A. Loke and F. Ali, "Direct conversion radio for digital mobile phones design issues, status, and trends," *IEEE Trans. Microw. Theory Techn.*, vol. 50, no. 11, pp. 2422–2435, Nov. 2002.
- [67]. Pekka Kyösti, Tommi Jämsä, and Jukka-Pekka Nuutinen, "Channel Modelling for Multiprobe Over-the-Air MIMO Testing," *International Journal of Antennas and Propagation*, vol. 2012, 2012.

- [68]. Inamori, M.; Bostamam, A.M.; Sanada, Y.; Minami, H., "IQ imbalance compensation scheme in the presence of frequency offset and dynamic DC offset for a direct conversion receiver," *IEEE Trans. Wireless Comm.*, vol.8, no.5, pp.2214-2220, May 2009
- [69]. Petermann, M., Stefer, M., Ludwig, F., Wubben, et al., "Multi-User Pre-Processing in Multi-Antenna OFDM TDD Systems with Non-Reciprocal Transceivers," *IEEE Trans. Comm.*, vol. 61, pp. 3781-3793, Aug. 2013
- [70]. Jinho Choi, "Downlink Multiuser Beamforming With Compensation of Channel Reciprocity From RF Impairments," *IEEE Transactions Comm.* , vol.63, no.6, pp.2158-2169, June 2015
- [71]. Pekka Kyösti, Tommi Jämsä, and Jukka-Pekka Nuutinen, "Channel Modelling for Multiprobe Over-the-Air MIMO Testing," *International Journal of Antennas and Propagation*, vol. 2012, 2012. doi:10.1155/2012/615954.
- [72]. Ya Jing; Zhu Wen; Hongwei Kong; Duffy, S.; Rumney, M., "Two-stage over the air (OTA) test method for MIMO device performance evaluation," *IEEE Intern. Symp. on Antennas and Propagation (APSURSI)*, pp.71-74, 3-8 July 2011
- [73]. Moray Rumney, Ryan Pirkel, Markus Herrmann Landmann, and David A. Sanchez-Hernandez, "MIMO Over-The-Air Research, Development, and Testing," *International Journal of Antennas and Propagation*, vol. 2012, 2012. doi:10.1155/2012/467695
- [74]. Wence, Zhang, Hong, Ren, Cunhua, Pan, Ming, Chen, et al., "Large-Scale Antenna Systems With UL/DL Hardware Mismatch: Achievable Rates Analysis and Calibration," *IEEE Trans. Comm.*, vol.63, pp. 1216-1229, Apr. 2015
- [75]. Upamanyu M., "Introduction to Communication Systems," *Cambridge University Press*, Nov. 2014, p.12
- [76]. Changyong Shin; Heath, R.W.; Powers, E.J., "Blind Channel Estimation for MIMO-OFDM Systems," *IEEE Trans. Vehicular Tech.*, vol.56, no.2, pp.670-685, Mar. 2007
- [77]. H. Aumann, A. Fenn, and F. Willwerth, "Phased array antenna calibration and pattern prediction using mutual coupling measurements," *IEEE Trans. Ant. Prop.*, vol. 37, no. 7, July 1989.
- [78]. T. Takahashi, Y. Konishi, S. Makino, H. Ohmine and H. Nakaguro, "Fast Measurement Technique for Phased Array Calibration," in *IEEE Transactions on Antennas and Propagation*, vol. 56, no. 7, pp. 1888-1899, July 2008.
- [79]. O'Keeffe, C. and M. O'Brien (2015). Communication system, apparatus and methods for calibrating an antenna array, *Patent, US9035828 B2*, May 19, 2015.
- [80]. Höhne, H. T., et al, "Calibration of smart antenna systems", *Google Patents, US20120020396 A1*, Jan 26, 2012.
- [81]. Hall, D., Hinde, A., Yupeng, Jia, "Multi-channel Phase-Coherent RF Measurement System Architectures and Performance Considerations," *IEEE Military Comm. Conf. (MILCOM)*, pp. 1318-1323, Oct. 2014
- [82]. Hung Nguyen-Le; LE-NGOC, THO; Chi Chung Ko, "RLS-Based Joint Estimation and Tracking of Channel Response, Sampling, and Carrier Frequency Offsets for OFDM," *IEEE Trans. Broadcasting*, vol.55, no.1, pp.84-94, Mar. 2009
- [83]. Bjornson, E.; Matthaiou, M.; Debbah, M., "Massive MIMO with Non-Ideal Arbitrary Arrays: Hardware Scaling Laws and Circuit-Aware Design," *IEEE Trans. Wireless Comm.* , vol.14, no.8, pp.4353-4368, Aug. 2015

- [84]. [Shannon, C.E.](#), "[A Mathematical Theory of Communication](#)", *Bell System Technical Journal*, 27, pp. 379–423 & 623–656, July & October, 1948.
