Coarsely Quantized Massive MU-MIMO Uplink with Iterative Decision Feedback Receiver

by

Zeyang Zhang

B.Eng, Capital Normal University, 2015

Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

MASTER OF APPLIED SCIENCE

in the Department of Electrical and Computer Engineering

Supervisory Committee

Coarsely Quantized Massive MU-MIMO Uplink with Iterative Decision Feedback Receiver

by

Zeyang Zhang

B.Eng, Capital Normal University, 2015

Supervisory Committee

Dr. Michael L. McGuire

Department of Electrical and Computer Engineering

Supervisor

Dr. T. Aaron Gulliver

Department of Electrical and Computer Engineering

Department Member

Abstract

Supervisory Committee

Dr. Michael L. McGuire

Department of Electrical and Computer Engineering

Supervisor

Dr. T. Aaron Gulliver

Department of Electrical and Computer Engineering

Department Member

Massive MU-MIMO (Multiuser-Multiple Input and Multple Output) is a promising technology for 5G wireless communications because of its spectrum and energy efficiency. To combat the distortion from multipath fading channel, the acquisition of channel state information is essential, which generally requires the training signal that lowers the data rate. In addition, coarse quantization can reduce the high computational energy and cost, yet results in the loss of information.

In this thesis, an iterative decision feedback receiver, including iterative Channel Estimation (CE) and equalization, is constructed for a Massive MU-MIMO uplink system. The impact of multipath distortion and coarse quantization can be gradually reduced due to the iterative structure that exploits extrinsic feedback to improve the CE and data detection, so that the data rate is improved by reducing training signals for CE and by using low precision quantization. To observe and evaluate the convergence behaviour, an Extrinsic Information Transfer (EXIT) chart method is utilized to visualize the performance of the iterative receiver.

Index Terms *- Massive MIMO, coarse quantization, iterative decision feedback, Channel Estimation, Zero-Forcing equalization, MMSE equalization, EXIT chart*

Table of Contents

List of Tables

List of Figures

List of Acronyms

SISO Soft-input and Soft-Output

AWGN Additive White Gaussian Noise

ADC/DAC Analog-to-Digital/Digital-to-Analog Converter

BER Bit-Error-Rate

CE Channel Estimation

CSI Channel State Information

CIR Channel Impulse Response

CP Cyclic Prefix

DFT Discrete Fourier Transform

ECC Error-Correcting Code/Error Control Coding

EXIT Chart Extrinsic Information Transfer Chart

FIR Finite Impulse Response

FFT/IFFT Fast Fourier Transform/Inverse Fast Fourier Transform

IDFE Iterative Decision Feedback Equalization/Equalizer

ISI Inter-Symbol Interference

LTE Long Term Evolution

LLR Log-Likelihood Ratio

MIMO Multiple-input and Multiple-output

MU-MIMO Multiple-user Multiple-input and Multiple-output

MAP Maximum A Posteriori

MMSE Minimum Mean Square Error

MFC Multipath Fading Channel

(NR)LDPC (New Radio) Low Density Parity Check

(O)FDM (Orthogonal) Frequency-Division Multiplexing

PAPR Peak-to-Average Power Ratio

PSK Phase-Shift Keying

QAM Quadrature Amplitude Modulation

RF Radio Frequency

SC-TDE/FDE Single-Carrier Time/Frequency Domain Equalization

SVD Singular Value Decomposition

TDD Time-Division Duplexing

ZF Zero-Forcing

Acknowledgements

First and foremost, I would like to express the most sincere gratitude to Dr. Michael L. McGuire - my supportive supervisor, the esteemed Chair of Department of Electrical and Computer Engineering, whose guidance and patience throughout my Master's program. As well as Dr. Aaron Gulliver, who provides significant guidance and knowledge to dispel my confusion.

Last but not least, I also appreciate the ATS group in Thales Canada, especially Josephine Sung, Alex Babut, Daniel Vijayakumar and Aditya Chandramouli (in no particular order). Their dedicated guidance inspired my passion for the engineering industry.

Dedication

*To my mother -**Lan Yang**,*

an ordinary and hardworking woman who devotes everything to her child.

For her parenting, endless love and support.

*To my uncle - **Fan Mo**,*

an entrepreneur, book enthusiast, educator and most importantly, a guide of me.

For his deep faith in me and encouragement of me studying abroad.

*To my grandparents - **Yuancheng Yang**, **Guiying Zhang**,*

a selfless and honored couple who raised me the best way they know.

For their honest love and support.

1 Introduction

1.1 Motivation and Objective

Wireless communication is a communication method that exchanges information using electromagnetic wave signals that propagate in free space. Nowadays, wireless communication is widely used in mobile devices, Wi-Fi networks, satellite televisions and Global Positioning Systems (GPS) as an indispensable part of sharing information and communication on a daily basis. To meet the exponentially increasing demand for radio data rate, new methodologies which enable higher capacity radio are always needed.

This thesis focuses on iterative Channel Estimation (CE) and equalization which join together in an iterative decision feedback receiver for a Massive Multiple-input and Multiple-output (MIMO) system in the Uplink Channels (UCs). Less training signals for CE are required with an iterative decision feedback receiver because correct data detection/decoding decisions from previous iterations are as additional training signals. However, care must be taken in iterative receivers to ensure that incorrect decisions do not reinforce themselves leading to instability. Iterative receivers can also help to mitigate the effect of coarse quantization created by the use of low-resolution Analog-to-Digital Converters (ADCs) in the receiver. To observe and analyze the iterative behavior, an Extrinsic Information Transfer (EXIT) chart method is used, from which we can substitute variables to predict and evaluate different setups of a given system.

1.2 Background

The first radios used only a single antenna at both the transmitting and receiving end in a configuration known as a single-input and single-output antenna system. There are multiple propagation paths from the transmitting antenna to the receiving antenna. In the real physical world, temperature, humidity [1] and geological altitude [2] combine with multipath propagation to create random channel gains. Channel noise is introduced by the thermal noise in the radio receiver [3]. The received signal, as described in Fig. 1.1, is the sum of the channel noise with transmitted signal after it has traveled over multiple propagation paths with different gains, which leads to distortions in terms of Inter-Symbol Interference (ISI). This introduces errors in the decision process at the radio receiver. The process of estimating the transmitted signal at the radio receiver is referred to as equalization. To make a proper estimation of the transmitted signal, estimation of the radio multipath channel is needed, which is referred to as CE. By knowing the channel, equalization reverses the distortions of multipath propagation and mitigates the impact of ISI. Accurate equalization requires an accurate estimation of the multipath channel parameters. [4].

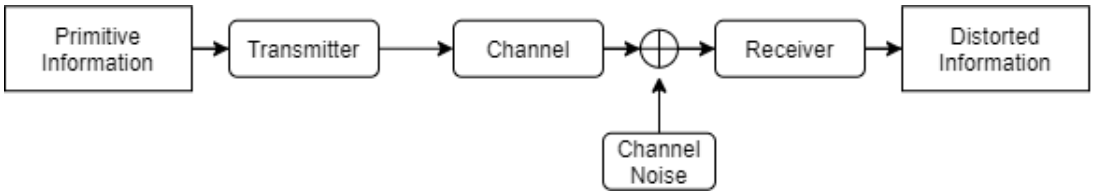

Figure 1.1: Basic communication block diagram

The radio channels that the users use to communicate to the Base Station (BS) are known as the Uplink Channels (UCs). The radio channels over which the BS sends data signals to users are defined as the Downlink Channels (DCs). In Time-Division Duplexing (TDD), a user transmits signal via the UCs to the BS at one time slot, and the BS transmits down to the user at the same frequency in a later slot. If the channel conditions do not change between the uplink and downlink time slots, the radio channel has reciprocity, which means the UCs are equal to the DCs [5]. If the

channel is reciprocal, then the BS can pre-distort the signals it sends on the DCs, based on the results from channel state estimation during uplink communication, so that the user terminals have simplified channel estimation and equalization. [6]. TDD has been the leading technology in cellular communications standards since the 1990s [7].

Aside from the single-input and single-output antenna systems from the first radios, modern radios are classified as single-input and multiple-output (SIMO) antenna systems, multiple-input and single-output (MISO) antenna systems and multiple-input and multiple-output (MIMO) antenna systems.

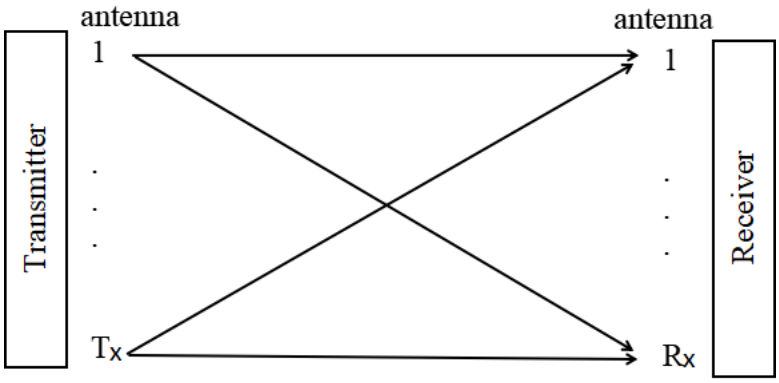

Figure 1.2: Wireless communication model

MIMO has been widely applied to essential standards of wireless communication, by multiplying the capacity of a single-input and single-out radio link using multiple transmitting and receiving antennas in order to counteract the multipath propagation [8].

In the 1990s, Point-to-Point MIMO, where an active user possesses multiple transmitting antennas, was the main topic in MIMO research [9]. However, the design of multi-antenna terminals (mobile devices) is complicated, and it is also limited by the required distance between antennas on a mobile device for the channels for each antenna to be independent. For the channel gains from a transmitter to two antennas at the receiver to be independent, a spacing of at least $\frac{1}{2}$ of a wavelength is

required [10]. The most common frequency band that is currently being used in the North American cellular network is $\sim$ 900 MHz [11], resulting in a wavelength $\sim$ 33 centimeters. For signals of this frequency, antenna separations of $\frac{33}{2} = 16.5$ cm are required between two antennas for those two antennas to have independent channel gains. Such separation is difficult to achieve at portable devices' dimensions, which is why this thesis will focus on single antenna user terminals. In millimeter wave (mmWave) 5G cellular systems, terminals with much higher carrier frequencies and thus proportionally shorter radio wavelengths are possible so multiple antenna user terminals will be explored in future work[12].

In the 2000s, researchers started to seriously investigate MU-MIMO (Multi-user MIMO). MU-MIMO describes a set of single-antenna user terminals that communicate with a BS equipped with multiple receiving antennas. In Fig. 1.3, each user possesses one single-antenna mobile device. In order to communicate with other users, the information will be sent through the UCs to the designated BS, then the BS will beamform to destination on the DCs. To detect data transmission over a MIMO channel, equalization in the UCs requires more mathematical operations because of the multiple receiving antennas. However, the BS is able to better interpret the received signal from the users' terminals because more independent information about each user's signal is available [13]. With the multiple measurements of the transmitters' signals with independent distortion and noise, a more accurate estimation of the channel leads to better performance of the equalization, which generally leads to a higher capacity [14].

Adding more receiving antennas actually allows the BS to get better estimates of the transmitted signals [7], but having a high-precision receiver for every antenna at the BS is very expensive [9]. To keep the benefits of multiple antennas without increasing the overall receiver's cost, sub-optimal CE and equalization calculations are used which have been shown to have results almost as good as the optimal solution, so long as a large number of measurements are taken [7][15]. In this thesis, we discuss the use of Massive MU-MIMO (also known as Large-Scale MIMO or Large-Scale

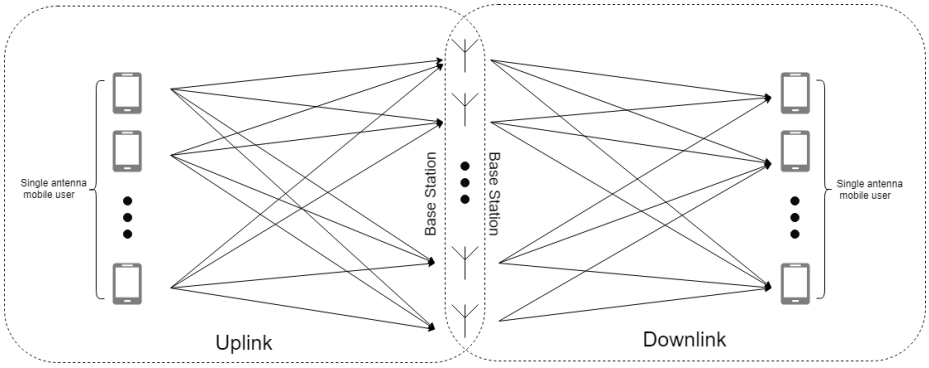

Figure 1.3: MU-MIMO uplink and downlink model

Antenna Systems) where single-antenna users communicate with a BS equipped with a large number of antennas (where the number of antennas exceeds the number of user terminals). Robustness has also become a perk of Massive MU-MIMO systems comparing to the conventional MU-MIMO, because the failure of one or a few of the antenna units has only a small effect on the radio link quality [16]. As a result, Massive MU-MIMO has become one of the key technologies in the new era of wireless communication - 5G.

To make Massive MIMO systems financially acceptable for commercial implementations, cheap individual receivers are needed, which can be achieved by lowering the precision of measurement at each antenna's receiver. Modern radio systems require Analog-to-Digital Converters (ADCs) to convert the measured radio signals to numerical form for the digital recevires, and Digital-to-Analog Converters (DACs) to convert the samples of the transmitted signal generated by the digital transmitter to a continuous time radio frequency signal which can be sent over the radio carrier.

The operation of an ADC is modeled as a two-phase process, sampling and quantization. Continuous analog signals in the form of time-varying sources of voltage are sampled at a designated sampling frequency (or sampling rate) into discrete-time signals. Mathematically, quantization takes each real number and maps to a finite set which is decided by the precision of quantization, as in word-length or bits, output discrete values are quantized signals. Commonly used in audio processing, 8-bit,

16-bit and 24-bits quantization provided satisfying results without much loss of information. Typically, high-precision quantization for radio communication is 8 to 12 bits [17], while low-precision, also referred to as coarse quantization, is 1 to 4 bits. It has been well proven that the higher precision we choose, the less quantization distortion the signal has [3]. However, this is a trade-off. Higher precision ADC requires more computational cost and more sophisticated hardware, which leads to a much more expensive receiver which requires a more complicated circuit design to perform higher precision quantization. For Massive MIMO systems, a well-functioning antenna-array with higher-order quantization Radio Frequency (RF) antennas will be financially challenging.

Overall, people always want to keep communication devices portable and cheap. The cost of the RF front-end devices for each antenna and the cost of the computational hardware in the receivers are the two major components in the overall cost of a radio receiver. At the present time, when users demand video streaming and high-quality real-time interaction, data rate and accuracy have arisen along with the modern radio communication technology, demand for low-error high data rate communications has grown which makes it necessary to develop low power radio receivers because of the limited development of chemical battery technology. Therefore, expensive operations such as high quality RF front end, computationally expensive operations such as CE and equalization, are chosen to be done at the BS which has access to wired power as opposed to the user terminals which have limited battery power supplies.

1.3 Literature Review

As a consequence of limited favorable radio spectrum and data traffic growth, the employment of the Massive MIMO transmission techniques have led the 5G communication systems to achieve high spectral-efficiency [18][16][19][20] and high energy-efficiency [16][21][22].

CE has always been an important topic in digital radio communications research. Kalman filters were applied for CE and data detection to combat the effects of the multipath fading channel in [23] which inspired us to implement a joint CE and equalization on a Massive MU-MIMO system. In [23][24][25][26] and [27], various training-based CE and optimization techniques were used to estimate Channel State Information (CSI). Using the known pilot signal transmitted from a user, also known as a training signal, and the resulting received noisy copy of the pilot signal at the receiver, the BS can obtain the information about the CSI which is represented by estimates of the channel coefficients. Since bandwidth is a limited resource, there is a great deal of work on reducing the required length of pilot sequences for CSI estimation while maintaining an acceptable degree of accuracy. Semi-blind CE, where the number of pilot measurements is less than the number of channel parameters to be estimated, and blind CE, where no pilot signals are used, have been of great interest to the research community. Unfortunately, the higher computational cost of semi-blind and blind CE has prevented these techniques from being deployed in commercial systems [28]. For the Massive MIMO systems, due to the large scale antenna array, more CSI needs to be estimated by high computational complexity algorithms, therefore blind and semi-blind CE which requires much higher computational cost to achieve the same performance as compared to conventional pilot-aided CE in Massive MIMO systems [29].

Given the estimated CSI, linear and non-linear equalization can be used to recover transmitted symbols. A non-linear equalizer processes the received signals with a non-linear filter, such as Least-Squared-Error equalization [30]. However, non-linear equalizers often require higher computational complexity [31]. V-BLAST is a simple non-linear equalization algorithm for multi-user detection over radio channels which initially estimates the most powerful users' signals and then removes their interference before estimating the less powerful users' signals [32]. However, if all users' signals are near equal power, the algorithm can fail [32]. Thus, linear equalization is taken into consideration of this thesis. In [33], a priori information from the decoder and CE

error were considered into a Linear Minimum Mean Squared Error (LMMSE) based iterative equalizer for a standard MIMO system. However, the prior information may contains incorrect decisions, feedback which can cause error enhancement and lead to instability [34].

In 1995, inspired by an iterative decoder used in so-called turbo codes, a model for turbo equalization with the exchange of soft decision was given in [35], where the extrinsic soft decisions were extracted from the detection and decoding steps and then used at the following iteration. This innovative system presented a new method to counteract the ISI. Based on which, the first Iterative Decision Feedback Equalizer (IDFE or IDE, also known as Turbo Equalization) for quantized Point-to-Point MIMO system was introduced in [36]. A LMMSE equalizer was implemented in [36] for iterative equalization. They modeled a receiver with IDFE operating on data with general quantization precision. In the end, their research shows that an IDFE in Point-to-Point MIMO gives a better performance in terms of BER. The iterative equalization in this thesis, based on a linear equalizer, will be introduced in Chapter 3.

To construct a BS with a large-scale antenna array, low-cost and power-efficient antennas have become more necessary. A quantized Massive MU-MIMO uplink systems was explored in [37], they employed the pilot-aided Maximum A-Posteriori (MAP) CE and LMMSE-based iterative equalization to study the associated performance and quantization precision trade-offs. The conclusion from [37] showed that 4 to 6 bits, depending on the ratio between the number of BS antennas and the number of users entails no obvious performance loss in the Massive MU-MIMO systems comparing to infinite-precision quantization. While in our thesis, a LMMSE and extrinsic soft-decision based iterative CE was applied. With iterative equalization, optimal quantized bit lowers down to 4 or even 3 bits with no additional cost in terms of transmitting energy and baseband processing complexity.

1.4 Contributions

The main contributions of this thesis are:

- We apply an Iterative Decision Feedback module on both detection and CE for the Massive MU-MIMO uplink systems then analyze the improvement. comparing to conventional pilot-aided CE. The performance of an iterative receiver which mitigates both the multipath propagation as well as estimating the CSI for short pilot-sequences is presented.
- We simulate then evaluate the performance of the iterative receiver in Massive MU-MIMO uplink systems from 1- to 4-bit coarse quantizations.
- The 5G standard LDPC code is applied, as well as a low constraint length convolutional code, so that we showed that the system we are presenting with is not restricted to any certain channel coding scheme. In fact, given any ECC and its corresponding EXIT chart, we can predict and analyze its performance when used in Massive MU-MIMO uplink systems.

1.5 Outline

In Chapter 1, starting with a brief motivation and objections, the background of wireless communication with regard to CE, equalization and quantization in the MIMO systems are introduced. In Chapter 2, we build a standard mathematical model for a Massive MU-MIMO uplink system. In Chapter 3, an iterative decision feedback receiver, as in iterative CE and equalization, is defined. Simulations are performed and analyzed in Chapter 4. Lastly, in Chapter 5, we conclude the entire thesis and set up future work.

2 Massive MU-MIMO System Uplink Model

This chapter introduces the Massive MU-MIMO system uplink model on multipath fading channels. Section 2.1 defines a mathematical model of multipath fading channels. Section 2.2 describes and illustrates several key steps of constructing a transmitted signal from bit sequences generated by a user. Section 2.3 describes the essential technologies of processing the received signal at the BS.

2.1 Multipath Fading Channel

In wireless communication, radiated signals from transmitters take different paths and arrive at the destination at different times. On all these paths, distances and obstacles may vary. Each path of multipath fading channels is characterized by two aspects - the delay caused by propagation and the attenuation from the scattering.