- [85]. Arikan, E., "Channel Polarization: A Method for Constructing Capacity-Achieving Codes for Symmetric Binary-Input Memoryless Channels". *IEEE Transactions on Information Theory*. 55 (7): 3051–73. [arXiv:0807.3917v5](#), July 2009.
- [86]. L.Bahl, J.Cocke, F.Jelinek, and J.Raviv, "Optimal Decoding of Linear Codes for minimizing symbol error rate", *IEEE Transactions on Information Theory*, vol. IT-20(2), pp. 284-287, March 1974.
- [87]. J. P. Woodard and L. Hanzo, "Comparative study of turbo decoding techniques: An overview," *IEEE Trans. Veh. Technol.*, vol. 49, no. 6, pp. 2208-2233, Nov. 2000.
- [88]. Y. G. Debessu, H. C. Wu, H. Jiang and S. Mukhopadhyay, "New Modified Turbo Decoder for Embedded Local Content in Single-Frequency Networks," *IEEE Trans. Broadcast.*, vol. 59, no. 1, pp. 129-135, March 2013.
- [89]. "IEEE Standard for Air Interface for Broadband Wireless Access Systems," *IEEE Std 802.16-2012* (Revision of IEEE Std 802.16-2009) , vol., no., pp.1-2542, Aug. 17 2012.
- [90]. LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); Multiplexing and Channel Coding, ETSI, Sophia Antipolis, France, Feb. 2013.
- [91]. L. Christodoulou, O. Abdul-Hameed, A. M. Kondo and J. Calic, "Adaptive Subframe Allocation for Next Generation Multimedia Delivery Over Hybrid LTE Unicast Broadcast," *IEEE Trans. Broadcast.*, vol. 62, no. 3, pp. 540-551, Sep. 2016.
- [92]. Altera, "3GPP LTE Turbo Reference Design", Jan. 2011
- [93]. 3GPP Technical Specification: Group Radio Access Network, Evolved Universal Terrestrial Radio Access, Multiplexing and Channel Coding (Release 8), TS 36.212 v8.3.0, May 2007.
- [94]. L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inf. Theory*, vol. 20, no. 2, pp. 284–287, Mar. 1974.
- [95]. Hamid R. Sadjadpour; Maximum a posteriori decoding algorithms for turbo codes. *Proc. SPIE 4045, Digital Wireless Communication II*, 73, July 26, 2000.
- [96]. Seokhyun, Y. and Y. Bar-Ness. "A parallel MAP algorithm for low latency turbo decoding." *IEEE Communications Letters* 6(7): 288-290, 2002.
- [97]. A. Tarable, G. Montorsi, and S. Benedetto, "Mapping of interleaving laws to parallel turbo decoder architectures," in *Proc. 3rd Int. Symp. Turbo Codes and Related Topics*, Brest, France, pp. 153–156, Sep. 2003.
- [98]. O. Y. Takeshita, "On maximum contention-free interleavers and permutation polynomials over integer rings," *IEEE Trans. on Info. Theory*, vol. 52, no. 3, pp. 1249-1253, March 2006.
- [99]. Z. Yan, G. He, We. He, S. Wang, Z. Mao, "High performance parallel turbo decoder with configurable interleaving network for LTE application", *Integration, the VLSI Journal*, Vol. 52, pp. 77-90, Jan. 2016.
- [100]. Butler, J.T., Sasao, T., "Fast hardware computation of $x \bmod z$ ". *Proc. 18th Reconfigurable Architectures Workshop (RAW 2011)*, Anchorage, Alaska, USA, pp. 294–267, May 2011.
- [101]. D. Vogrig, A. Gerosa, A. Neviani, A. Graell I Amat, G. Montorsi, and S. Benedetto, "A 0.35- μm CMOS analog turbo decoder for the 40-bit rate 1/3 UMTS channel code," *IEEE J. Solid-State Circuits*, vol. 40, no. 3, pp. 753–762, 2005.

- [102]. Q. T. Dong, M. Arzel, C. J. Jago, and W. J. Gross, "Stochastic decoding of turbo codes." *IEEE Trans. Signal Processing*, vol. 58, no. 12, pp. 6421–6425, Dec. 2010.
- [103]. R. G. Maunder, "A Fully-Parallel Turbo Decoding Algorithm," *IEEE Trans. Communications*, vol. 63, no. 8, pp. 2762–2775, Aug. 2015.
- [104]. Altera, "Baseband-C-RAN", <https://www.altera.com/solutions/industry/wireless/applications/baseband/c-ran.html>, Jun. 2016.
- [105]. Sadjadpour, H. R., et al. (2001). "Interleaver design for turbo codes." *IEEE Journal on Selected Areas in Communications* 19(5): 831–837.
- [106]. S. Benedetto and G. Montorsi, "Design of parallel concatenated convolutional codes," *IEEE Trans. Commun.*, vol. 44, pp. 591–600, May 1996.
- [107]. O. Y. Takeshita and D. J. Costello, "New deterministic interleaver designs for turbo codes," in *IEEE Transactions on Information Theory*, vol. 46, no. 6, pp. 1988–2006, Sept. 2000. doi: 10.1109/18.868474
- [108]. Jing, S. and O. Y. Takeshita, "Interleavers for turbo codes using permutation polynomials over integer rings." *IEEE Transactions on Information Theory* 51(1): 101–119, 2005
- [109]. P. Robertson , P. Hoeher , E. Villebrun, "Optimal and sub-optimal maximum a posteriori algorithms suitable for turbo decoding." *European Transactions on Telecommunications* 8(2): 119–125, 1997.
- [110]. J. Erfanian, S. Pasupathy and G. Gulak, "Reduced complexity symbol detectors with parallel structure for ISI channels," in *IEEE Transactions on Communications*, vol. 42, no. 234, pp. 1661–1671, Feb/Mar/Apr 1994.
- [111]. S. Dolinar and D. Divsalar, "Weight Distributions for Turbo Codes using Random and Nonrandom Permutations", *Tech. Rep.*, JPL, August 1995.
- [112]. C. Berrou and A. Glavieux, "Near Optimum Error Correcting Coding and Decoding: TurboCodes", *IEEE Transactions on Communications*, vol. 44, no. 10, pp. 1261–1271, Oct 1996.
- [113]. 3GPP, "R1-104661 Time-domain eICIC Solutions from Interference Avoidance Perspective", pp. 23–27, Madrid, Spain, 23rd–27th, Aug. 2010.
- [114]. Saad SA, Ismail M, Nordin R. "Partially compensated power control technique for LTE-A macro-femto networks", In: *Advanced computer and communication engineering technology*. Springer; 2015:331–341.
- [115]. M. Andrews, V. Capdevielle, A. Feki, and P. Gupta, "Autonomous Spectrum Sharing for Mixed LTE Femto and Macro Cells Deployments", in *INFOCOM IEEE Conference on Computer Communications Workshops*, 2010, 2010, pp.1–5.
- [116]. Yi Youwen, Zhang Jin, Zhang Qian, and Jiang Tao, "Spectrum leasing to femto service provider with hybrid access", in *INFOCOM, 2012 Proceedings IEEE*, 2012, pp. 1215–1223.
- [117]. Khan, L. U. "Visible light communication: Applications, architecture, standardization and research challenges." *Digital Communications and Networks*.
- [118]. "[500 Megabits/Second with White LED Light](#)" (Press release). Siemens. Jan 18, 2010.
- [119]. http://www.ted.com/talks/harald_haas_wireless_data_from_every_light_bulb.html
- [120]. <http://www.independent.co.uk/news/science/li-fi-revolution-internet-connections-using-light-bulbs-are-250-times-faster-than-broadband-8909320.html>

- [121]. S. Haruyama Y. Ito and M. Nakagawa. "rate-adaptive transmission on a wavelength dependent channel for underwater wireless communication using visible light leds". 105:127{132, Feb 2006. 5
- [122]. S. Rajagopal, R. D. Roberts and S. K. Lim, "IEEE 802.15.7 visible light communication: modulation schemes and dimming support," in *IEEE Communications Magazine*, vol. 50, no. 3, pp. 72-82, March 2012.
- [123]. H. Burchardt, N. Serafimovski, D. Tsonev, S. Videv and H. Haas, "VLC: Beyond point-to-point communication," in *IEEE Communications Magazine*, vol. 52, no. 7, pp. 98-105, July 2014.
- [124]. J. Vucic, C. Kottke, S. Nerreter, K. Habel, A. Buettner, K.-D.Langer, J.Walewski, 125 Mbit/s over 5 m wireless distance by use of OOK-modulated phosphorescent white LEDs, in: *Proceedings of 35th European Conference on Optical Communication*, pp.1-2, 2009
- [125]. D. Shan Shiu and J. Kahn, "Differential pulse-position modulation for power-efficient optical communication," *IEEE Transactions on Communications*, vol.47, no. 8, pp.1201-1210, Aug. 1999. doi: 10.1109/26.780456.
- [126]. S. H. Lee, S.- Y. Jung, and J. K. Kwon, "Modulation and coding for dimmable visible light communication," *IEEE Communications Magazine*, vol. 53, no. 2, pp.136-143, Feb. 2015. doi: 10.1109/MCOM.2015.7045402.