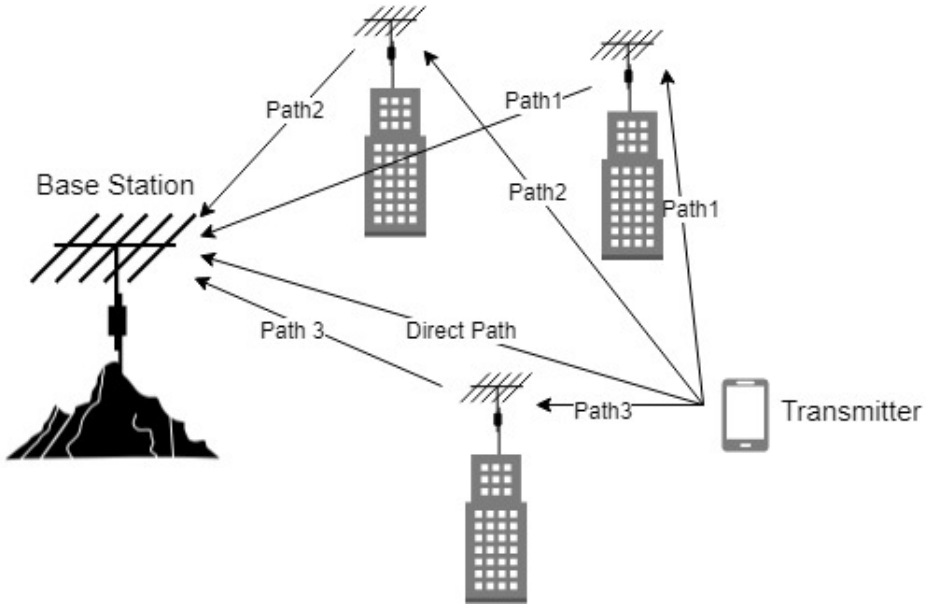

Figure 2.1: Multipath propagation

We assume, on the l^{th} path, the complex attenuation a_l is a complex Gaussian random variable with independent and identically distributed (i.i.d) real and imaginary components. Each path for a time sample t has a path impulse response $a_l\delta(t-\tau_l)$,

where the $\delta(t-\tau_l)$ represents the shifted delta function at a delay τ_l. The multipath response for a single transmitter to a receiving antenna pair is the summation of individual paths:

$$h(t) = \sum_{l=1}^{L} a_l \delta(t-\tau_l) \tag{1}$$

The multipath fading Channel Impulse Response (CIR) $h(t)$ varies depending on the various channel attenuation and delay factors. When several paths with small relative shifts add together, the summed signal can have a magnitude greater than the magnitude of any individual component, and we say that constructive interference arises. Destructive interference occurs when the magnitude of the summation of signals is less than the magnitude of each component (often referred to as a null or deep fade scenario). To receive data signals with low error rate, a strong enough signal is needed for proper reception. In addition, the signals from different transmitters must be resolvable from each other. To counteract destructive interference, the diversity of a multiple resolvable paths to the receiver antenna system is crucial. Multiple links between transmitters and receivers will dramatically degrade the probability of communication disruption caused by deep fade due to the independence of the gains for all of the multiple paths. The chance that all independent channels from an active user to the BS are in deep fade becomes unlikely as the number of receivers increases [18][20].

For the work in this thesis, it is assumed that there is a multipath fading channel from each transmitter to each antenna with L paths with the average power at each path being $1/L$, i.e there is no dominant path. The multipath fading channel for a given pair of transmitter and receiver antennas is modeled as a linear Finite Impulse Response (FIR) filter. Let $x_{tx}[n]$ denote the n^{th} sample sent by a transmitter tx through the channel, $y_{rx}[n]$ denote the n^{th} sample received at a receiver rx, we have:

$$y_{rx}[n] = \sum_{tx=1}^{Tx} \sum_{l=0}^{L-1} h_{rx,tx}[l] x_{tx}[n-l] + v_{rx}[n] \tag{2}$$

where

$$v_{rx}[n] \sim \mathcal{N}(0,\, {\sigma_v}^2)$$

denotes complex Additive White Gaussian Noise (AWGN) with zero mean and variance ${\sigma_v}^2$. $h_{rx,tx}[l]$ denotes the sampled CIR on the path with a delay of l samples between a transmitting antenna tx and a receiving antenna rx.

2.2 User Terminal (Transmitter)

In a digital cellular network, every mobile device is equipped with a low power transceiver (transmitter and receiver). In a wireless uplink system, a transmitting antenna of a mobile user sends signals to the BS. In this section, we focus on how the transmitted signal is generated at the user terminal.

2.2.1 Error-Correcting Code (ECC)

An Error-Correcting Code, or Error Correction Code (ECC) technique is required to detect/correct random errors generated from noise. A functional ECC enables a system to reach a high degree of reliability, so that the effect from the presence of noise can be mitigated [38]. In addition to the transmitted original data bits from users, some additional redundant check bits are also attached or inserted or embedded into the data sequences [38].

Suppose Tx users send information as transmitters. For a transmitting antenna/user tx, we have raw binary data $\mathbf{x}_{raw,tx}$. To encoded the raw binary data string:

$$\mathbf{x}_{coded,tx} = \mathbf{C}\{\mathbf{x}_{raw,tx}\} \tag{3}$$

where $\mathbf{x}_{raw,tx}$ denotes raw data from user tx, and the $\mathbf{C}\{\cdot\}$ operator represents the encoding function. In the encoding function, a length D_{raw} string of bits generates a length D_{coded} ($D_{coded} > D_{raw}$) string with redundancy which the receiver can use to

detect/correct errors. In Fig. 2.2, the raw data sequence $\mathbf{x}_{\text{raw,tx}}$ from a user is divided into several blocks [$\mathbf{x}_{\text{blk\#1}}$, $\mathbf{x}_{\text{blk\#2}}$ ...], the encoder applies ECC to each of the blocks. Lastly, to match the code rate, a certain bits who contain relatively low information are removed as a process known as puncturing, these punctured bits are assumed to be zeros in the decoding algorithm in the receiver.

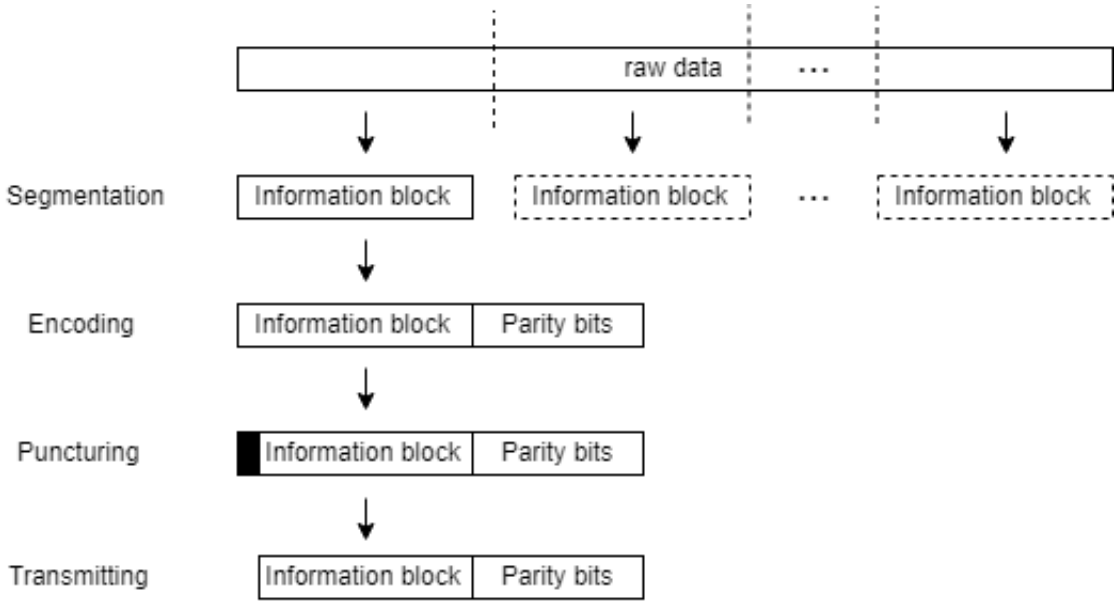

Figure 2.2: Encoding Process

A fading channel can create error bursts where there are multiple bit errors all located close together. These errors can overwhelm the ability of the ECC, if the number of error bits in a given region is greater than the number of bits the code can correct. Such burst errors occur in a contiguous manner in wireless channels [39], this may cause the failure of ECCs, because ECCs encodes raw data bits contiguously, a serial encoded bits missing or inaccuracy can cause trouble in retrieving the original data bits. To counteract this negative effect, many communications systems use an interleaver where the data bits are reordered, so the bits from several error correction regions are uniformly spread over a long time period. This causes a bit errors from a short error burst to be spread over several codewords with a low number of errors in each codeword, so that the correction ability of the ECC is less likely to be over-

whelmed, resulting in an improvement in performance versus when an interleaver is not used.

At receivers, decoder output depends on the range of input samples. These dependencies result in correlations between feedback decisions and input samples, which lead to dependencies between feedback decisions [40]. Feedback decisions are updated by a soft-input/soft-output decoder in the iterative system. These decisions are then converted to the soft symbol values by taking a statistical expectation with the assumption of feedback decisions being independent of each other [41]. To ensure that feedback decisions being as independent as possible of each other, interleaving at receivers is a necessary step in our system. At transmitters, an interleaver is needed as well so that the orders of information bits can be consistent at transmitters and receivers. We here define an interleaver function $\Pi\{\cdot\}$ as follow:

$$\mathbf{x}_{\pi,tx} = \Pi\{\mathbf{x}_{coded,tx}, st\} \tag{4}$$

where st is a random permutation vector, known to the receiver, with a length equal to the length of the transmitted bit sequence, so that input sequence $x_{coded,tx}$ reordered into $x_{\pi,tx}$ as described in the permutation vector st.

This thesis examines the use of both a non-systematic convolutional code [42] and a New Radio Low Density Parity Check block code from 5G standard (5G NR-LDPC) [43]. Both ECCs correct errors by maintaining parity bits for a selection of the data bits. The parity bits describe the linear relations that a corresponding codeword must satisfy, based on which a parity-check matrix can be formed [38]. Several other channel coding options for the Massive MIMO systems are discussed and compared in [44].

• Convolutional code

The ECC from the IEEE 802.11 standard is a non-systematic convolutional code [45] with code rate $R_c = \frac{1}{2}$, is applied in this thesis. It has reasonable performance at a reasonable decoding computational cost. To encode a data convolutionally, K

memory registers are needed, each holding one input bit for K-state for the generator polynomial.

The encoder takes K bits of raw data into combinations of Boolean XOR operations for each output coded bit. To decode a convolutional code, the Viterbi algorithm is commonly used for both hard-decision and soft-decision when the memory of the previous register states is short, i.e low constraint. The Viterbi decoder provides the Maximum Likelihood (ML) [42] source bit string, as a single-path algorithm that returns the closest codeword to the received bit string. On the other hand, an algorithm that invented by Bahl, Cocke, Jelinek and Raviv (BCJR algorithm) [46] gives Maximum A Posteriori (MAP) estimate of the transmitted sequence, which is an iterative decoder that makes decisions to minimize Bit-Error-Rate (BER). The maximum posteriori probability can be used as the soft decision in the iterative decision feedback system. Therefore, to obtain such soft decision, a BCJR decoder is used in this thesis.

• **LDPC code**

Robert G. Gallager invented the LDPC ECC concept in 1962. LDPC codes are designed to support high throughput [43]. A LDPC code is defined by its parity-check matrix, which is mostly filled with 0s with some 1s such that it can be said that the matrix has low density. Generally, the parity check matrix is defined so that no two rows or columns have more than a single 1 in common [39]. The challenge was the design of the codes such that encoding and decoding algorithms can recover the original codeword in heavy noise. However, new analytic methods [47] and evolution of computation devices make it possible to solve the design problem, therefore LDPC codes now have been rediscovered its potentials due to its capacity-achieving channel coding schemes [48]. LDPC codes have been adopted into the 5G cellular-radio standard. We simulate and analyze the performance for this popular 5G standard code with a Massive MIMO antenna arrangement and iterative receivers.

2.2.2 Modulation

Modulation is defined as the process by which some characteristic of a sinusoidal waveform is varied in accordance with a modulating waveform [49]. A modulator encodes M-ary binary information into a carrier wave by varying the waveform's properties. A carrier wave is normally a sinusoidal with a constant amplitude and frequency. There are four popular modulation techniques for radio data transmission. They are known as Amplitude-Shift Keying (ASK), Frequency-Shift Keying (FSK), Phase-Shift Keying (PSK) and Quadrature Amplitude Modulation (QAM) [50]. The alphabet of the output signals consists of $M = 2^b$ symbols, each symbol contains b bits information. A symbol rate of r_s symbols per second results in a data rate (or bit rate) of $r_S b$ bits per second.

Signals are in sinusoidal form in wireless communication, a sinusoid can be decomposed into two amplitude-modulated sinusoids - In-phase (I) and Quadrature (Q) components - with the same frequency and a relative phase shift of 90 degrees [51].

To clearly represent modulation schemes, a constellation diagram is often to be used, where shows the output symbols on a complex axis as in Fig. 2.3. The real and imaginary axes are the (I) and (Q) components respectively.

Input bit pair	Phase (radians)	I	Q	Symbol value
0, 0	$\pi/4$	0	0	1 + j
1, 0	$3\pi/4$	1	0	-1 + j
1, 1	$5\pi/4$	1	1	-1 - j
0, 1	$7\pi/4$	0	1	1 - j

Table 2.1: QPSK/4-QAM bit to symbol mapping

The Massive MIMO of 5G standard supports modulation schemes from 4-QAM/QPSK up to 1024-QAM [52]. Simulations throughout this thesis are conducted according to $M = 4$, 4-QAM (4-QAM and QPSK share identical constellations, as in Fig. 2.4).

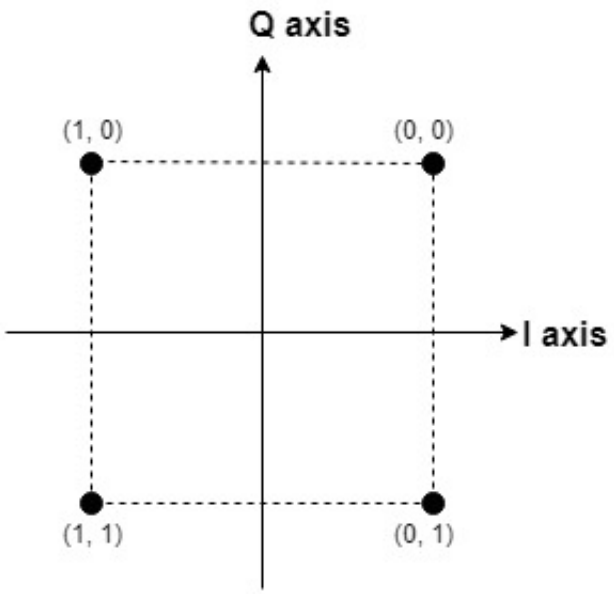

Figure 2.3: QPSK/4-QAM constellation

Table 2.1 indicates the relation between input bits and output symbols for $M = 4$. At a time point t, with T_s symbol duration, the continuous In-phase component $I(t)$ and Quadrature component $Q(t)$ are defined as:

$$I(t) = \sum_{n=-\infty}^{\infty} p(t - nT_s)I[n] \tag{5}$$

$$Q(t) = \sum_{n=-\infty}^{\infty} p(t - nT_s)Q[n] \tag{6}$$

where p denotes the pulse shape function, $I[n]$ and $Q[n]$ as the n^{th} consecutive bit pair from a bit sequence for In-phase and Quadrature component respectively. Mathematically, the signal from a transmitting antenna tx passes through a modulator can be modeled as:

$$\mathbf{s}_s(t)_{tx} = \cos(2\pi f_c t)I(t) + \sin(2\pi f_c t)Q(t) \tag{7}$$

where f_c is the carrier frequency. In the real world, this continuous time signal will be sent as the transmitted signal at the RF antenna. Generally, if the baseband samples of the signal pass through the modulator, it can be abstracted as the function: $M\{x_{\pi,tx}, S_{mod}\}$:

$$\mathbf{s}_{tx} = \mathbf{M}\{\mathbf{x}_{\pi,tx}, S_{mod}\} \tag{8}$$

where $\mathbf{M}\{x_{\pi,tx}, S_{mod}\}$ takes the input bit sequence $\mathbf{x}_{\pi,tx}$ and generates the sequence of modulated symbols $\mathbf{s}_{tx}$ where S_{mod} is the vector of complex modulation constellation points of length M. Each sub-sequence of $\log_2 M$ bits from $\mathbf{x}_{\pi,tx}$ is used to generate an index to find the corresponding modulated symbol from S_{mod} which is placed into the appropriate place in $\mathbf{s}_{tx}$.

2.2.3 OFDM and Single Carrier

The modulated symbols are allocated in frequency/frequencies within a channel bandwidth to be sent through a wireless channel. Orthogonal Frequency-Division Multiplexing (OFDM) forms the basis of for Fourth-Generation (4G) wireless communication systems. It separates the channel bandwidth into several narrow-band frequency bins, where each narrow-band carries a modulated signal. However, unlike traditional Frequency-Division Multiplexing (FDM) which uses guard frequency bands to prevent interference between sub-carriers and wasting scarce wireless spectrum between sub-channels [53], OFDM uses sub-carriers that are all orthogonal to each other, which avoids interference between sub-carriers with sub-channel signals having overlapping yet non-interfering signals [53]. OFDM symbols are constructed in the frequency domain by mapping the input bits on the (I) and (Q) components of the modulated symbols, which is done in the modulation block. In fact, Inverse Fast Fourier Transform (IFFT) and Fast Fourier Transform (FFT) [54] blocks in the transmitter are interchangeable as long as their duals are used in the receiver. The size of IFFT, N, indicates the number of sub-carriers. A Serial-to-Parallel (S/P) module is required to convert serial data sequences to N parallel streams for convenience, so that each multi-carrier is mapped to their corresponding stream at the designated frequency.

Given the symbol rate f_s, the symbol transmission rate per carrier is f_s/N, where N denotes the number of carriers in a OFDM system. Comparing to a conventional Single carrier with Time Domain Equalization (SC-TDE), OFDM processing only requires $\log_2 N$ multiplications per data symbol [55], because the system uses the

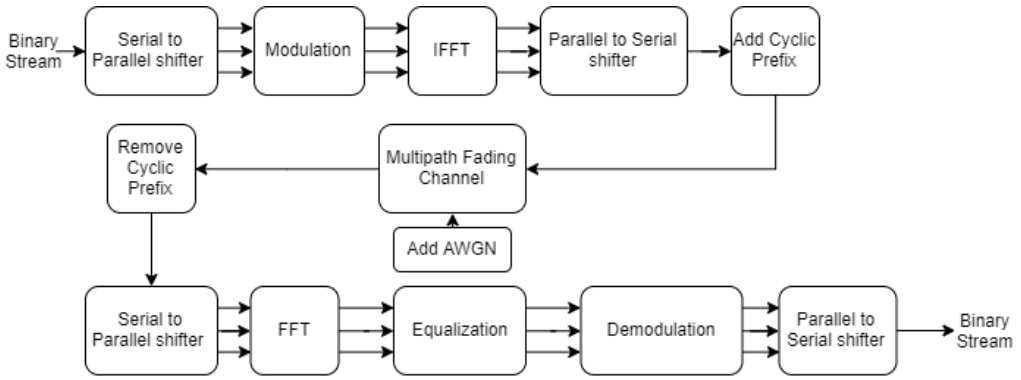

Figure 2.4: OFDM block diagram

efficient FFT algorithm which only demands on the order of $N \log_2 N$ computations. While SC-TDE requires multiplications per symbol that is proportional to the number of data symbols spanned by the multipath [56] because finding the coefficient values of the TDE filter is expensive, even though the equalization is linear with respect to the length of the CIR.

Despite these benefits, OFDM has a high peak-to-average power ratio (PAPR) in the time domain [57]. Since many sub-carrier components are added together in the IFFT operation, when all the components achieve the maximum value simultaneously, the summation in the time domain creates a high peak. Due to the large number of independently modulated sub-carriers in the OFDM system, the peak can be very high compared to the average output value of the system [57]. This side-effect makes transmitter power amplifier design challenging [58], particularly for mobile devices.

On the other hand, a Single Carrier (SC) system transmits the data signal on a single carrier. SC transmission only uses one RF carrier (frequency band) to carry the information. It has been the traditional modulation scheme in digital communications. Frequency domain linear equalization in a SC system (SC-FDE) is simply the frequency domain analog of what is done by a conventional linear TDE [56]. SC-FDE has been proven to give a similar performance to the conventional OFDM system [56][59]. It even gives slightly better performance than OFDM for low constraint length convolutional codes in [60]. SC presents a low PAPR comparing to the OFDM

systems [61], which results in a simpler RF system and better energy efficiency [62]. All of the SC symbols' signals are spread over all sub-carriers, in both the ones in deep fade and less-faded channels, so that information loss due to the deep fade is unlikely [63].

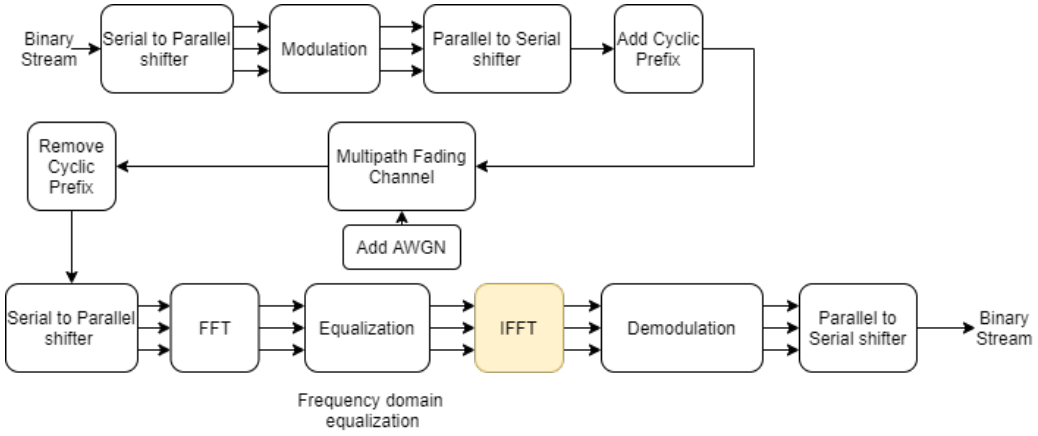

Figure 2.5: Single carrier block diagram

Traditionally, to compensate ISI from previous OFDM or SC symbol blocks due to multipath propagation, a Cyclic Prefix (CP) with length L_{cp} greater than the maximum channel propagation lag of L samples must be transmitted at the beginning of each symbol block. The CP is a repetition of the last data symbols in a block. Essentially, it makes the linear convolution of the CIR with the transmitted signal for the current OFDM or SC block equivalent to a circular convolution. This allows the effect of multipath propagation in the frequency domain to be modeled as product of the FFT of the transmitted signal with the FFT of the CIR, so that the effect of time domain multipath propagation is modelled as a single complex gain applied to each frequency bin of the transmitted signal. This greatly reduces the cost of equalization which is one of the main reasons for the popularity of OFDM and SC modulation in wireless systems.

With CP, for a user tx, we have transmitted symbols:

$$\mathbf{x}_{tx} = \begin{bmatrix} s_{1,1} & s_{1,2} & \cdots & s_{1,N+L_{cp}} \\ s_{2,1} & s_{2,2} & \cdots & s_{2,N+L_{cp}} \\ \vdots & \vdots & \ddots & \vdots \\ s_{B+P,1} & s_{B+P,2} & \cdots & s_{B+P,N+L_{cp}} \end{bmatrix} \tag{9}$$

where B denotes the number of data blocks, $s_{b,n}$ represents modulated n_{th} symbol at a b^{th} block (P denotes the pilot block, which will be introduced in Section 2.3.2).

2.3 Base Station (Receiver)

A Base Station (BS) refers to a wireless base station installed at a fixed location. The BS plays an important role in modern wireless communication, its transceivers allow it to serve both in uplink systems as a receiver and in downlink systems as a transmitter.

As the receiver throughout this thesis, a well-functioning BS that supports Massive MIMO is expected to achieve several goals. First of all, a massive receiving antenna-array is required to receive analog signals, which pass through ADC and are quantized to digital signals. Secondly, since signals are transmitted through unknown multipath fading channels, in order to estimate the transmitted signals, CE is required, based on which, an equalizer processes the signals to estimate the transmitted signals then de-map it into coded bit streams. Lastly a deinterleaving module and a decoder are also necessary as the last steps of reordering then correcting the data bits.

Within this thesis, it is assumed that the UCs between Tx single-antenna users and Rx antennas BS remains constant over the period of time of one data block is being transmitted, so that they can be modeled as linear time-invariant systems in the CE and equalization [64].

2.3.1 Coarse Quantization

In modern communications systems, the analog signal received on the antennas must be converted to a digital signal for processing in the receiver by an ADC module. An ADC contains two main processes: sampling and quantization. Continuous analog signals are sampled at the sampling rate of f_s samples per second, equivalently, a sampling period $T_s = \frac{1}{f_s}$. Let $\mathbf{y}_{rx}$ denote the received continuous signal, the sampled signal $\mathbf{y}_{s,rx}$ at a receiver rx can be expressed as:

$$\mathbf{y}_{s,rx}[n] = \mathbf{y}_{rx}(nT_s) \tag{10}$$

With a sufficiently high sampling rate, the original signals can be recovered by assembling samples via ideal low pass filtering with a bandwidth of $f_s/2$. The sampling rate must be larger than twice the maximum frequency of the signal for the original continuous time signal to unambiguously represented by the samples. This sampling rate is defined as the the Nyquist sampling rate $f_{Nyquist}$ [65]:

$$f_s > f_{Nyquist} = 2f_{max} \tag{11}$$

In the present 5G standard, signal bandwidths on the order of MHz are being used [19], the higher frequency signal requires higher sampling rate. The cost of the ADC increases with both the sampling rate and the number of bits per sample (also known as the quantization rate) [66][67]. If coarse quantization can be used, meaning only a low number of bits is measured from each sample, then the cost of the ADC and thus the receiver can be reduced.

Analog signals have a continuous range of amplitudes. Digital systems must always have a finite number of amplitude levels, with larger numbers of levels leading to more complicated ADCs with higher power requirements [68]. The quantizer $\mathrm{Q}\{\cdot\}$ is applied at the ADC after the sampling procedure, and is chosen by quantization-bit (or precision) and step size. By definition, 4 or less bit quantization is considered as coarse quantization, quantization precision greater than 5 bits is considered as regular quantization [50]. Uniform quantizers have the same step size, meaning the

distances between all adjacent quantization levels are the same. On the other hand, non-uniform quantizers have inconsistent step sizes [67]. With the same number of quantization levels, when the quantization step size gets smaller, the distortion for quantized signals will get lower as well [66]. Uniform quantizers are the most commonly used due to their simplicity and low cost [68]. Therefore, for choosing the quantization level, this thesis applies a uniform quantizer, in which the received normalized power at each receiver equals to 1.

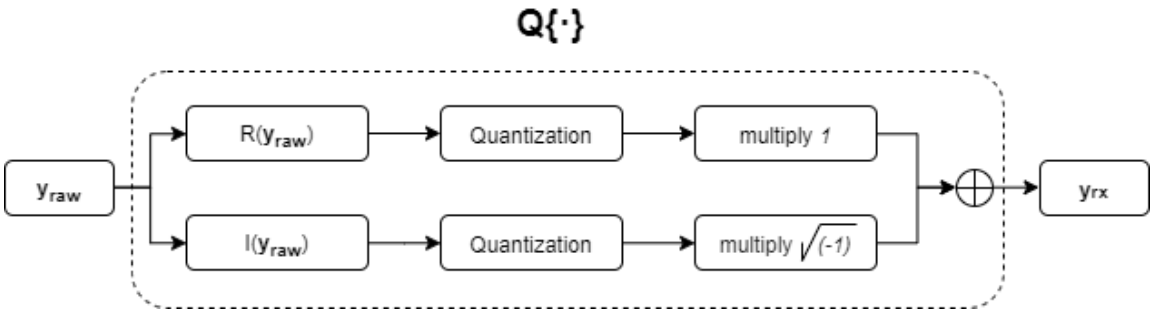

Figure 2.6: Quantization of complex signals

As in Fig. 2.6, the real and imaginary part represent the (I) and (Q) components for a signal in the real world. Mathematically, therefore, a quantizer in ADC actually separately quantizes real and imaginary parts.

Each sample value is mapped to a discrete level which is chosen according to the received power. This conversion of analog samples of the signal into digital form is called the quantizing process [50]:

$$\mathbf{y}_{rx} = \mathbf{Q}\{\mathbf{y}_{s,rx}\} \tag{12}$$

where $\mathbf{y}_{rx} \in \mathbb{C}$ represents quantized symbols at a receiver rx.

2.3.2 Pilot-aided Channel Estimation

The typical method for performing CE is to send a signal known at the receiver as a pilot signal from the transmitter, so that the receiver can estimate the radio channel using basic system estimation methods, providing the number of measurements

resulting from the pilot signal exceeds the number of channel parameters to be estimated. Unfortunately, during the transmission of the pilot signal, the transmitter is not sending a data signal, so the overall efficiency of the communications system is reduced. In this thesis, therefore, we focus on pilot-aided CE but use an iterative CE method to reduce the number of pilot blocks needed for an acceptable accuracy of the estimations.

Pilot-aided CE is a popular scheme for CE, though a conventional pilot-aided CE suffers from unavoidable pilot contamination. [69] illustrates that cross contamination between pilot signals is caused by correlation of pilot signals from different users, it can be suppressed by using a larger number of pilots sequences to lower the correlation among users [70] because two long random vectors with i.i.d elements have a near zero normalized dot product, as the length of the two vectors go to infinity. Unfortunately, increasing the pilots sequence length sacrifices the data rate [69][71][72] because transmitting extra pilot bits results in additional cost. Moreover, in large-scale MU-MIMO, this method also restricts the number of active users, because the more users in the system, the more pilots sequences are required, consequently, it becomes more likely that pilot sequences will suffer from cross-correlation [69]. The pilot signals which are used to estimate the channels can be contaminated as a result of reusing non-orthogonal pilot signals in a multi-user system [71]. Such contamination leads to inaccurate estimation. In [7], non-orthogonal pilot-aided CE always has pilot contamination that will remain even as $R_x \to \infty$. In this thesis, orthogonality of pilot sequences is not a necessity, because the CE method introduced in Chapter 3 only fractionally depends on short pilot sequences. The correction of the estimated channel is achieved by data sequences. Long data sequences are much more likely to be orthogonal to each other.

Pilot-aided CE methods generally are classified into block-type where all sub-carriers are sent at a time reserved for the pilot symbols transmission and comb-type where certain frequency sub-carriers are reserved for pilot transmission, as in Fig. 2.7 (a) and (b) respectively. Block-type and comb-type pilot-aided CE for the Massive MIMO

systems are investigated by [73] which proves that the comb-type slightly outperforms the block-type in terms of BER. However, for higher efficiency, SC-FDE is used in this thesis which does not allow each sub-carrier to be assigned for pilot transmission. As a result, instead of the comb-type that requires multiple carriers, block-type pilot insertion is more compatible with SC-FDE.

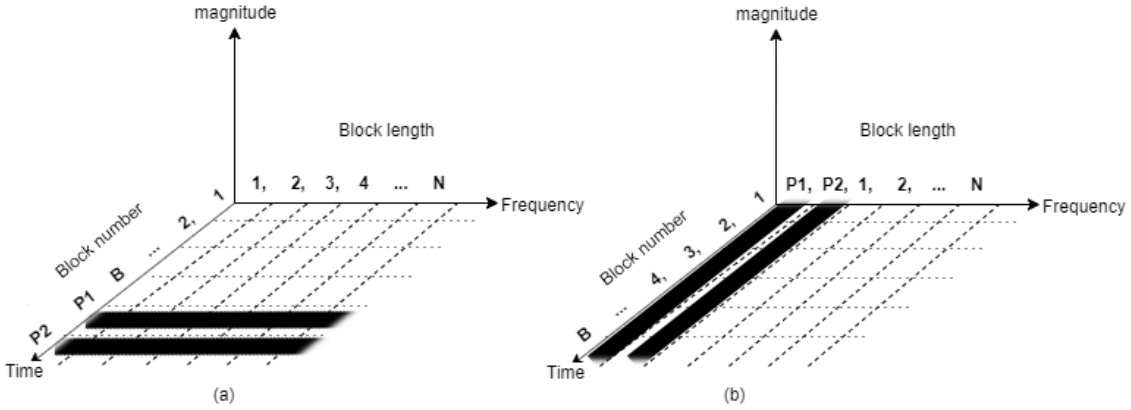

Figure 2.7: Schematics of block-type (a) and comb-type (b) pilot insertion

Channel parameter estimation is performed in the frequency domain, so it is necessary to convert (2) to its frequency domain equivalent. Let $\tilde{\mathbf{h}}$ denote the channel coefficients in the frequency domain. The FFT of the channel coefficients for the radio channel from a transmitter tx to a receiver rx is calculated as:

$$\tilde{\mathbf{h}}_{rx,tx}[k] = \sum_{n=0}^{N} \mathbf{h}_{rx,tx}[n]\omega^{kn} \tag{13}$$

where $\omega = e^{-2\pi j/N}$, $\mathbf{h}_{\mathrm{rx,tx}} = [h_0, ..., h_{L\text{-}1}, 0, ..., 0]_{1 \times N}$ as the CIR $h_{rx,tx}$ with padding of zeros. The FFT of the k^{th} sub-carrier block transmitted from the transmitter tx is denoted as $\tilde{\mathbf{x}}_{tx}[k]$:

$$\tilde{\mathbf{x}}_{tx}[k] = \sum_{n=0}^{N} \mathbf{x}_{k,tx}[n]\omega^{kn} \tag{14}$$

Similarly, the frequency domain channel noise at the receiver rx can be expressed as:

$$\tilde{\mathbf{v}}_{rx}[k] = \sum_{n=0}^{N} \mathbf{v}_{k,tx}[n]\omega^{kn} \tag{15}$$

From (13) - (15), the frequency domain system model from (2) becomes (16). Noting that convolution in the time-domain becomes simple multiplication in the frequency

domain:

$$\tilde{\mathbf{y}}_{rx}[k] = \sum_{tx=1}^{Tx} \tilde{\mathbf{h}}_{rx,tx}[k]\tilde{\mathbf{x}}_{tx}[k] + \tilde{\mathbf{v}}_{rx}[k], \tag{16}$$

where $\tilde{\mathbf{y}}_{rx}[k]$ refers to the received symbol vector at an antenna rx on the k^{th} sub-carrier, $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{v}}$ denote the frequency domain transmitted symbol vector and AWGN noise vector respectively.

For simplicity of expressions, we define $\mathcal{F}(\cdot)$ as the FFT operator. Let $\mathbf{F}$ denote the FFT matrix with the size of the block length and taps of the fading channel. The transformation matrix $\mathbf{F}$ is defined as $\mathbf{F} = (\omega^{n\times l})_{n=0\ldots N-1, l=0\ldots L-1}$, or equivalently:

$$\mathbf{F} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & \omega & \cdots & \omega^{(L-1)} \\ 1 & \omega^2 & \cdots & \omega^{(L-1)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{(N-1)} & \cdots & \omega^{(N-1)(L-1)} \end{bmatrix}$$

where $\omega = e^{-2\pi j/N}$. Now, the conversion from time to frequency domain can be seen as:

$$\mathcal{F}(\mathbf{y}_{rx}[p]) = \sum_{tx}^{Tx} \mathcal{F}(\mathbf{x}_{tx}[p] * \mathbf{h}_{rx,tx}) + \mathcal{F}(\mathbf{v}_{rx}[p]) \tag{17}$$

$$\tilde{\mathbf{y}}_{rx}[p] = \sum_{tx}^{Tx} \tilde{\mathbf{x}}_{tx}[p] \circ \tilde{\mathbf{h}}_{rx,tx} + \tilde{\mathbf{v}}_{rx}[p] \tag{18}$$

where $*$ denotes the convolution operation, $\circ$ denotes the element-wise product (also known as the Hadamard product [74]), $\mathbf{y}_{rx}[p]$ and $\tilde{\mathbf{y}}_{rx}[p]$ represent the time and frequency domain reception of a pilot block p at receiver rx, respectively.

A frequency domain measurement matrix in a linear system can be derived from the pilot observation at the receiver, so that a frequency domain measurement equation with the time domain CIR is created:

$$\tilde{\mathbf{y}}_{rx}[p] = \sum_{tx=1}^{Tx} \tilde{\mathbf{x}}_{tx}[p] \circ \mathbf{F}\mathbf{h}_{rx,tx} + \tilde{\mathbf{v}}_{rx}[p] \tag{19}$$

A measurement matrix $\mathbf{M}_{tx}$[p] for the pilot block $\tilde{\mathbf{x}}_{tx}[p]$ is defined as:

$$\mathbf{M}_{tx}[p] = Diag\{\tilde{\mathbf{x}}_{tx}[p]\}\mathbf{F} \tag{20}$$

where $Diag\{\tilde{\mathbf{x}}_{tx}[p]\}$ is the diagonal matrix created with the vector $\tilde{\mathbf{x}}_{tx}[p]$ specifying the diagonal entries. Thus, we have:

$$\tilde{\mathbf{y}}_{rx}[p] = \mathbf{M}_{tx}[p]h_{rx,tx} + \tilde{\mathbf{v}}_{rx}[p] \tag{21}$$

Commonly, multiple pilot blocks are used. The size of the final measurement matrix $\mathbf{M}$ is decided by the number of pilot blocks and fading channel paths. The overall measurement matrix for all pilot blocks $p = 1, 2 \cdots P$ for all Tx users is constructed as:

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_1[1] & \mathbf{M}_2[1] & \cdots & \mathbf{M}_{Tx}[1] \\ \mathbf{M}_1[2] & \mathbf{M}_2[2] & \cdots & \mathbf{M}_{Tx}[2] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{M}_1[P] & \mathbf{M}_2[P] & \cdots & \mathbf{M}_{Tx}[P] \end{bmatrix}$$

Here, we define

$$[\mathbf{A}; \mathbf{B}] = \begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix}$$

so ; inside of an matrix indicates stacking. The received pilot vector can be expressed with stacked received pilot vectors from each pilot block as:

$$\tilde{\mathbf{y}}_p = [\tilde{\mathbf{y}}_{rx}[1]; \tilde{\mathbf{y}}_{rx}[2] \cdots \tilde{\mathbf{y}}_{rx}[P]]$$

Similarly, the noise vectors can be derived as:

$$\tilde{\mathbf{v}}_p = [\tilde{\mathbf{v}}_{rx}[1]; \tilde{\mathbf{v}}_{rx}[2] \cdots \tilde{\mathbf{v}}_{rx}[P]]$$

Here, for Tx users, the estimation at a receiver rx of CIR matrix is expressed as:

$$\hat{\mathbf{h}} = [h_{rx,1}, h_{rx,2} \cdots h_{rx,Tx}]^T$$

where $\hat{\mathbf{h}}$ is defined as channel parameters vector for the receiver rx and $(\cdot)^T$ denotes vector transpose operation. The LMMSE has the lowest mean square error through the optimization problem [75], so the estimation of channel parameters can be derived as:

$$minimize \quad \mathbf{E}\{||(\hat{\mathbf{h}} - \mathbf{h})||_2^2\} \tag{22}$$

$$\textit{subject to} \begin{cases} \tilde{\mathbf{y}}[p] = \sum_{tx=1}^{Tx} \tilde{\mathbf{x}}[p] \circ \mathbf{F}\mathbf{h} + \tilde{\mathbf{v}}[p] \ , \ p = 1 \cdots P \\ \mathbf{E}\{\tilde{\mathbf{v}}\tilde{\mathbf{v}}^H\} = \sigma_v^2 \mathbf{I} \\ \mathbf{E}\{\mathbf{h}\mathbf{h}^H\} \approx \frac{1}{L \times Tx}\mathbf{I} \end{cases}$$

where $\mathbf{E}\{\cdot\}$ denotes expectation. The linear solution to the MMSE estimator in (22) is given by [76]:

$$\hat{\mathbf{h}} = \mathbf{R}_{\mathbf{h}\tilde{\mathbf{y}}_p} \mathbf{R}^{-1}_{\tilde{\mathbf{y}}_p \tilde{\mathbf{y}}_p} \tilde{\mathbf{y}}_p \tag{23}$$

with CE error in terms of Mean Squared Error (MSE) [76]:

$$e_{ce} = \mathbf{E}\{(\hat{\mathbf{h}} - \mathbf{h})^2\} = \mathbf{R}_{\mathbf{hh}} - \mathbf{R}_{\mathbf{h}\tilde{\mathbf{y}}_p} \mathbf{R}^{-1}_{\tilde{\mathbf{y}}_p \tilde{\mathbf{y}}_p} \mathbf{R}_{\tilde{\mathbf{y}}_p \mathbf{h}} \tag{24}$$

where $\mathbf{R}\{\cdot\}$ denotes the covariance matrix:

$$\mathbf{R}_{\mathbf{h}\tilde{\mathbf{y}}_p} = \mathbf{E}\{\mathbf{h}(\mathbf{M}\mathbf{h} + \mathbf{v})^H\} = \mathbf{M}^H \sigma_h^2 \tag{25}$$

$$\mathbf{R}_{\tilde{\mathbf{y}}_p \tilde{\mathbf{y}}_p} = \mathbf{E}\{(\mathbf{M}\mathbf{h} + \mathbf{v})(\mathbf{M}\mathbf{h} + \mathbf{v})^H\} = \mathbf{M}\mathbf{M}^H \sigma_h^2 + \sigma_v^2 \mathbf{I} \tag{26}$$

Noting that we assume the transmitted power is equally distributed between the L channel taps, therefore:

$$\mathbf{R}_{\mathbf{hh}} \approx \frac{1}{L \times Tx}\mathbf{I} \tag{27}$$

where L is the number of channel taps and Tx is the number of user terminals.

With the help of (25) - (27), from (23), we have:

$$\begin{aligned} \hat{\mathbf{h}} &= \mathbf{M}^H \sigma_h^2 (\mathbf{M}\mathbf{M}^H \sigma_h^2 + \sigma_v^2 \mathbf{I})^{-1} \tilde{\mathbf{y}}_p \\ &= \mathbf{M}^H (\mathbf{M}\mathbf{M}^H + \sigma_v^2 (L \times N \times \mathbf{I}))^{-1} \tilde{\mathbf{y}}_p \end{aligned} \tag{28}$$

By the push-through identity [77][1], the linear system is solved in a different order, and the complexity of matrices calculation is dropped from $(P \times N)$ by $(L \times Tx)$ to $(L \times Tx)$ by $(L \times Tx)$. Note that $(P \times N)$ has to be greater than $(L \times Tx)$, because the measurements of pilots have to be greater or equal to the unknown parameters

[1] In [77], the push-through identity is defined as $(I + UV)^{-1}U = U(I + VU)^{-1}$, where U is a $n \times k$ and V is a $k \times n$ matrix, matrix $I + UV$ is non-singular, i.e invertible.

of CSI so that the statistical estimation can be made. By applying the push-through identity to (28) for dimension reduction, we have:

$$\hat{\mathbf{h}} = (\mathbf{M}^H\mathbf{M} + \sigma_v^2(L \times N \times \mathbf{I}))^{-1}\mathbf{M}^H\tilde{\mathbf{y}}_p \tag{29}$$

To obtain the frequency domain channel efficiencies for all sub-carriers, simply padding zeros to the CIR then apply FFT. Thus, the channel coefficients vector from a user tx to a receiving antenna rx can be expressed as:

$$\hat{\mathbf{H}}_{rx,tx} = \mathcal{F}\{[\hat{\mathbf{h}}, 0, \cdots, 0]\} \tag{30}$$

Assuming the transmitted power at each transmitter is $\frac{1}{Tx}$, so that at each receiving antenna, the receivequalized signal's ed power is $\frac{1}{Tx} \times Tx \times \frac{1}{L} \times L = 1$, then the measurement error variance from CE over L taps is:

$$e_{mea} = L \times e_{ce} \tag{31}$$

and this error is added to the noise and treated as channel noise because we assumed that the channel and data symbols are independent:

$$\sigma_{\dot{v}}^2 = \sigma_v^2 + e_{mea} \tag{32}$$

The equalizer requires knowledge of the channel to estimate transmitted symbols. Commonly, we do not have perfect knowledge of the channel, thus we must use the CE error to compute an estimate of the uncertainty in the signal for the equalizer. Therefore, the reliability of the equalized signal can be computed as described in section 2.3.3.