- [127]. Y. Zeng, R. Green, and M. Leeson, "Multiple pulse amplitude and position modulation for the optical wireless channel," in *Proc.10th Anniversary International Conference on Transparent Optical Networks (ICTON'08)*, vol. 4, Athens, Greece, Jun.22-262008, pp.193-196. doi: 10.1109/ICTON.2008.4598766
- [128]. CIE, "Commission Internationale de l'Eclairage Proc." 1931
- [129]. "IEEE Standard for Local and Metropolitan Area Networks—Part 15.7: Short-Range Wireless Optical Communication Using Visible Light," *IEEE Std 802.15.7-2011*, pp. 1 –309, 6 2011.
- [130]. Dissanayake, S. D. and J. Armstrong (2013). "Comparison of ACO-OFDM, DCO-OFDM and ADO-OFDM in IM/DD Systems." *Journal of Lightwave Technology* 31(7): 1063-1072.
- [131]. D. Tsonev, S. Sinanovic and H. Haas, "Novel Unipolar Orthogonal Frequency Division Multiplexing (U-OFDM) for Optical Wireless," *2012 IEEE 75th Vehicular Technology Conference (VTC Spring)*, Yokohama, 2012, pp. 1-5.
- [132]. O. Kharraz, D. Forsyth, Performance comparisons between PIN and APD photodetectors for use in optical communication systems, *Opt. Int. J. Light Electron Opt.* 124 (13) (2013) 1493–1498.
- [133]. K.-D. Langer and J. Grubor. Recent developments in optical wireless communications using infrared and visible light. In *Transparent Optical Networks, 2007. ICTON '07. 9th International Conference on*, volume 3, pages 146{151, July 2007. doi: 10.1109/ICTON.2007.4296267. 23
- [134]. T. Komine and M. Nakagawa. Fundamental analysis for visible-light communication system using led lights. *Consumer Electronics, IEEE Transactions on*, 50(1):100-107, 2004. ISSN 0098-3063. doi: 10.1109/TCE.2004.1277847. 23, 24, 121
- [135]. J.B. Carruthers and J.M. Kahn. Modeling of nondirected wireless infrared channels. *Communications, IEEE Transactions on*, 45(10):1260{1268, 1997. ISSN 0090-6778. doi: 10.1109/26.634690. 24
- [136]. J. Grubor, S. Randel, K.-D. Langer, and J.W. Walewski. Broadband information broadcasting using led-based interior lighting. *Lightwave Technology, Journal of*, 26 (24):3883{3892, Dec 2008. ISSN 0733-8724. doi: 10.1109/JLT.2008.928525. 17, 21, 23, 25, 37, 60, 121

- [137]. D. O'Brien, "Visible light communications: Challenges and potential," in *Proc. IEEE Photon. Conf.*, 2011, pp. 365–366
- [138]. Q. Wang, Z. Wang and L. Dai, "Multiuser MIMO-OFDM for Visible Light Communications," in *IEEE Photonics Journal*, vol. 7, no. 6, pp. 1-11, Dec. 2015.
- [139]. Zhenhua Y, Robert J. B., G. Tong Z: "Multi-User MISO Broadcasting for Indoor Visible Light Communication," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2013, pp. 4849–4853
- [140]. Hao H, Lutz L., Steve H.: 'Robust MMSE Linear Precoding for Visible Light Communication Broadcasting Systems', *IEEE Globecom Workshops*, Dec 2013, pp. 1081–1086
- [141]. Qi W., Zhaocheng W, Lingdong D.: 'Multiuser MIMO-OFDM for Visible Light Communications', *IEEE Photonics Journal*, 2015, 7, pp. 1-11.
- [142]. Ozgur E., Ergin D., Ozgur B.A.: 'Communicate to illuminate: State-of-the-art and research challenges for visible light communications', *Physical Communication*, 2015, 17, pp. 72-85
- [143]. Harald B., Nikola S., Dobroslav T., Stefan V., Harald H.: 'VLC: Beyond point-to-point communication', *IEEE Comm. Mag.*, 2014, 52, pp. 98-105.
- [144]. Christian B. P., Bertrand M. H., A. Lee S.: 'A vector-perturbation technique for near-capacity multiantenna multiuser communication-part I: channel inversion and regularization', *IEEE Trans. on Comm.*, 2005, 53, pp. 195-202
- [145]. H. Luo, Y. Zhang, L. k. Huang and J. Cosmas, "Low latency turbo decoder implementation for future broadcasting systems," *2017 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, Cagliari, Italy, 2017, pp. 1-4.
- [146]. H. Luo; Y. Zhang; W. Li; L. K. Huang; J. Cosmas; D. Li; C. Maple; X. Zhang, "Low Latency Parallel Turbo Decoding Implementation for Future Terrestrial Broadcasting Systems," in *IEEE Transactions on Broadcasting* , vol.PP, no.99, pp.1-9