In the end, the estimated frequency domain CSI matrix $\hat{\mathbf{H}}$ at a k^{th} sub-carrier can be written as:

$$\hat{\mathbf{H}}[k] = \begin{bmatrix} \hat{\mathbf{H}}_{1,1}[k] & \hat{\mathbf{H}}_{1,2}[k] & \cdots & \hat{\mathbf{H}}_{1,Tx}[k] \\ \hat{\mathbf{H}}_{2,1}[k] & \hat{\mathbf{H}}_{2,2}[k] & \cdots & \hat{\mathbf{H}}_{2,Tx}[k] \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\mathbf{H}}_{Rx,1}[k] & \hat{\mathbf{H}}_{Rx,2}[k] & \cdots & \hat{\mathbf{H}}_{Rx,Tx}[k] \end{bmatrix} \tag{33}$$

And the system model with AWGN at a k^{th} sub-carrier can be defined in matrix form as:

$$\tilde{\mathbf{y}}[k] = \hat{\mathbf{H}}[k]\tilde{\mathbf{x}}[k] + \tilde{\mathbf{v}}[k] \tag{34}$$

2.3.3 Linear Equalization

Multipath propagation of communication channels causes ISI in which one symbol interferes with subsequent symbols. The procedure that mitigates ISI is called equalization. Several solutions were introduced by researchers, two typical linear equalizers in frequency domain will be discussed in this section.

Frequency Domain Equalization (FDE) has less computational complexity comparing to Time Domain Equalization (TDE), since OFDM and SC-FDE only requires the order of $\log_2 N$ multiplications per data symbol as explained in Section 2.2.3. Overall, as Table 2.2, SC-FDE has both low computational complexity and PAPR [56].

	OFDM	SC-TDE	SC-FDE
PAPR	High	Low	Low
Computational Complexity	Low	High	Low

Table 2.2: Anti-multipath schemes comparison [56][59][62]

2.3.3.1 Zero-Forcing Equalization

The Zero-Forcing (ZF) equalizer is formulated to restore the transmitted signal after a multipath propagation channel. It was first proposed by Robert Lucky, whose work of ZF equalization was illustrated in [78]. Generally, the frequency domain ZF equalizer applies the inverse of the channel frequency gains to the received signals to restore the signal after the channel [79].

The equalized signal at the receiver is denoted as $\hat{\tilde{\mathbf{x}}}$. The frequency domain reception $\tilde{\mathbf{y}}$ is defined in (34). To obtain the inverse of a rectangular matrix $\hat{\mathbf{H}}$, the equalizer is expressed as the pseudo-inverse:

$$\hat{\tilde{\mathbf{x}}} = (\hat{\mathbf{H}}^H \hat{\mathbf{H}})^{-1} \hat{\mathbf{H}}^H \tilde{\mathbf{y}} \tag{35}$$

Based on (35), we can derive the error of ZF equalization:

$$\hat{\tilde{\mathbf{x}}} = (\hat{\mathbf{H}}^H\hat{\mathbf{H}})^{-1}\hat{\mathbf{H}}^H\hat{\mathbf{H}}\tilde{\mathbf{x}} + (\hat{\mathbf{H}}^H\hat{\mathbf{H}})^{-1}\hat{\mathbf{H}}^H\tilde{\mathbf{v}} = \tilde{\mathbf{x}} + \underbrace{(\hat{\mathbf{H}}^H\hat{\mathbf{H}})^{-1}\hat{\mathbf{H}}^H\tilde{\mathbf{v}}}_{\text{error}} \tag{36}$$

Generally, ZF has one significant drawback: AWGN could get boosted by a huge factor which eventually destroys the equalized signal's SNR when deep nulls in the channel frequency response exist [80]. To observe the drawback, we simply apply the Singular Value Decomposition (SVD) on the channel matrix so that $\hat{\mathbf{H}} = \mathbf{U}\mathbf{D}\mathbf{V}^H$, $\mathbf{U}$ and $\mathbf{V}$ denote the left-singular matrix and the right-singular matrix, and $\mathbf{D}$ denotes the singular value matrix whose diagonal entries are the non-negative real singular values of $\hat{\mathbf{H}}$. The gains of ZF equalization after the SVD becomes:

$$(\hat{\mathbf{H}}^H\hat{\mathbf{H}})^{-1}\hat{\mathbf{H}}^H = \mathbf{V}\mathbf{D}^{-2}\mathbf{D}\mathbf{U}^H \tag{37}$$

If a singular value d_H of $\mathbf{D}$ nears zero, $\frac{d_H}{d_H^2} = \frac{1}{d_H}$ will lead to noise amplification. In other words, ZF neglects the noise effect on the transmitted signals for portions of the signal when the SNR is low.

Note that (35) requires the computational complexity of $O(Tx^3)$ due to the matrix inversion of a size Tx. To save the cost from computation of matrix inversion, a Jacobi algorithm [15] is introduced. The Jacobi method is an iterative algorithm for solving matrix-vector linear systems where the matrix is diagonally dominant. $\hat{\mathbf{H}}^H\hat{\mathbf{H}}$ can be decomposed into a strictly diagonal matrix $\mathbf{D}_s$ and an off-diagonal matrix $\mathbf{R}$. $\mathbf{D}_s$ is a strictly diagonal when the entries outside the main diagonal are all zero. On the contrary, $\mathbf{R}$ is an off-diagonal matrix when any entry of $\mathbf{R}$ that is not on its main diagonal.

Result: Estimated solution of the linear system

$\hat{\mathbf{H}}^H\hat{\mathbf{H}} = \mathbf{D}_s + \mathbf{R}$

$\hat{\tilde{\mathbf{x}}}_0 = \mathbf{D}_s^{-1}\hat{\mathbf{H}}^H\tilde{\mathbf{y}}$

$k = 1$

while *convergence not reached* **do**

- $e_k = \hat{\mathbf{H}}^H - \mathbf{D}_s\hat{\tilde{\mathbf{x}}}_{k-1}$
- $\hat{\tilde{\mathbf{x}}}_k = \mathbf{D}_s^{-1}e_k$
- $k = k + 1$

The trade-off of the Jacobi algorithm is that off-diagonal elements are neglected, which leads to estimation errors. When the matrix is diagonally or diagonally dominant, such as $\hat{\mathbf{H}}^H\hat{\mathbf{H}}$ due to the Gaussian distribution of the elements of the channel matrix $\hat{\mathbf{H}}$, Jacobi iteration performs greatly [81]. In this thesis, inspired by the Jacobi algorithm, we here apply the Jacobi estimation with 1 iteration, the estimation error in the algorithm is replace by the Successive Interference Cancellation (SIC), which will iteratively lower the estimation error in Chapter 3. Here, each diagonal element $a_{tx,tx}$ of $\mathbf{D}_s$ is approximated by their corresponding inversions $1/a_{tx,tx}$:

$$\hat{\tilde{\mathbf{x}}} \approx 1./Diag(\hat{\mathbf{H}}^H\hat{\mathbf{H}}) \circ (\hat{\mathbf{H}}^H\tilde{\mathbf{y}}) \tag{38}$$

where ./ denotes element-wise inversion and $\circ$ denotes element-wise product. In this case, first the element-wise inverse is applied on the main diagonal entries of $\hat{\mathbf{H}}^H\hat{\mathbf{H}}$, then element-wise product of $1./Diag(\hat{\mathbf{H}}^H\hat{\mathbf{H}})$ and $\hat{\mathbf{H}}^H\tilde{\mathbf{y}}$ is applied. With calculating the diagonal-entry inversions first, the computational complexity becomes $O(Tx^2)$.

It is useful for the equalizer to also calculate the variance of the estimated transmitted signal which is used by the symbol detection algorithm. Noting that we assume symbols in $\tilde{\mathbf{x}}$ and $\hat{\tilde{\mathbf{x}}}$ are i.i.d., to compute reliability values for its detection values, this variance δ_{ZF} is calculated as:

$$\mathbf{E}\{\tilde{\mathbf{x}}\hat{\tilde{\mathbf{x}}}^H\} = \mathbf{E}\{\hat{\tilde{\mathbf{x}}}^H\tilde{\mathbf{x}}\} = \sigma_x^2\mathbf{I} \tag{39}$$

$$\mathbf{E}\{\hat{\tilde{\mathbf{x}}}\hat{\tilde{\mathbf{x}}}^H\} = \sigma_x^2\mathbf{I} + \sigma_{\dot{v}}^2(\hat{\mathbf{H}}^H\hat{\mathbf{H}})^{-1} \tag{40}$$

$$\begin{aligned}
\delta_{ZF} &= \mathbf{E}\{(\tilde{\mathbf{x}} - \hat{\tilde{\mathbf{x}}})(\tilde{\mathbf{x}} - \hat{\tilde{\mathbf{x}}})^H\} \\
&= \mathbf{E}\{\tilde{\mathbf{x}}\tilde{\mathbf{x}}^H - \tilde{\mathbf{x}}\hat{\tilde{\mathbf{x}}}^H - \hat{\tilde{\mathbf{x}}}\tilde{\mathbf{x}}^H + \hat{\tilde{\mathbf{x}}}\hat{\tilde{\mathbf{x}}}^H\} \\
&= \sigma_x^2\mathbf{I} - \sigma_x^2\mathbf{I} - \sigma_x^2\mathbf{I} + \sigma_x^2\mathbf{I} + \sigma_{\dot{v}}^2(\hat{\mathbf{H}}^H\hat{\mathbf{H}})^{-1} \\
&= \sigma_{\dot{v}}^2(\hat{\mathbf{H}}^H\hat{\mathbf{H}})^{-1}
\end{aligned} \tag{41}$$

where σ_x^2 denotes variance of signal vector $\tilde{\mathbf{x}}$, i.e the power of the signal from an individual user, $\sigma_{\dot{v}}^2$ is defined in (32), representing the variance of AWGN noise with the CE error being considered as an independent white noise [82].

2.3.3.2 Minimum Mean Square Error Equalization

The MMSE equalization algorithm computes the estimated symbol values which minimize the squared error between the estimated and true symbol values. The MMSE equalizer uses knowledge of statistical characteristics of the transmitted signal, noise, and channel parameters to perform equalization.

Since our model (34) is a noisy linear system, the LMMSE equalization problem can be solved as:

$$\begin{aligned} & minimize \quad \mathbf{E}\{||\hat{\tilde{\mathbf{x}}} - \tilde{\mathbf{x}}||_2^2\} \\ subject\ to & \begin{cases} \tilde{\mathbf{y}}[b] = \hat{\mathbf{H}}[b]\tilde{\mathbf{x}}[b] + \tilde{\mathbf{v}}[b]\ ,\ b = 1 \cdots B \\ \mathbf{E}\{\tilde{\mathbf{v}}\tilde{\mathbf{v}}^H\} = \sigma_v^2\mathbf{I} \\ \mathbf{E}\{\tilde{\mathbf{x}}\tilde{\mathbf{x}}^H\} = \sigma_x^2\mathbf{I} \end{cases} \end{aligned} \tag{42}$$

By solving the MMSE detector of (42) [76], we have:

$$\hat{\tilde{\mathbf{x}}} = \mathbf{R}_{\tilde{\mathbf{x}}\mathbf{y}}\mathbf{R}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}^{-1}\tilde{\mathbf{y}} \tag{43}$$

The cross-covariance matrix of $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$ and self-covariance matrix of $\tilde{\mathbf{y}}$ can be purposed as follow:

$$\mathbf{R}_{\tilde{\mathbf{x}}\tilde{\mathbf{y}}} = \mathbf{E}\{\tilde{\mathbf{x}}(\mathbf{H}\tilde{\mathbf{x}} + \mathbf{v})^H\} = \hat{\mathbf{H}}^H\sigma_x^2 \tag{44}$$

$$\mathbf{R}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}} = \mathbf{E}\{(\mathbf{H}\tilde{\mathbf{x}} + \mathbf{v})(\mathbf{H}\tilde{\mathbf{x}} + \mathbf{v})^H\} = \hat{\mathbf{H}}\hat{\mathbf{H}}^H\sigma_x^2 + \sigma_v^2\mathbf{I} \tag{45}$$

Put (44) and (45) into (43), the final LMMSE estimation becomes:

$$\hat{\tilde{\mathbf{x}}} = \hat{\mathbf{H}}^H\sigma_x^2(\hat{\mathbf{H}}\hat{\mathbf{H}}^H\sigma_x^2 + \sigma_v^2\mathbf{I})^{-1}\tilde{\mathbf{y}} \tag{46}$$

Note that $\hat{\mathbf{H}}$ has a dimension $Rx \times Tx$, calculating (46) results in a matrix inversion of a $Rx \times Rx$ matrix, which is expensive and inefficient.

Suppose we have the MMSE estimation $\hat{\tilde{\mathbf{x}}}$, by the push-through identity [77], we have:

$$\hat{\tilde{\mathbf{x}}} = \hat{\mathbf{H}}^H\sigma_x^2(\hat{\mathbf{H}}\hat{\mathbf{H}}^H\sigma_x^2 + \sigma_v^2\mathbf{I})^{-1}\tilde{\mathbf{y}} = \sigma_x^2(\hat{\mathbf{H}}^H\hat{\mathbf{H}}\sigma_x^2 + \sigma_v^2\mathbf{I})^{-1}\hat{\mathbf{H}}^H\tilde{\mathbf{y}} \tag{47}$$

which lowers the $Rx \times Rx$ matrix inversion down to $Tx \times Tx$. Meanwhile, consider the factorization (SVD) of channel matrix $\hat{\mathbf{H}} = \mathbf{U}\mathbf{D}\mathbf{V}^{\hat{H}}$. The error covariance δ_{MMSE} is a figure of merit that indicates the quality of the current symbol estimation. It can be formulated as follow:

$$\begin{aligned}
\delta_{MMSE} &= \mathbf{E}\{(\tilde{\mathbf{x}} - \hat{\tilde{\mathbf{x}}})(\tilde{\mathbf{x}} - \hat{\tilde{\mathbf{x}}})^H\} \\
&= \mathbf{E}\{\tilde{\mathbf{x}}\tilde{\mathbf{x}}^H - \tilde{\mathbf{x}}\hat{\tilde{\mathbf{x}}}^H - \hat{\tilde{\mathbf{x}}}\tilde{\mathbf{x}}^H + \hat{\tilde{\mathbf{x}}}\hat{\tilde{\mathbf{x}}}^H\} \\
&= \sigma_x^2\mathbf{I} - \mathbf{R}_{\tilde{\mathbf{x}}\tilde{\mathbf{y}}}\mathbf{R}_{\mathbf{y}\tilde{\mathbf{y}}}^{-1}\mathbf{R}_{\tilde{\mathbf{x}}\mathbf{y}}^H - \mathbf{R}_{\tilde{\mathbf{x}}\tilde{\mathbf{y}}}\mathbf{R}_{\mathbf{y}\mathbf{y}}^{-1}\mathbf{R}_{\tilde{\mathbf{x}}\tilde{\mathbf{y}}}^H + \mathbf{R}_{\tilde{\mathbf{x}}\tilde{\mathbf{y}}}\mathbf{R}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}^{-1}\mathbf{R}_{\tilde{\mathbf{x}}\tilde{\mathbf{y}}}^H \\
&= \sigma_x^2\mathbf{I} - \mathbf{R}_{\tilde{\mathbf{x}}\tilde{\mathbf{y}}}\mathbf{R}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}^{-1}\mathbf{R}_{\tilde{\mathbf{x}}\tilde{\mathbf{y}}}^H \\
&= \sigma_x^2\mathbf{I} - \hat{\mathbf{H}}^H(\hat{\mathbf{H}}\hat{\mathbf{H}}^H + \mathbf{I}\frac{\sigma_{\dot{v}}^2}{\sigma_x^2})^{-1}\hat{\mathbf{H}}\sigma_x^2 \\
&= \sigma_x^2\mathbf{I} - \mathbf{V}\mathbf{D}^H\mathbf{U}^H(\mathbf{U}\mathbf{D}\mathbf{D}^H\mathbf{U}^H + \mathbf{U}\mathbf{U}^H\frac{\sigma_{\dot{v}}^2}{\sigma_x^2})^{-1}\mathbf{U}\mathbf{D}\mathbf{V}^H\sigma_x^2 \\
&= \sigma_x^2\mathbf{I} - \mathbf{V}\mathbf{D}^H(\mathbf{D}\mathbf{D}^H + \mathbf{I}\frac{\sigma_{\dot{v}}^2}{\sigma_x^2})^{-1}\mathbf{D}\mathbf{V}^H\sigma_x^2
\end{aligned} \tag{48}$$

Through the Woodbury identity [83][84] (also known as the Matrix Inversion Lemma)[2], the simplification of (48) is:

$$\delta_{MMSE} = \left(\frac{1}{\sigma_x^2}\mathbf{I} + \frac{\mathbf{V}\mathbf{D}^H\mathbf{D}\mathbf{V}^H}{\sigma_{\dot{v}}^2}\right)^{-1} \tag{49}$$

Consider the system model (34), given the estimation of $\hat{\tilde{\mathbf{x}}}$ at (46), its corresponding MSE is given by the trace of the error covariance matrix at (48). As the error covariance of the MMSE equalizer in terms of the reliability δ_{MMSE} will be used for detector in Section 3.1.

$$e_{mmse} = \mathrm{tr}\{\delta_{MMSE}\} = \mathrm{tr}\left\{\sigma_x^2\mathbf{I} - \mathbf{V}\mathbf{D}^H(\mathbf{D}\mathbf{D}^H + \mathbf{I}\frac{\sigma_{\dot{v}}^2}{\sigma_x^2})^{-1}\mathbf{D}\mathbf{V}^H\sigma_x^2\right\} \tag{50}$$

[2]In [83], the Woodbury identity was found by Max A. Woodbury in 1950 as: $(A+UCV)^{-1} = A^{-1} - A^{-1}U(C^{-1} + VA^{-1}U)^{-1}VA^{-1}$, where A is $n \times n$, U is $n \times k$, C is $k \times k$ and V is $k \times n$, $C^{-1} + VA^{-1}U$ is invertible.

3 Iterative Decision Feedback Receiver

To construct an iterative decision feedback receiver, a soft decision flow is required. A symbol detection algorithm that produces soft decisions and a soft-input/soft-output decoder are briefly described at the beginning of this chapter, both will support iterative CE and equalization. Then we show how the soft decisions are used to assist the CE and equalization for iteratively improving performance. Lastly, two equalization algorithms from Chapter 2 are compared for iterative equalization.

A block diagram of the iterative decision feedback receiver is presented as Fig. 3.1. At a receiver rx, the quantized received symbol vector $\mathbf{y}_{rx}$ will be transformed into the frequency domain for FDE as $\tilde{\mathbf{y}}$. At the first iteration, the CE is performed using only pilot symbols, producing the initial estimation $\hat{\mathbf{H}}_{(1)}$ for equalization, after which an IFFT operation brings the equalized symbols back to the time domain for the deinterleaving, detection and decoding process.

In Section 3.1, a definition of soft decision is given. The soft decision indicates the probability of a bit value being 0 or 1. The extrinsic soft decision $\mathcal{L}_{ex}$ is calculated after the decoder, as the input bits' soft decision values subtracted from the output bits' soft decision values. An extrinsic soft decision value of zero indicates that the system has no information about the bit values, the uncertain bits will be re-estimated in the next iteration with more available information to assist. After interleaving, the extrinsic soft bit decisions will be remapped to symbol values as $\bar{\mathbf{x}}_{(i-1)}$, then converted into the frequency domain, and then the resulting frequency domain signal is subtracted from the received signal's frequency domain representation to get $\Delta\tilde{\mathbf{y}}_{(i)}$ $(i \geq 2, i \in \mathcal{Z})$.

At this moment, a new iteration has begun. Updated mean symbols $\bar{\mathbf{x}}_{(i-1)}$ in each iteration will have lower error than the previous iteration until its convergence, help each iteration's data-aided CE to have lower MSE than the previous iteration, so long as the extrinsic information is reliable. Since the CE is more accurate, the equalizer

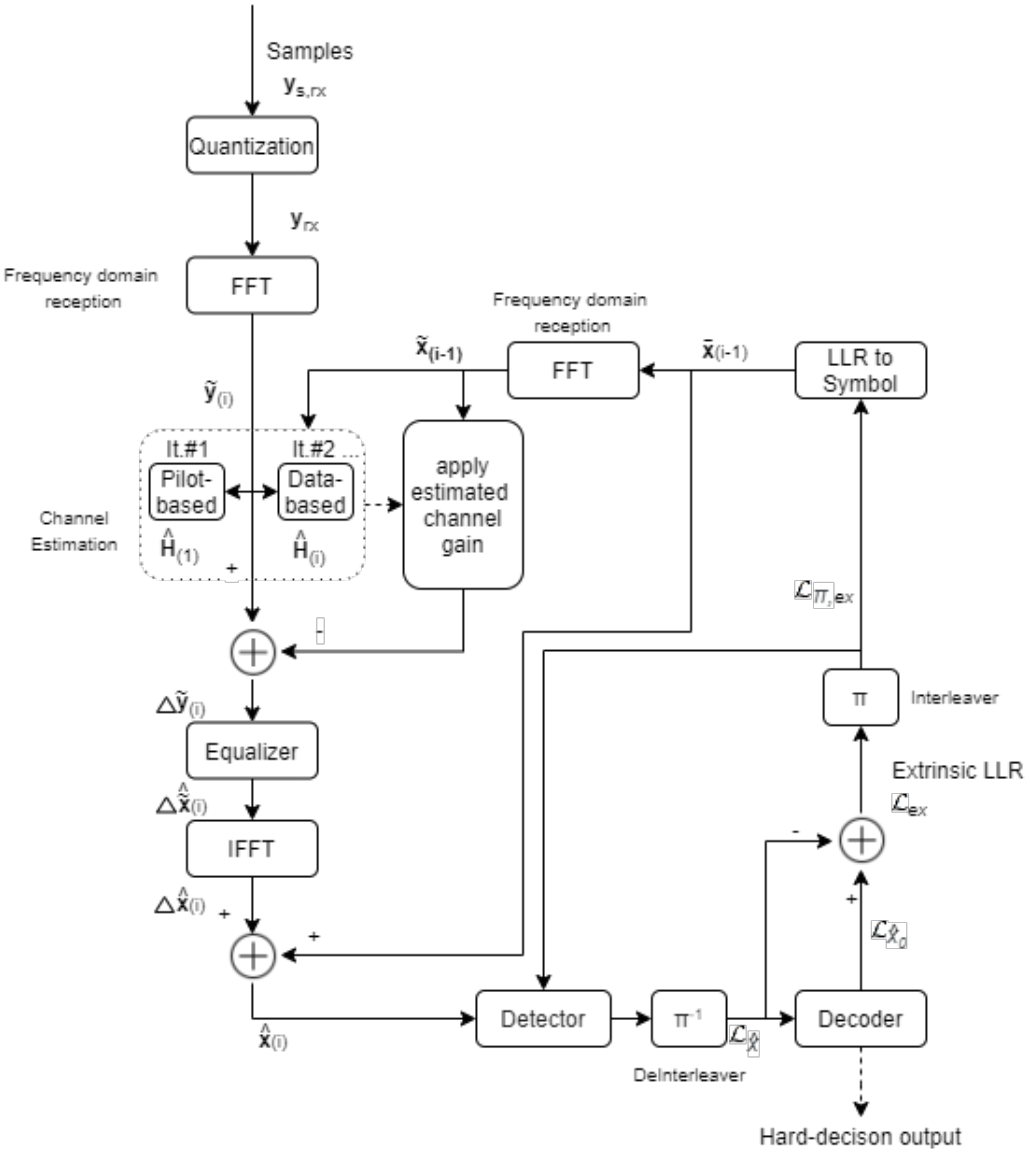

Figure 3.1: Iterative receiver

in each successive iteration can perform a better estimation of the transmitted signal. More accurate equalizer output results in the reliability of the detector estimates improving and a lower BER.

3.1 Symbol Detection and Soft-Input/Soft-Output Decoder

Symbol detection, also known as demodulation, aims to convert the time domain equalized symbol back to bit sequences. In this section, we present a soft-input/soft-output decoder which takes in the bit probabilities from the previous iteration's decoder as well as the estimated symbols signal from the equalizer to calculate updated probabilities of the data bits.

To demodulate symbols, the squared distance of each element of the equalized symbol vector, $\hat{\tilde{\mathbf{x}}}$, from a possible symbol value s is defined as:

$$\mathbf{d}(\hat{\tilde{\mathbf{x}}}|s) = \left|(\hat{\tilde{\mathbf{x}}} - s\mathbf{1})\right|^2 \tag{51}$$

where $\mathbf{1}$ denotes a $1 \times NB$ vector with all 1 entries, the vector $\mathbf{d}(\hat{\tilde{\mathbf{x}}}|s)$ contains the squared distance from each equalized symbol to a transmitted symbol s.

We assume the real and imaginary components of the error of the equalized signals are i.i.d Gaussian random variables. That is, it is assumed that equalization errors are circularly symmetric complex Gaussian random variables. For each element d in $\mathbf{d}(\hat{\tilde{\mathbf{x}}}|s)$, the conditional complex Gaussian density of the noise energy, given that the transmitted symbol is s, can be calculated with the help of error variance:

$$f(\hat{\tilde{x}}|s) = K \exp\left\{\frac{-d}{\delta}\right\} \tag{52}$$

where K represents an appropriate scalar so the probability density function in (52) integrates to 1, δ denotes the variance of the equalization estimation error which assumed to be equal for all symbol signal indices and is calculated from the mean of main diagonal value of the error covariance matrix provided by the equalization from (41) or (49). The symbol likelihoods are transformed into bit Log Likelihood Ratios (LLRs). For a bit x_0, the LLR $\mathcal{L}_{x_0}$ is defined as:

$$\mathcal{L}_{x_0} \triangleq \ln\left(\frac{\Pr(x_0 = 0|y_0)}{\Pr(x_0 = 1|y_0)}\right) \tag{53}$$

where y_0 represents the measurement and prior information on the bit values used to compute the bit value probability, $\Pr(x|y)$ denotes the conditional probability of

x given y. Whether the probability of a bit value being 0 or 1 is in the numerator or denominator of the expression inside of the logarithm in (53) is designer dependent. The absolute value of the LLR represents the certainty of the bit, and the sign of the LLR is the hard-decision with a positive LLR indicating the bit is more likely to have a 0 value, and a negative LLR indicating the bit is more likely to have a 1 value.

To convert from symbols probabilities to LLR, based on the definition of (53), the n^{th} bit's LLR $\mathcal{L}_n$ of a received symbol vector $\hat{\tilde{\mathbf{x}}}$ is approximated as:

$$\mathcal{L}_n = \frac{\sum\limits_{s \in S_{mod}^{n,0}} f(\hat{\tilde{\mathbf{x}}}|s)}{\sum\limits_{s \in S_{mod}^{n,1}} f(\hat{\tilde{\mathbf{x}}}|s)} \tag{54}$$

where $S_{mod}^{n,0}$ represents the set of symbols whose n^{th} bit is a 0, $S_{mod}^{n,1}$ represents the set of symbols whose n^{th} bit is a 1, $f(\hat{\mathbf{x}}|s)$ denotes the vector consisted of the noise density from all symbols derived in (52).

With a soft decision (LLR) being calculated from the symbol detector, a soft-input/soft-output decoder [85] takes in numerical soft decision values indicating the bit decision reliability in terms of LLR of the bit decisions from the symbol detector sub-system and outputs numerical values indicating the reliability of its own decisions on the encoded data bits.

This is different from a hard-input/hard-output decoder which takes in only the most likely bit values from the detector and output only the most likely calculated values of the encoded bits. Simply using only zeros and ones usually results in a loss of information [85], therefore soft-input/soft-output provides superior performance.

3.2 Iterative Data-aided Channel Estimation

An iterative CE method can reduce the need for a long pilot sequence to achieve a required level of CE accuracy. The users data sequences have all the properties

required of ideal pilot sequences with the notable exception that they are unknown to the receiver at the start of the receiver algorithm calculations. In this thesis, pilot sequences are used for the initial estimation of the channel to support the initial estimation of the users' data sequences.

In the initial iteration of the receiver, the channel is estimated using (29) based on the use of the pilot symbols. The CE error from (24) is considered in conjunction with the unknown symbol values to compute a new measurement noise variance in (32). The equalization produces the error covariance for the data detection to compute detector output LLR in Section 3.1.

The information in the receiver's first iteration about the transmitted data after the detector and de-interleaver Π^{-1} is represented by the LLR values $\mathcal{L}_{\hat{X}}{}^{(1)}$. After decoding first iteration, soft-output values in a form of LLR — $\mathcal{L}_{\hat{X}_o}{}^{(1)}$ denotes the posterior LLR of corresponding uncoded bit sequence X_o. The subtraction of the prior from the posterior LLR forms the extrinsic LLR. The extrinsic LLR illustrates what the decoder contributes to knowledge of the transmitted data in the current iteration.

We define $\mathcal{L}_{ex}^{(1)}$ as the extrinsic posterior LLR in the first iteration:

$$\mathcal{L}_{ex}^{(1)} = \mathcal{L}_{\hat{X}_o}^{(1)} - \mathcal{L}_{\hat{X}}^{(1)} \tag{55}$$

$\mathcal{L}_{ex}$ represents the extrinsic posterior LLR of the coded bits. Bit locations with zero extrinsic LLR means that the decoder is not providing any more information on those transmitted bits than the detector. Only the extrinsic LLR value is fed back to the next iteration. That is, only new information calculated by the decoder about a given bit is used for the feedback. This reduces the chance that a bad bit decision by the detector reinforces itself by contaminating the CE and equalization in the following iterations.

From the extrinsic bit LLRs, the receiver must compute an estimate of the transmitted signal for the usage of the iterative CE and equalizer. The extrinsic outputs are

assumed to be independent of each other [40]. The interleaver is applied on $\mathcal{L}_{ex}$ as (4) to remove the dependencies by random permutation:

$$\mathcal{L}_{\pi,ex} = \Pi\{\mathcal{L}_{ex}, st\} \tag{56}$$

In order to fit into the receiving data symbols, we need to obtain the estimated symbols based on $\mathcal{L}_{\pi,ex}$. From (56) and the definition of the LLR in (53), the bits probabilities can be derived as:

$$\begin{cases} \Pr(\hat{x} = 0) = \frac{1}{1+e^{(\mathcal{L}_{\pi,ex})}} \\ \Pr(\hat{x} = 1) = \frac{1}{1+e^{(-\mathcal{L}_{\pi,ex})}} \end{cases} \tag{57}$$

Using the formulas in Table 3.1, the probability of each possible symbol value for 4-QAM/QPSK symbol with Gray coding is calculated where $\Pr(x = b)_{re}$ is the probability of the bit values b being 0 or 1, for the bit assigned to the real axis, and $\Pr(x = b)_{im}$ is the probability of the bit values for the bit assigned to the imaginary axis. The mean of estimated symbols is calculated by multiplying normalized symbol value row vector S_{mod} with the symbol probabilities matrix $\hat{\mathbf{P}}_{\mathrm{M\times NB}}$, in which M represents the M possible symbol values ($M = 4$ in 4-QAM/QPSK). The mean symbol vector $\bar{\mathbf{x}}_{\text{(i-1)}}$ of a user tx from iteration $i-1 (i \geq 2, i \in \mathcal{Z})$ is:

$$\bar{\mathbf{x}}_{(i-1)} = S_{mod}\hat{\mathbf{P}}_{(i-1)} \tag{58}$$

Let $|\cdot|^2$ denote the element-wise squared absolute value of a vector. The corresponding variance vector is:

$$\sigma^2_{\bar{\mathbf{x}}_{(i-1)}} = |S_{mod}|^2\hat{\mathbf{P}}_{(i-1)} - |\bar{\mathbf{x}}_{(i-1)}|^2 \tag{59}$$

Consider the previous mean symbols from estimation vector $\bar{\mathbf{x}}_{i-1}$. The channel can be estimated by using these estimated data symbols as virtual pilot symbols using a variation of the methodology described in Section 2.3.2 as described below. Assume that $\tilde{\mathbf{x}}_{(i-1)}[b]$ is the estimated b^{th} transmitted data block in the frequency domain, computed by taking the the FFT of $\bar{\mathbf{x}}_{(i-1)}[b]$. We have the measurement equation:

$$\tilde{\mathbf{y}}_{rx}[b] = \sum_{tx}^{Tx} \tilde{\mathbf{x}}_{tx,(i-1)}[b] \circ \mathbf{F}\mathbf{h}_{rx,tx} + \tilde{\mathbf{v}}^{+}_{rx,(i)}[b] \tag{60}$$

Estimate bit-pair	Symbol Probability	Symbol value
0, 0	$\Pr(x=0)_{re} \times \Pr(x=0)_{im}$	1 + j
1, 0	$\Pr(x=1)_{re} \times \Pr(x=0)_{im}$	-1 + j
1, 1	$\Pr(x=1)_{re} \times \Pr(x=1)_{im}$	-1 - j
0, 1	$\Pr(x=0)_{re} \times \Pr(x=1)_{im}$	1 - j

Table 3.1: Estimated QPSK/4-QAM symbol from bit probabilities

where $\circ$ represents the element-wise product. The uncertainty of feedback data is modeled as additional white noise with zero mean at a receiving antenna rx, so $\tilde{\mathbf{v}}^{+}_{rx,(i)}$ represents the AWGN with a variance of the measurement noise increased by a factor to consider the uncertainty in the feedback data knowledge which is Gaussian distributed. $\mathbf{F}$ denotes the FFT matrix, defined as $\mathbf{F} = (\omega^{n\times l})_{n=0\ldots N-1, l=0\ldots L-1}$. Assume the power distribution is $\frac{1}{L}$ on each path (as discussed in Section 2.1), so the additional noise from overall L paths can be added into the original AWGN $\tilde{\mathbf{v}}$. a The variance of the measurement noise $\sigma^2_{v^{+}_{(i)}}$ for the CE using the data symbols as the virtual pilot symbols is given by:

$$\sigma^2_{v^{+}_{(i)}} = \sigma^2_v + \sigma^2_{\tilde{x}_{(i-1)}} \times \frac{1}{L} \times L \tag{61}$$

To once again construct the measurement matrix $\mathbf{M}[b]$ for a b^{th} data block of $\tilde{\mathbf{x}}_{tx,(i-1)}$, it is computed as in (20):

$$\mathbf{M}_{tx}[b] = Diag\{\tilde{\mathbf{x}}_{tx,(i-1)}[b]\}\mathbf{F} \tag{62}$$

Then, the measurement model of (61) based on the b^{th} data block at receive antenna rx becomes:

$$\tilde{\mathbf{y}}_{rx}[b] = \mathbf{M}_{tx}[b]h_{rx,tx} + \tilde{\mathbf{v}}^{+}_{rx,(i)}[b] \tag{63}$$

Consider all Tx active users and B data blocks, the overall measurement matrix is constructed as:

$$\mathbf{M}_x = \begin{bmatrix} \mathbf{M}_1[1] & \mathbf{M}_2[1] & \cdots & \mathbf{M}_{Tx}[1] \\ \mathbf{M}_1[2] & \mathbf{M}_2[2] & \cdots & \mathbf{M}_{Tx}[2] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{M}_1[B] & \mathbf{M}_2[B] & \cdots & \mathbf{M}_{Tx}[B] \end{bmatrix}$$

Correspondingly, at a receiver rx, the received data vector with dimension NB is:

$$\tilde{\mathbf{y}}(rx) = [\tilde{\mathbf{y}}_{rx}[1]; \tilde{\mathbf{y}}_{rx}[2] \cdots \tilde{\mathbf{y}}_{rx}[B]]_{NB}$$

The $NB \times 1$ noise vector can be derived as:

$$\tilde{\mathbf{v}}_i^+(rx) = [\tilde{\mathbf{v}}_{rx,(i)}^+[1]; \tilde{\mathbf{v}}_{rx,(i)}^+[2] \cdots \tilde{\mathbf{v}}_{rx,(i)}^+[B]]_{NB}$$

The channel parameter matrix for all Tx active users as previously defined in Section 2.3.2:

$$\mathbf{h} = [h_{rx,1}, h_{rx,2} \cdots h_{rx,Tx}]^T$$

Therefore, at a receiving antenna rx, we have:

$$\tilde{\mathbf{y}} = \mathbf{M}_x \mathbf{h} + \tilde{\mathbf{v}}_{(i)}^+ \tag{64}$$

To derive the combination of CE using the pilots and CE using the estimated data symbols, define:

$$\hat{\mathbf{h}}_{(i)} = \hat{\mathbf{h}}_{(1)} + \Delta\hat{\mathbf{h}}_{(i)} \tag{65}$$

where $\hat{\mathbf{h}}_{(\mathrm{i})}$ denotes the estimation of the channel at iteration i, $\hat{\mathbf{h}}_{(1)}$ denotes the initial estimation of (29) based on pilot symbols. $\Delta\hat{\mathbf{h}}_{(\mathrm{i})}$ denotes the innovation of CE at iteration i based on using the estimated data symbols as virtual pilots. We have the LMMSE problem [76] for the CE formulated as:

$$\underbrace{\tilde{\mathbf{y}} - \mathbf{M}_x \mathbf{h}_{(1)}}_{\text{innovation vector}} = \mathbf{M}_x \Delta\hat{\mathbf{h}}_{(i)} + \tilde{\mathbf{v}}_{(i)}{}^+ \tag{66}$$

The innovation vector corresponds to the difference between the measured signal and the expected data signal measurements, if the CE using the pilots and data estimates from the previous iteration are both correct. For a simpler expression, define:

$$\Delta\tilde{\mathbf{y}}_{(i)} = \tilde{\mathbf{y}} - \mathbf{M}_x \mathbf{h}_1$$

The optimization problem in terms of MSE is:

$$minimize \quad \mathbf{E}\{||(\hat{\mathbf{h}}_{(1)} + \Delta\hat{\mathbf{h}}_{(i)} - \mathbf{h})||_2^2\} \tag{67}$$

Since $\hat{\mathbf{h}}_1$ is fixed and given in the first iteration, (67) becomes:

$$minimize \quad \mathbf{E}\{||(\Delta\hat{\mathbf{h}}_{(i)} - \mathbf{h})||_2^2\} \tag{68}$$

$$subject\ to \begin{cases} \tilde{\mathbf{y}}[b] = \sum_{tx=1}^{Tx} \tilde{\mathbf{x}}[b] \circ \mathbf{F}\mathbf{h} + \tilde{\mathbf{v}}[b]\ ,\ b = 1 \cdots B \\ \mathbf{E}\{\tilde{\mathbf{v}}^{+}\tilde{\mathbf{v}}^{+H}\} = \sigma^2_{v^{+}_{(i)}}\mathbf{I} \\ \mathbf{E}\{\mathbf{h}\mathbf{h}^H\} \approx \frac{1}{L \times Tx}\mathbf{I} \\ i > 1\ ,\ i \in \mathcal{Z} \end{cases}$$

By solving (68) as a standard LMMSE estimator [76], we have:

$$\hat{\mathbf{h}}_{(i)} = \hat{\mathbf{h}}_{(1)} + \mathbf{M}_x^H(\mathbf{M}_x\mathbf{M}_x^H + \sigma^2_{v^{+}_{(i)}}(L \times N \times \mathbf{I}))^{-1}\Delta\tilde{\mathbf{y}}_{(i)} \tag{69}$$

where $\mathbf{I}$ is an appropriately sized identity matrix. Note that the number of data symbols is significantly greater than the pilot symbols, so the measurement matrix $\mathbf{M}_{\mathrm{x}}$ has the dimensions $(N \times B)$ by $(L \times Tx)$, thus a dimension reduction is desired. As in (29), the estimation with dimension reduction by the push-through identity [77] is:

$$\hat{\mathbf{h}}_{(i)} = \hat{\mathbf{h}}_{(1)} + (\mathbf{M}_x^H\mathbf{M}_x + \sigma^2_{v^{+}_{(i)}}(L \times N \times \mathbf{I}))^{-1}\mathbf{M}^H\Delta\tilde{\mathbf{y}}_{(i)} \tag{70}$$

Similar to (30), the channel matrix in the frequency domain can be calculated through the Fourier Transform as:

$$\hat{\mathbf{H}}[k]_{(i)} = \mathcal{F}\{[\hat{\mathbf{h}}_{(i)}, 0, \cdots, 0]\} \tag{71}$$

Then the MSE of the estimation at an iteration i can be defined as [76]:

$$e_{ce_{(i)}} = \mathbf{E}\{(\hat{\mathbf{h}}_{(i)} - \mathbf{h})^2\} \tag{72}$$

The measurement error from CE over L taps is:

$$e_{mea_{(i)}} = L \times e_{ce_{(i)}} \tag{73}$$

As in Section 2.3.2, the equalizer expects knowledge of the channel to estimate transmitted symbols. The CE error affects the performance of the equalizer. Since the CE

error is independent of the transmitted signals, this estimation error can be treated as an additional noise [82]:

$$\sigma^2_{\ddot{v}_{(i)}} = \sigma^2_{v^+_{(i)}} + e_{mea_{(i)}} \tag{74}$$

The CE is based on estimation noise which is related to the SNR. The figure of merit used for communication systems is the energy per bit to noise power spectral density ratio (E_b/N_o, in dB). It is useful to calculate the SNR in terms of E_b/N_o for the purpose of evaluating our system. E_b/N_o is a normalized way to look at SNR, and SNR is normalized to the bit rate [86].

We measure the SNR $= \frac{\mathbf{E}\{\mathbf{x}^H\mathbf{H}^H\mathbf{H}\mathbf{x}\}}{\sigma_v^2}$ at the each receiver. With the assumption of equal power distribution at each path, and the received power at each receiver equals to 1, the E_b/N_o for all antennas is defined as:

$$E_b/N_o = 10\log_{10}\left(\frac{\mathbf{E}\{\mathbf{x}^H\mathbf{H}^H\mathbf{H}\mathbf{x}\}}{\sigma_v^2}\right) - 10\log_{10}\left(R_c \times b\right) - 10\log_{10}\left(\frac{Tx}{Rx}\right) \tag{75}$$

where R_c denotes the code rate of the ECC, b refers to the modulation order (each symbol contains 2 bits in QPSK/4-QAM).

Fig. 3.3 displays the MSE of CE. The lower bound is the ideal case where all transmitted signals are known perfectly at the BS. With $Tx = 10, Rx = 100$ QPSK/4QAM system with a 9 tap multipath fading channel, 2 pilot blocks and 64 data blocks with individual block length of 256 symbols, four iterations are needed to achieve the same MSE as the ideal case at $\sim$ 3.5 dB E_b/N_o. This shows that in the above-described system, when $E_b/N_o \approx 3.5$ dB, the iterative CE performs as good as the ideal case. The more iterations, the sooner the system converges to the lower bound of CE error.

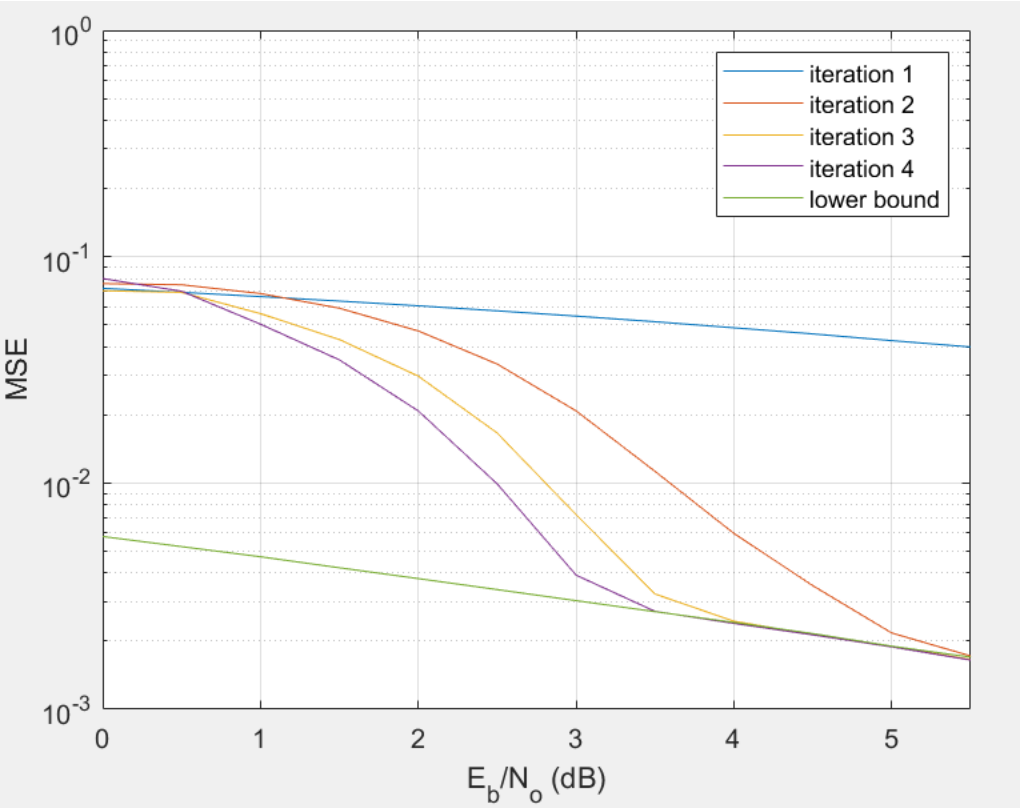

Figure 3.3: Mean square error of the CE versus E_b/N_o over different iteration with Tx=10, Rx=100 and infinite precision quantization

Fig. 3.4 shows BER results for the system described with iterative CE. For comparison, results are also plotted for systems where the CSI is obtained only from pilot signal measurements, for systems where the pilot signals are as long as the data signal (labelled as "All Pilot CE"), for systems where the system has ideal CSI, and for digital communication over a non-fading channel with AWGN (labelled as "AWGN"). Performance in terms of BER for iterative CE gives an almost exact results of using all data for CE. In other words, the result will not significantly improve, as the number of pilot symbols increases, because the iterative CE has already almost achieved the lower bound.

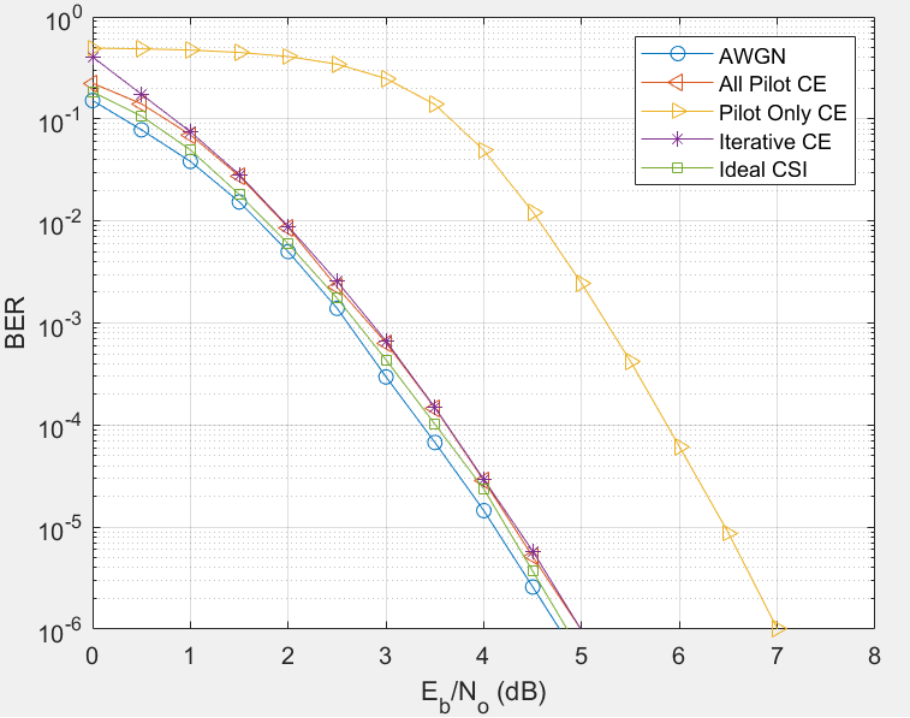

Figure 3.4: BER of CSI in ideal, iterative CE, pilot-only CE, all pilot CE and AWGN over E_b/N_o at Tx=10, Rx=100 with infinite precision quantization

Fig. 3.5 gives results for our system (2 pilot blocks and 64 data blocks) with 10 active users with $R = 50$, 100, and 200 compared with systems ideal CSI, and CSI obtained only from pilot signals. As the graph indicates, performance in terms of BER of the iterative receiver is almost as good as the case where the CSI in perfectly known. In this simulation, each block has a length of 256 symbols.

In Fig. 3.5, from the cases where CE is performed only with pilots, we can find that for a fixed E_bN_o, as the number of receiving antennas increases, the SNR decreases according to (75). The CE error is directly determined by the SNR at receiving antennas [4], as the CE error goes up, the system performs poorly in terms of BER. This phenomenon supports the illustration of importance of CE in [4][20][18]. Without enough knowledge of the CSI, a system cannot exploit the advantages of a large-scale receiving antenna array [20], so the performance gain of Massive MIMO is partially lost.

3.3 Iterative Data Detection

Modern wireless systems must deal with interference between multiple transmitters' signals. There have been several techniques to deal with this problem being proposed [87][34]. Successive Interference Cancellation (SIC), where an iterative receiver uses the estimated signal from one iteration to provide an estimate of the interference for the next iteration for improved data estimation, has been shown to improve performance [41][88]. A problem with this technique is that incorrect decisions made in an early iteration have the possibility of error enhancement in a later iteration leading to instability [34][89]. One solution is to use extrinsic LLR feedback similar to (55) above for SIC.

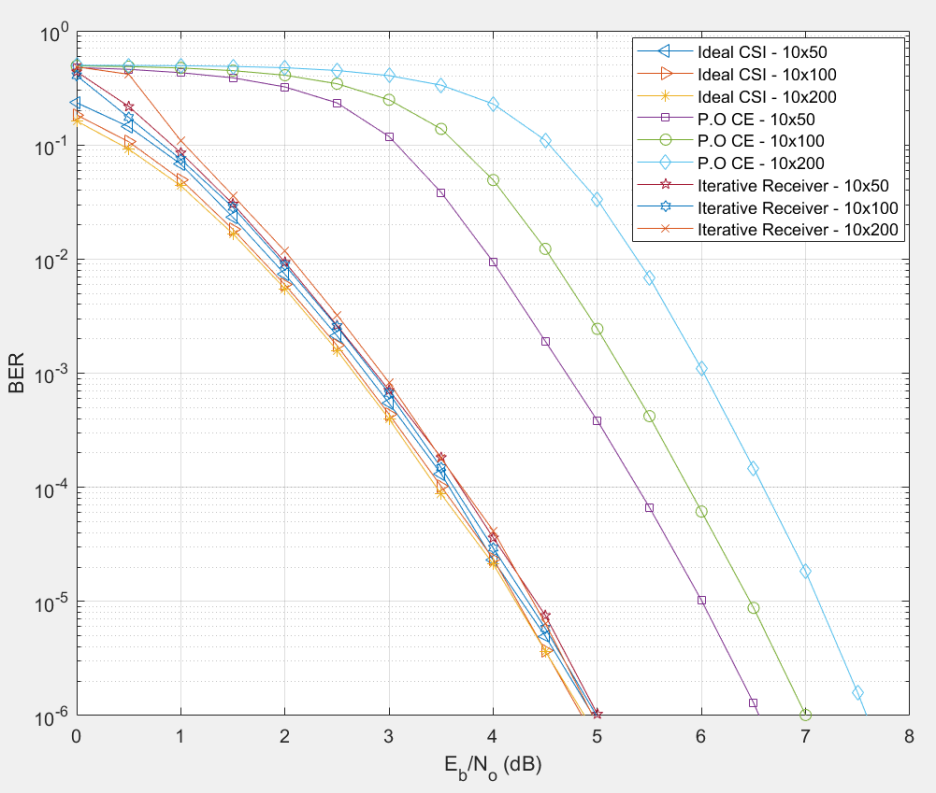

Figure 3.5: Comparison of ideal channel state information, pilot-only CE and iterative Data-aid CE

3.3.1 Iterative Decision Feedback Equalization (IDFE)

The Iterative Decision Feedback Equalization (IDFE), also known as turbo equalization, was first proposed for single antenna wireless systems in [88]. At an iteration $i(i \geq 2, i \in \mathcal{Z})$, we have symbols $\bar{\mathbf{x}}_{(i-1)}$ that are estimated through extrinsic LLR from the preceding iteration $i-1$. Generally, SIC first detects the strongest signals and then removes the resulting estimated interference from the received signal to support a more accurate estimation of the weaker signals [90]. This operation removes the interference generated by each symbol [32]. To remove the previously determined stronger symbols from the last iteration, a SIC is defined in advance of an equalizer:

$$\Delta\tilde{\mathbf{y}}_{(i)} = \tilde{\mathbf{y}}_{(i)} - \hat{\mathbf{H}}_{(i)}\tilde{\mathbf{x}}_{(i-1)} \tag{76}$$

where $\Delta\tilde{\mathbf{y}}_i$ represents difference between received symbols and estimated received symbols, or the remaining relatively weak symbols. Fig. 3.6 is a block diagram of cooperation of equalization and CE, extracted from Fig. 3.1.

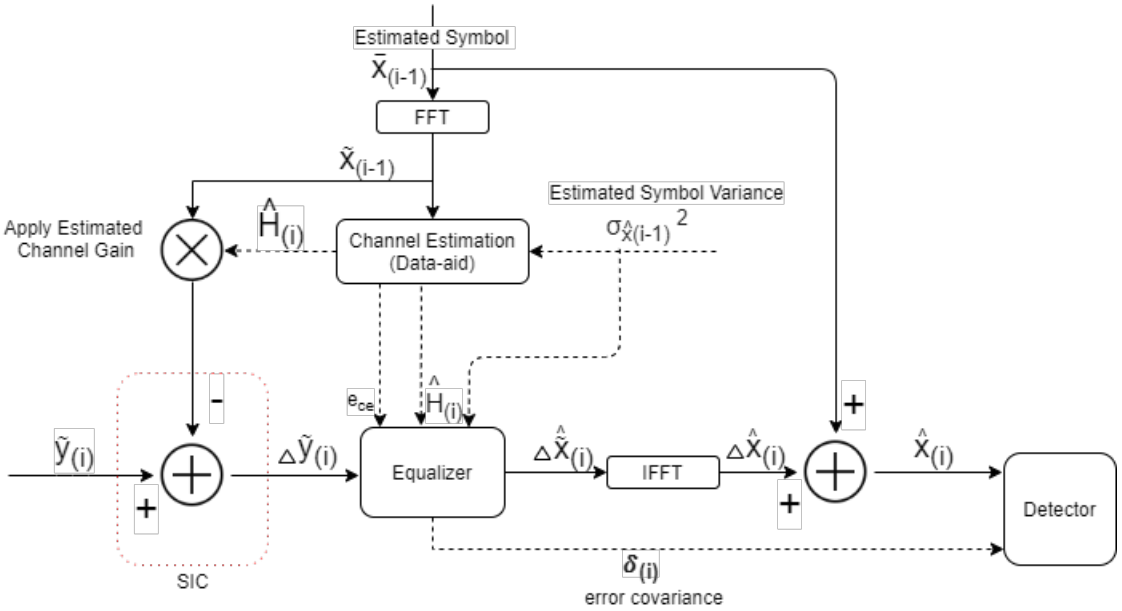

Figure 3.6: Cooperation of channel estimation and equalization in an iterative structure

The extrinsic LLRs, computed as in (55), are fed to the next iteration. The estimated symbol means and variances are computed as (58) and (59) in Section 3.2. The variance of the estimated symbols provides information about the reliability of the results of the previous iteration. The equalization in the following iteration only utilizes the estimated symbols with high reliability to avoid errors of detection in early iterations being repeated in the following iterations. Hence, the equalizer generates decisions of symbols whose reliability increases with each iteration monotonically [36].

Consider the linear equalization filter $\mathbf{G}^{(\mathrm{i})}$ (ZF or LMMSE in this thesis), derived from (35) and (47), the ZF and LMMSE filter at an iteration i are:

$$\mathbf{G}_{zf}^{(i)} = (\hat{\mathbf{H}}_{(i)}^H \hat{\mathbf{H}}_{(i)})^{-1} \hat{\mathbf{H}}_{(i)}^H \tag{77}$$

$$\mathbf{G}_{mmse}^{(i)} = (\hat{\mathbf{H}}_{(i)}^H \hat{\mathbf{H}}_{(i)} + \frac{\sigma_v^2}{\sigma_x^2} \hat{\mathbf{H}}_{(i)})^{-1} \hat{\mathbf{H}}_{(i)}^H \tag{78}$$

The extracted extrinsic LLRs from certain symbols with higher reliability improves the CE results. In a new iteration i, the equalizer performs better with more accurate estimation of the channel, because only the high-reliability symbols are used for CE. Consequently, the heavy interference within the residual symbols $\Delta\tilde{\mathbf{y}}_{(i)}$ can be mitigated, due to the improved estimation of the channel being applied in the equalizer. The equalized residual symbols $\Delta\hat{\tilde{\mathbf{x}}}_{(i)}$ are calculated through the equalizer as:

$$\Delta\hat{\tilde{\mathbf{x}}}_{(i)} = \mathbf{G}^{(i)}_{zf/mmse}\Delta\tilde{\mathbf{y}}_{(i)} \tag{79}$$

Whilst, the error covariance matrix $\delta^{(i)}_{ZF/MMSE}$ of the equalized symbols can be calculated similarly to (41) and (48). This updated reliability will be used in the symbol detector as described in Section 3.1. After applying the equalization filter, by adding back the subtracted mean symbols $\bar{\mathbf{x}}_{(i-1)}$ in the time domain, the equalized time domain symbol vector $\hat{\mathbf{x}}_{(i)}$ can be expressed as:

$$\hat{\mathbf{x}}_{(i)} = \bar{\mathbf{x}}_{(i-1)} + \Delta\hat{\mathbf{x}}_{(i)} \tag{80}$$

To show the performance of our data estimation/detection/decoding system, we simulate a $Tx = 10, Rx = 100$ QPSK/4QAM system with a 9 tap multipath fading channel, 2 pilot blocks and 64 data blocks with an individual block length of 256 symbols. Fig. 3.7 shows the BER performance of a rate $\frac{1}{2}$ convolutional code associating with the iterative receiver. The first iteration does not provide an optimal BER since the CE from the small number of pilot blocks leads to poor equalizer, detector, and decoder performance. Yet, starting at the second iteration, as the CE improves, the system performance is improved significantly. Eventually, the BER curves converge to a low value of 10^{-6} for sufficiently high E_b/N_0 values.

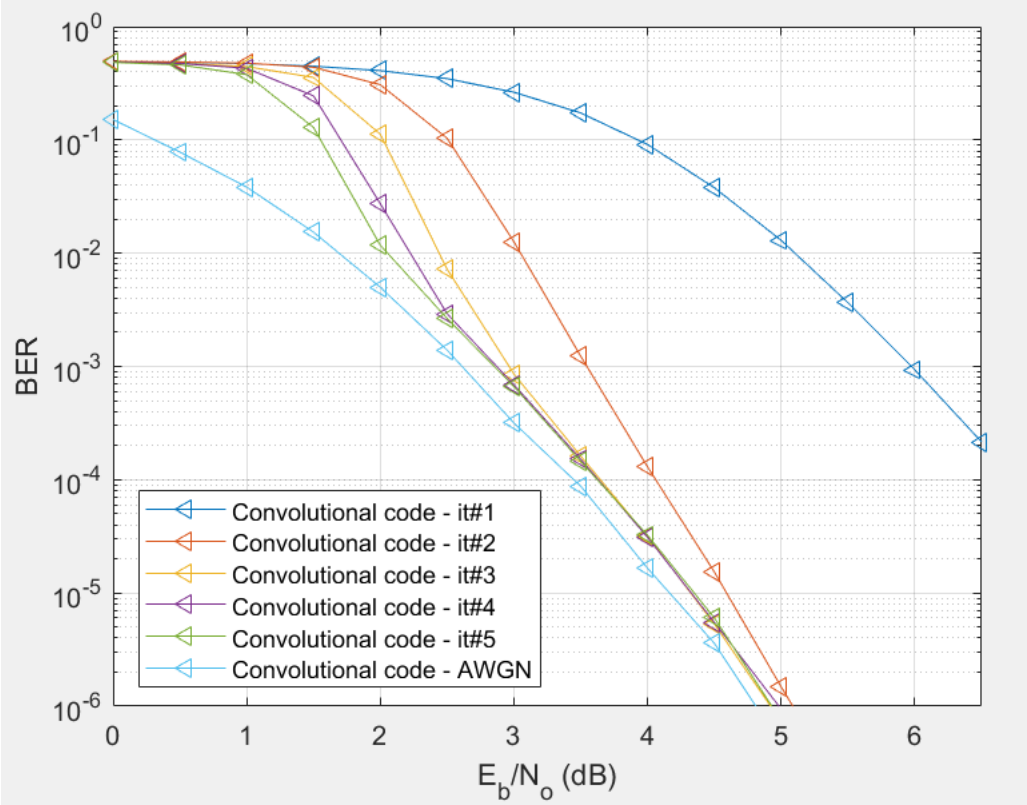

Figure 3.7: BER of an IEEE 802.11 standard R_c =1/2 convolutional code (block length = 16378) with iterative receiver (no quantization)

3.3.2 Comparison between ZF and LMMSE equalization

Estimated signals (equalized signals) are derived from LMMSE and ZF algorithm from previous section:

$$\begin{aligned} \hat{\tilde{\mathbf{x}}}_{zf} &= \mathbf{G}_{zf}\tilde{\mathbf{y}} \\ \hat{\tilde{\mathbf{x}}}_{mmse} &= \mathbf{G}_{mmse}\tilde{\mathbf{y}} \end{aligned} \tag{81}$$

Note that $\tilde{\mathbf{y}}$ denotes the frequency domain reception vector, also known as the observation vector.

Some equalization algorithms such as LMMSE cause inconsistency between the power of the estimated signals and the power of true signals [91], which migrates the symbol values from their actual values and leads to a failure in the detector due to inaccurate symbol values. Normalization makes the mean value of the estimation signal power

equal to the true signal power. To normalize the equalized signals, an inverse of gain is often required on equalized symbols after the equalizer.

Equalizer	Gain	Complexity	Normalization Factor
ZF	$[Diag(\hat{\mathbf{H}}^H\hat{\mathbf{H}})]^{-1}\hat{\mathbf{H}}^H$	$O(Tx^2)$	1
LMMSE	$\sigma_x^2(\hat{\mathbf{H}}^H\hat{\mathbf{H}}\sigma_x^2+\sigma_v^2\mathbf{I})^{-1}\hat{\mathbf{H}}^H$	$O(Tx^3)$	$\text{Gain}_{\text{mmse}}\hat{\mathbf{H}}$

Table 3.2: ZF and LMMSE normalization comparison

To evaluate the performance in terms of the gains, we apply SVD ($\hat{\mathbf{H}}$=$\mathbf{UDV}^H$, where $\mathbf{U}$ and $\mathbf{V}$ are orthonormal matrices and $\mathbf{D}$ is the so-called diagonal matrix with the non-negative real singular values of $\hat{\mathbf{H}}$ on the main diagonal) to both equalization gains of ZF (34) and LMMSE (48) with dimension reduction:

$$\begin{aligned}\mathbf{G}_{zf} &= (\hat{\mathbf{H}}^H\hat{\mathbf{H}})^{-1}\hat{\mathbf{H}}^H \\ &= (\mathbf{VD}^H\mathbf{DV}^H)^{-1}\mathbf{VD}^H\mathbf{U}^H \\ &= [\mathbf{V}(\mathbf{D}^H\mathbf{D})\mathbf{V}^H]^{-1}\mathbf{VD}^H\mathbf{DV}^H\end{aligned} \tag{82}$$

$$\begin{aligned}\mathbf{G}_{mmse} &= (\hat{\mathbf{H}}^H\hat{\mathbf{H}}+\mathbf{I}\frac{\sigma_v^2}{\sigma_x^2})^{-1}\hat{\mathbf{H}}^H \\ &= (\mathbf{VD}^H\mathbf{DV}^H+\mathbf{I}\frac{\sigma_v^2}{\sigma_x^2})^{-1}\mathbf{VD}^H\mathbf{U}^H \\ &= [\mathbf{V}(\mathbf{D}^H\mathbf{D}+\mathbf{I}\frac{\sigma_v^2}{\sigma_x^2})\mathbf{V}^H]^{-1}\mathbf{VD}^H\mathbf{U}^H\end{aligned} \tag{83}$$

From above expressions, the only difference between the two algorithms is the extra term — $\mathbf{I}\frac{\sigma_v^2}{\sigma_x^2}$ — in LMMSE gain. If one of the non-zero values of $\mathbf{D}$ is small, the inversion $\left(\mathbf{D}^H\mathbf{D}\right)^{-1}$ can have a very large value on its mean main diagonal leading to a large amplification of noise in the equalization. The extra term in the LMMSE case $\mathbf{D}^H\mathbf{D}+\mathbf{I}\frac{\sigma_v^2}{\sigma_x^2}$ reduces the impact of this problem.

It is assume that each element of $\mathbf{H}$ is i.i.d. In Massive MU-MIMO systems, consider the SVD of $\hat{\mathbf{H}}^H\hat{\mathbf{H}}$, as the number of receiving antennas increases, the values of singular value matrix $\hat{\mathbf{D}}^H\hat{\mathbf{D}}$ increase proportionally [92]. When the number of receiving

antennas significantly exceeds the number of users, the extra term of SNR, $\mathbf{I}\frac{\sigma_{\ddot{v}}^2}{\sigma_x^2}$, becomes numerically insignificant comparing to the diagonal elements of $\hat{\mathbf{H}}^H\hat{\mathbf{H}}$ [7][18]. With a large-scale receiving antenna array, we have:

$$(\hat{\mathbf{H}}^H\hat{\mathbf{H}} + \frac{\sigma_{\ddot{v}}^2}{\sigma_x^2}\mathbf{I})^{-1}\hat{\mathbf{H}}^H \simeq (\hat{\mathbf{H}}^H\hat{\mathbf{H}})^{-1}\hat{\mathbf{H}}^H \tag{84}$$

if and only if $\mathbf{D}^H\mathbf{D} \gg \mathbf{I}\frac{\sigma_{\ddot{v}}^2}{\sigma_x^2}$, which is being said at $Tx \ll Rx$. Because the diagonal dominant matrix $\hat{\mathbf{H}}^H\hat{\mathbf{H}}$ has dominating weight on the equalization, ZF and LMMSE equalizers have similar performance [5][91]. Figs. 3.8 and 3.9 present the MSE and linear SNR of the estimated data signal, respectively. Lower MSE leads to a low noise power, which eventually results in a higher SNR. This is the reason that generally, the LMMSE equalizer produces lower BER than ZF equalizer [75]. However, a Massive MIMO system with IDFE and iterative CE remedies the shortcoming of ZF equalizer.

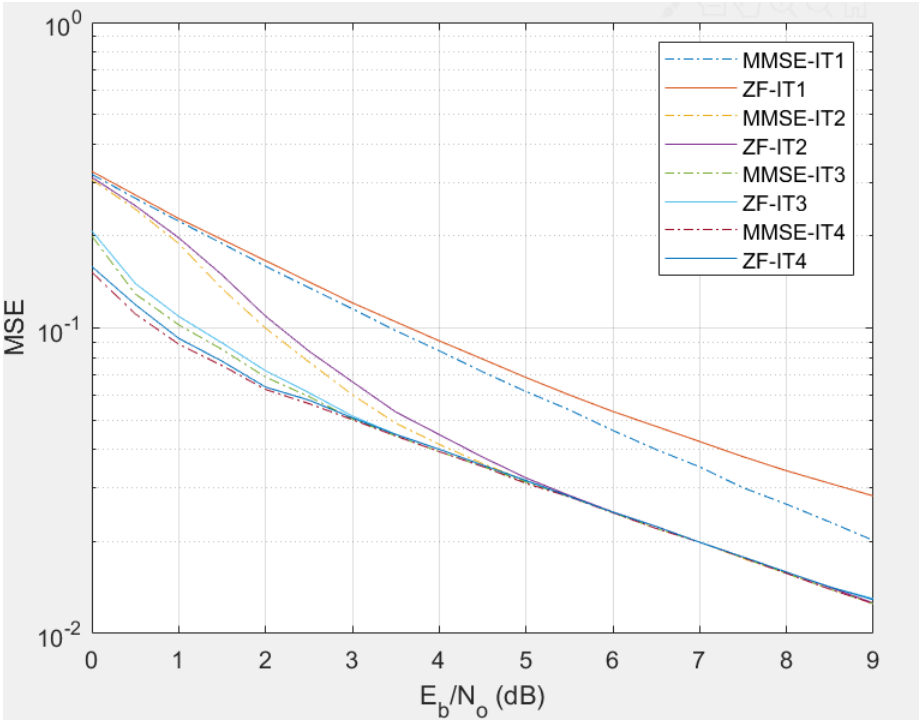

Figure 3.8: Mean square error comparison of iterative ZF and LMMSE

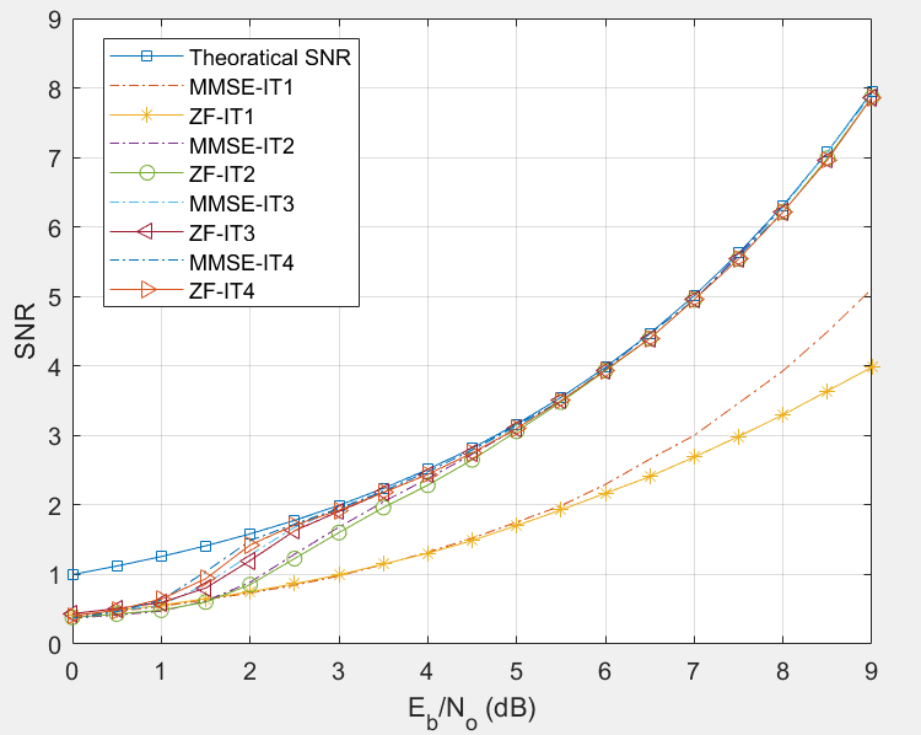

Figure 3.9: Linear SNR comparison versus E_b/N_0 for iterative ZF and LMMSE

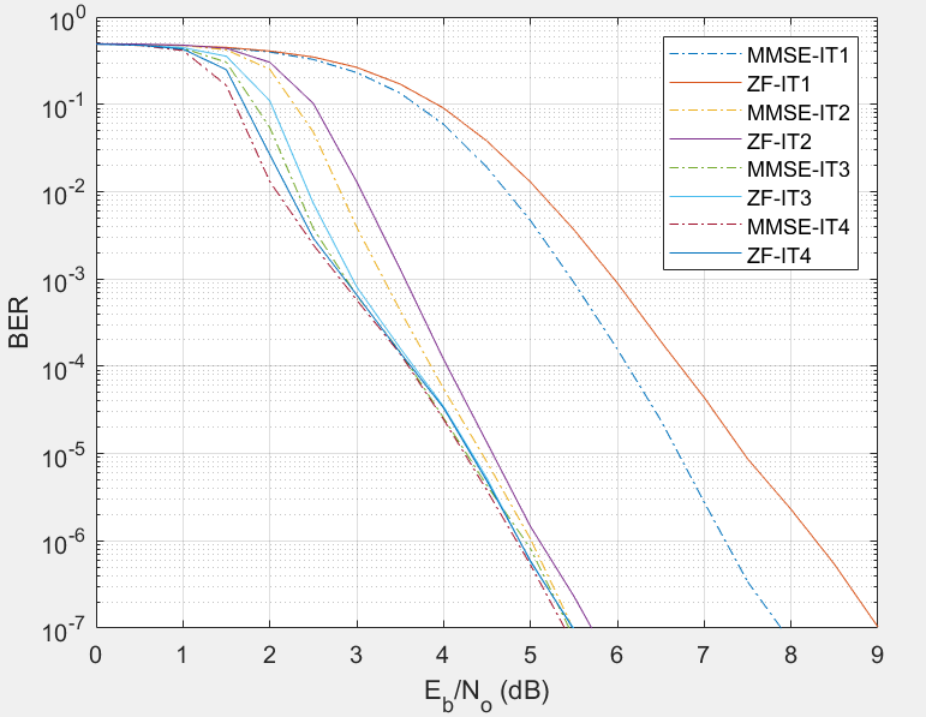

Figure 3.10: BER comparison of iterative ZF and LMMSE

Fig. 3.10 shows the simulation of a 4-QAM system with 10 users (Tx) and 100 receiving antennas (Rx) and ideal quantization. This figure shows that ZF and LMMSE equalizers have similar performance after 4 iterations. After the consideration of computational costs and difficulties of implementations, ZF is the cheaper option with satisfactory performance. Note that in Fig. 3.11, instead of sampling only 4 itera-

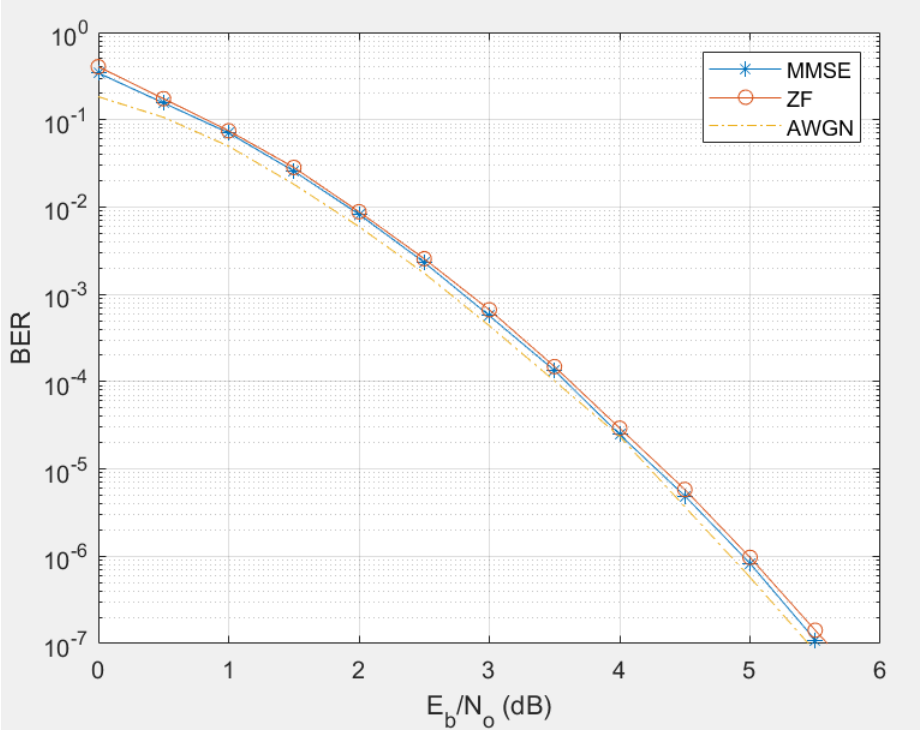

Figure 3.11: ZF and LMMSE final BER performance

BER	$\leq$ 10e-4	$\leq$ 10e-6	$\leq$ 10e-7
ZF	3.70	4.99	4.41
MMSE	3.66	4.98	4.37

Table 3.3: Required E_b/N_o over LMMSE and ZF equalizer at certain BERs, $Tx = 10$, $Rx = 100$

tions as Fig. 3.8 - 3.10, we set a tolerance δ of the MSE of the estimated data signal in between iterations for the system to converge. We define δ as:

$$\delta = |\hat{\tilde{x}}_i - \hat{\tilde{x}}_{(i-1)}|^2 \leq 10^{-6} \tag{85}$$

where $i \geq 2$, $i \in \mathcal{Z}$, $\hat{\tilde{x}}_1$=0.

The iterative system will continuously work until the squared distance satisfies (85), or the system reaches the maximum iteration, which is set to be 30 in our simulations. Table 3.4 shows the number of iteration for ZF and LMMSE to converge to the (85) at a given E_b/N_0 (Maximum iteration is set to be 30). To ensure an accurate result is taken, we run the system until 100 blocks with errors have been seen. As the E_b/N_0 increases, the convergence takes less iterations to reach the capacity until the number of iterations becomes invariant. In addition, no obvious difference in terms of the number of iterations is found between ZF and LMMSE equalizer after E_b/N_0 reaches 3.0 dB.

E_b/N_0	ZF (Mean of iterations)	LMMSE (Mean of iterations)
0	30.0	30.0
0.5	30.0	30.0
1.0	30.0	30.0
1.5	22.7	19.3
2.0	12.6	11.1
2.5	8.3	7.5
3.0	6.2	6.2
3.5	5.0	5.0
4.0	4.2	4.1
4.5	4.0	3.9
5.0	3.2	3.1
5.5	3.1	3.0

Table 3.4: Required iterations for convergence at $\delta = 10^{-6}$ for a given E_b/N_o over LMMSE and ZF equalizer with $Tx = 10$, $Rx = 100$ and infinite precision quantization

4 Performance Evaluation

4.1 Simulation Setup

In a practical Massive MU-MIMO system, the number of receiving antennas at the BS is greater than the number of active users. For the simulation in this thesis, we employed 10 single-antenna active users ($Tx = 10$). A BS with 50, 100 and 200 antennas is simulated ($Rx = [50, 100, 200]$). The modulation constellation of 4-QAM with SC technique is applied.

An IEEE 802.11 standard non-systematic convolutional code with code rate $R_c = 0.5$ with a generator polynomials $(133, 171)$ expressed in octal is used [45]. A block length of 16378 bits is used with this code, the number of blocks for user data is set to 64 ($B = 64$), and the number of pilot blocks is 2 ($P = 2$). For a block length of 8192 bits with a convolutional code, the number of blocks for user data is set to 16 ($B = 16$). A BCJR decoder is applied as the decoding algorithm because of the MAP is preferable for the extrinsic feedback. For a NR-LDPC code with a theoretical code rate $R_c = 0.5$, a block length of 8192 bits with puncturing first 192 bits at each block is used, the number of blocks for user data with 8192 NR-LDPC code is set to 16 ($B = 16$). The sum-product algorithm [93] is used as the LDPC decoding algorithm with 10 decoding iterations to satisfy the parity check matrix.

SC signals have a number of sub-carriers ($T = 256$) equal to the length of the block ($N = 256$). We use a cyclic prefix with 16 samples ($L_{cp} = 16$).

In our simulations, the radio channel between each transmitting and receiving antenna is independent of the radio channel for each other transmitting and receiving antenna. The radio channels are multipath radio channels with 9 taps with delays from 0 to 8 sample periods.

To limit the cost of the radio receivers, coarse quantization (1 to 4 bits) is used.

4.2 BER Results and Analysis

To ensure accurate BER results, for a given set of simulations we run the system until a minimum of 100 blocks with errors have been seen.

4.2.1 BER Performance with Differing Numbers of Receiving Antennas

The BER for a system with $Tx = 10$ transmit antennas and $Rx = 50$ receive antennas is shown in Fig. 4.1. 1-bit quantization has the worst performance, its BER does not drop down to 10^{-6} until 8.92 dB E_b/N_0. The performance improvement from 1-bit to 2-bit quantization is large as 5.98 dB E_b/N_o for 2-bit precision achieves a BER of 10^{-6} compared to 1-bit quantization, indicating an improved efficiency in terms of energy-saving up to 33.0%. In comparison, improvement from 2-bit to 3-bit precision at 10^{-6} BER level is 10.3%, 3 to 4-bit is only 2.05%. The required E_b/N_0 for several BER levels for different quantizations are shown in Table 4.1.

10x50	≤10E-4	≤10E-5	≤10E-6
Ideal CSI+Quan.	3.64	4.36	4.98
CE+Ideal Quan.	3.78	4.44	5.12
CE+4-bit Quan.	3.80	4.48	5.25
CE+3-bit Quan.	3.94	4.70	5.38
CE+2-bit Quan.	4.47	5.31	5.98
CE+1-bit Quan.	6.97	7.96	8.92

Table 4.1: Required E_b/N_o for CE and without CE in 1 to 4 and infinite quantization bits for a given BER at $Tx = 10$, $Rx = 50$

Fig. 4.2 indicates the BER performance with a system of $Tx = 10$ active single-antenna users and $Rx = 100$ receiving antennas at a BS. Table 4.2 illustrates the required E_b/N_0 for each BER level. To reach a BER with a range of 10^{-6}, a system

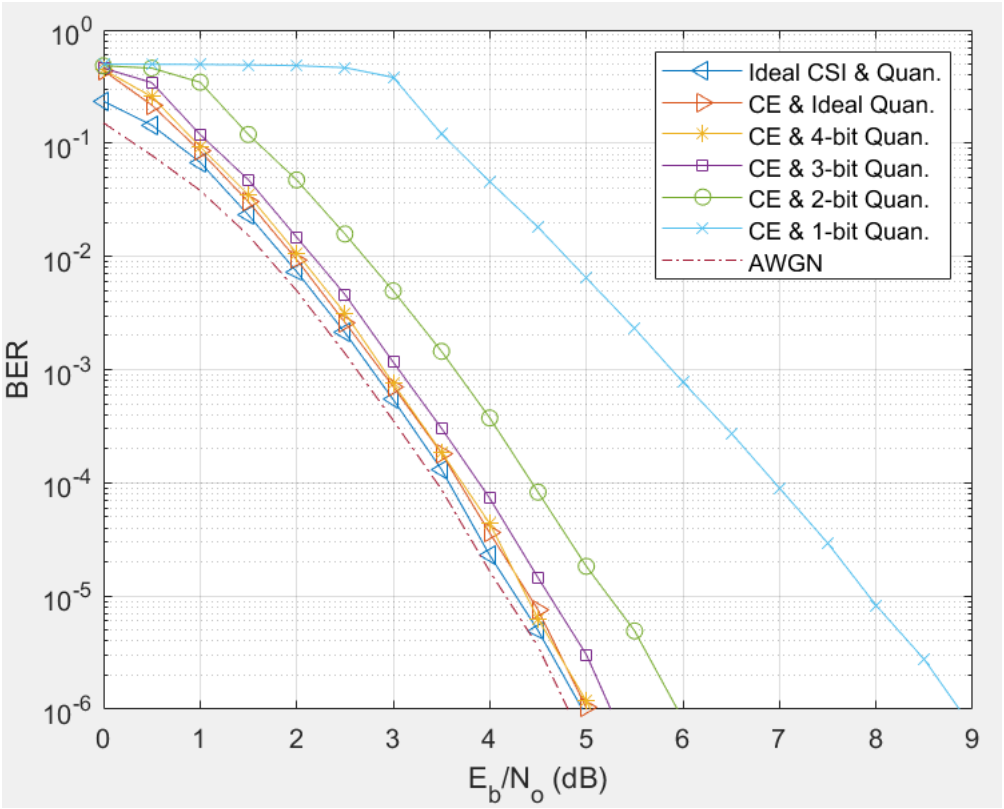

Figure 4.1: BER in regards to E_b/N_o with Tx=10, Rx=50, at 1 to 4 bits and infinite-precision quantization with CE and ideal CSI with infinite-precision quantization

of 1-bit quantizer needs 7.86 dB per bit, 2-bit quantizer requires 5.80 dB per bit, which saves up to 26.2% approximately. Comparing to the improvement of 8.70% from 2 to 3 bits precision, and 0.94% from 3 to 4 bits in the same criteria, we find the improvements is slightly less.

In Fig. 4.3, with $R_x = 200$ well-functional antennas at the BS, an ideal system (receivers with ideal quantization precision and ideal CSI without CE) only improves marginally over systems with $Rx = 200$ antennas. Yet, in terms of the quantization, 1-bit quantizer only requires 7.44 dB per bit for BER$\leq 10^{-6}$, 2-bit quantization will save transmitted energy 23.3%. 3-bit improves saving 9.31% more, and 4-bit improves merely 0.45%.

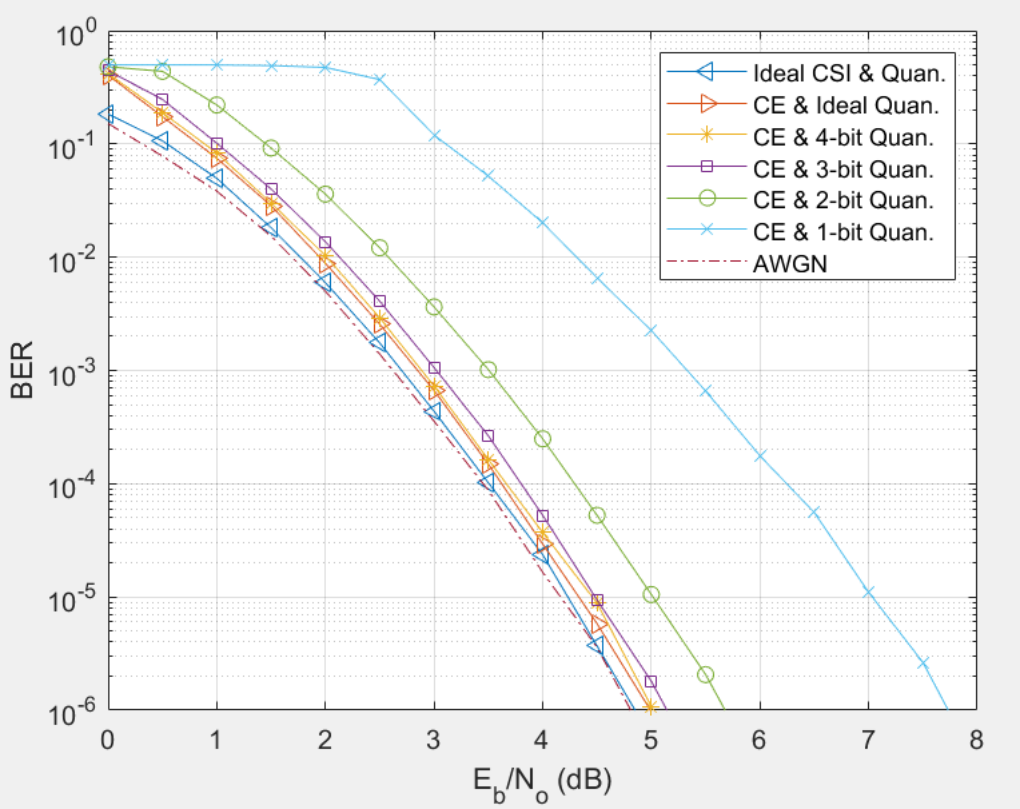

Figure 4.2: BER in regards to E_b/N_o with $Tx = 10$, $Rx = 100$, at 1 to 4 bits and infinite-precision quantization with CE and ideal CSI with infinite-precision quantization

10x100	≤10E-4	≤10E-5	≤10E-6
Ideal CSI+Quan.	3.51	4.34	4.93
CE+Ideal Quan.	3.70	4.41	5.01
CE+4-bit Quan.	3.75	4.48	5.04
CE+3-bit Quan.	3.90	4.50	5.26
CE+2-bit Quan.	4.38	5.03	5.80
CE+1-bit Quan.	6.32	7.06	7.86

Table 4.2: Required E_b/N_o for CE and without CE in 1 to 4 and infinite quantization bits for a given BER at $Tx = 10$, $Rx = 100$

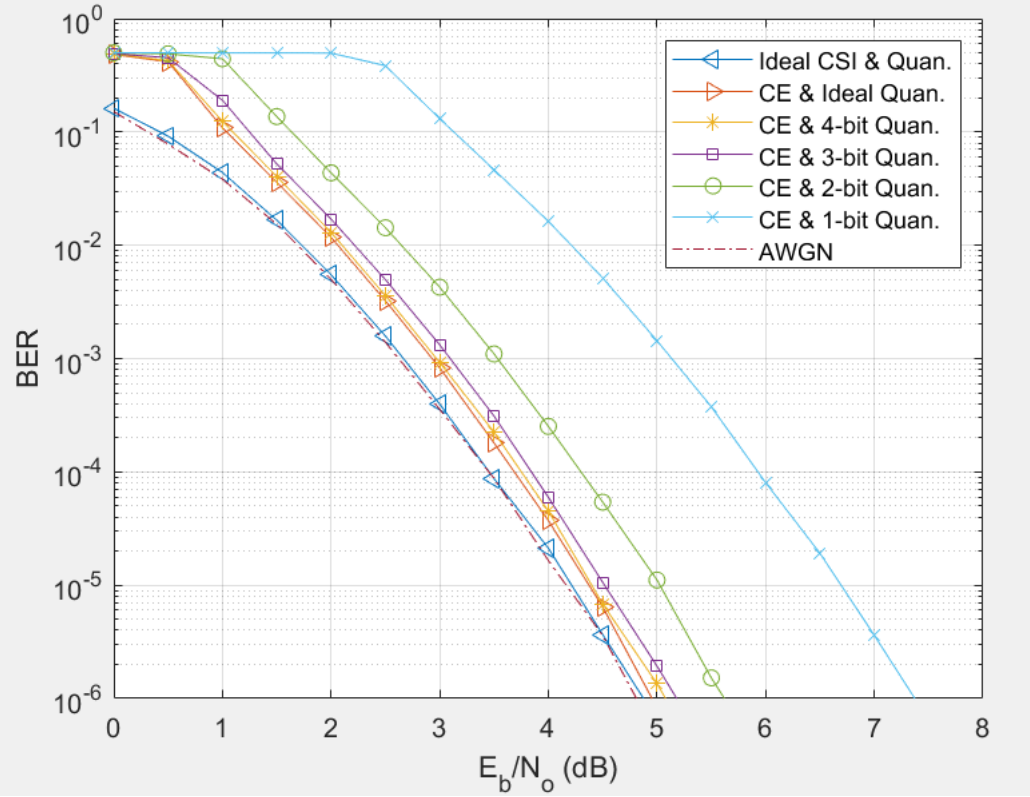

Figure 4.3: BER in regards to E_b/N_o with $Tx = 10$, $Rx = 200$, at 1 to 4 bits and infinite-precision quantization with CE and ideal CSI with infinite-precision quantization

10x200	≤10E-4	≤10E-5	≤10E-6
Ideal CSI+Quan.	3.480	4.32	4.94
CE+Ideal Quan.	3.78	4.44	4.99
CE+4-bit Quan.	3.85	4.46	5.15
CE+3-bit Quan.	3.92	4.54	5.29
CE+2-bit Quan.	4.38	5.06	5.71
CE+1-bit Quan.	5.97	6.79	7.44

Table 4.3: Required E_b/N_o for CE and without CE in 1 to 4 and infinite quantization bits for a given BER at $Tx = 10$, $Rx = 200$

4.2.2 BER Performance with Differing Quantization Precision

Figs. 4.4 to 4.7 show 1-bit to 4-bit quantization on different amounts of receiving antenna respectively.

For the BER of 1-bit quantization, Fig. 4.4 displays a pattern of decreasing BER for all receiving antenna configurations after $\sim$ 2.5 dB of E_b/N_o. Because the quantized signals with the lowest precision suffer the most severe distortion of quantization, higher energy is required at the receiver to compensate for the loss due to the coarse quantization and estimated the accurate symbols. Table 4.3 illustrates the required energy for each BER level at 50/100/200 receiving antenna systems. To achieve the BER of 10^{-6}, $Rx = 50$ requires 8.92 dB, $Rx = 100$ requires 7.86 dB which saves 12.0% of energy, while $Rx = 200$ requires 7.44 dB which saves merely 5.28%.

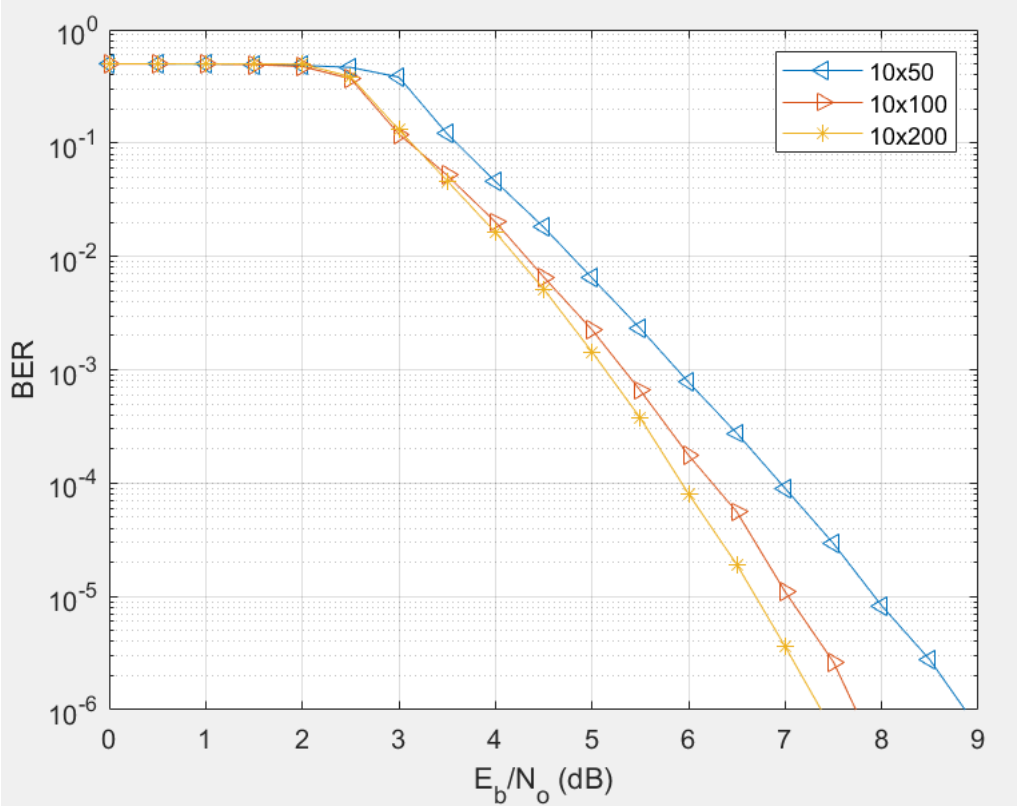

Figure 4.4: BER performance to E_b/N_o with $Tx = 10$, $Rx = 50/100/200$, at 1-bit quantization

Rx	≤10E-4	≤10E-5	≤10E-6
50	6.97	7.96	8.92
100	6.32	7.06	7.86
200	5.97	6.79	7.44

Table 4.4: Required E_b/N_o for 1-bit quantization for a given BER at $Tx = 10$, $Rx = 50/100/200$

As for 2-bit quantization, from both Fig. 4.5 and Table 4.5, at BER=10^{-6}, improvement is 3.01% from $Rx = 50$ to $Rx = 100$, 1.55% from $Rx = 100$ to $Rx = 200$. From the observation of 3-bit and 4-bit quantizations in Fig. 4.6 and Fig. 4.7, the improvement becomes more subtle, especially at 4-bit quantization.

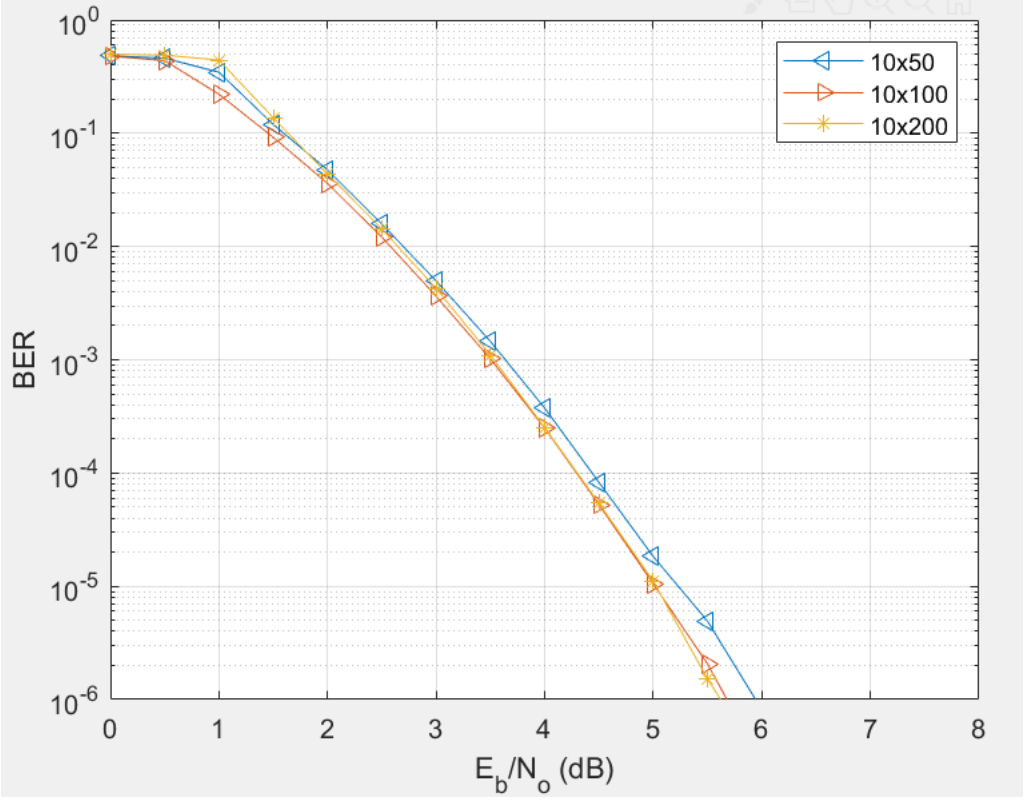

Figure 4.5: BER performance to E_b/N_o with $Tx = 10$, $Rx = 50/100/200$, at 2-bit quantization

Rx	≤10E-4	≤10E-5	≤10E-6
50	4.47	5.31	5.98
100	4.38	5.03	5.80
200	4.33	5.01	5.71

Table 4.5: Required E_b/N_o for 2-bit quantization for a given BER at $Tx = 10$, $Rx = 50/100/200$

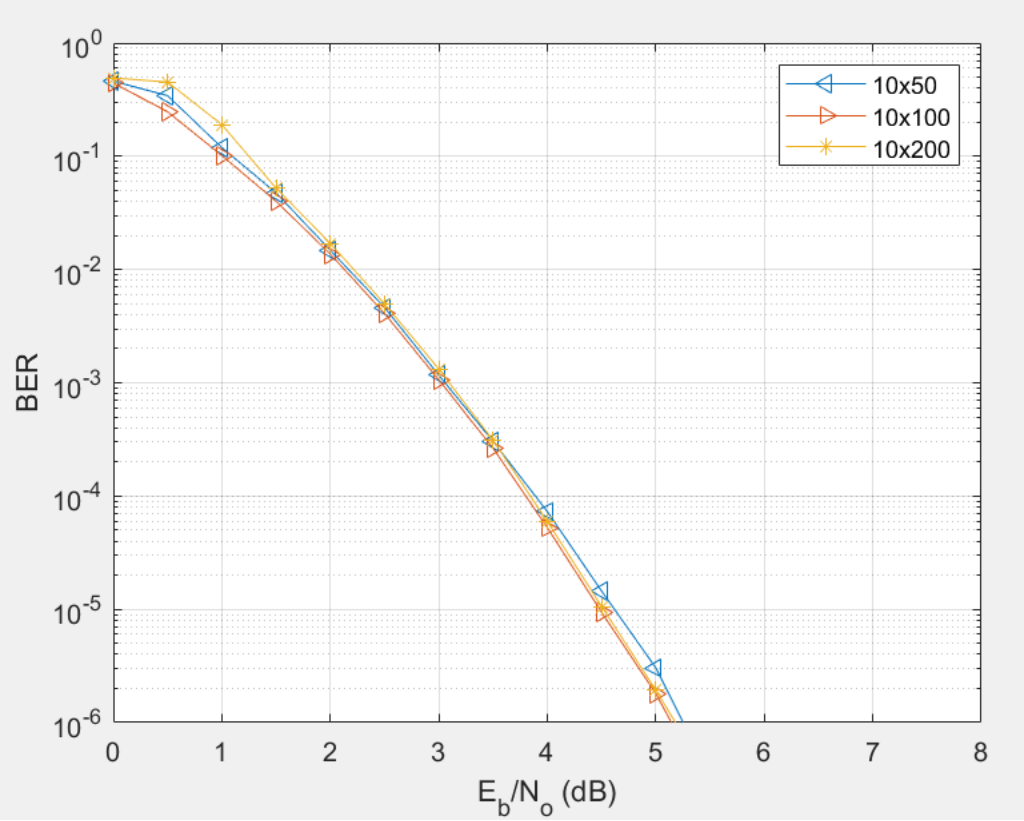

Figure 4.6: BER performance to E_b/N_o with $Tx = 10$, $Rx = 50/100/200$, at 3-bit quantization

To compare different numbers of antennas and different levels of quantization, we compare the E_b/N_0 values required to obtain a BER of 10^{-4}. To obtain accurate BER values, simulations are run until at least 1000 blocks with errors are received. Fig. 4.8 shows that the higher quantization precision we applied, the lower E_b/N_o is required to reach BER=10^{-4}. However, from 4-bit to infinite precision, there is

Rx	≤10E-4	≤10E-5	≤10E-6
50	3.94	4.70	5.38
100	3.89	4.49	5.26
200	3.86	4.48	5.24

Table 4.6: Required E_b/N_o for 3-bit quantization for a given BER at $Tx = 10$, $Rx = 50/100/200$

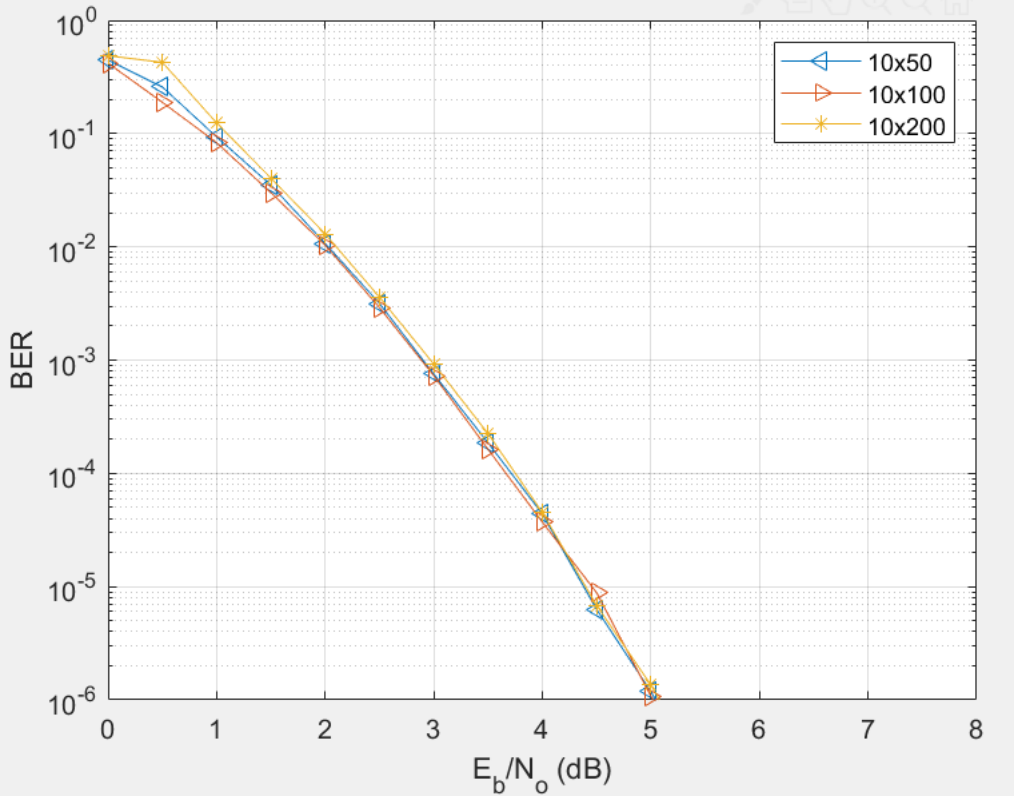

Figure 4.7: BER performance to E_b/N_o with $Tx = 10$, $Rx = 50/100/200$, at 4-bit quantization

not much potential for improvement, which means higher-order quantizations in our system will not make much difference as 4-bit quantization is near-optimal.

Rx	≤10E-4	≤10E-5	≤10E-6
50	3.81	4.48	5.16
100	3.77	4.48	5.16
200	3.75	4.46	5.15

Table 4.7: Required E_b/N_o for 4-bit quantization for a given BER at $Tx = 10$, $Rx = 50/100/200$

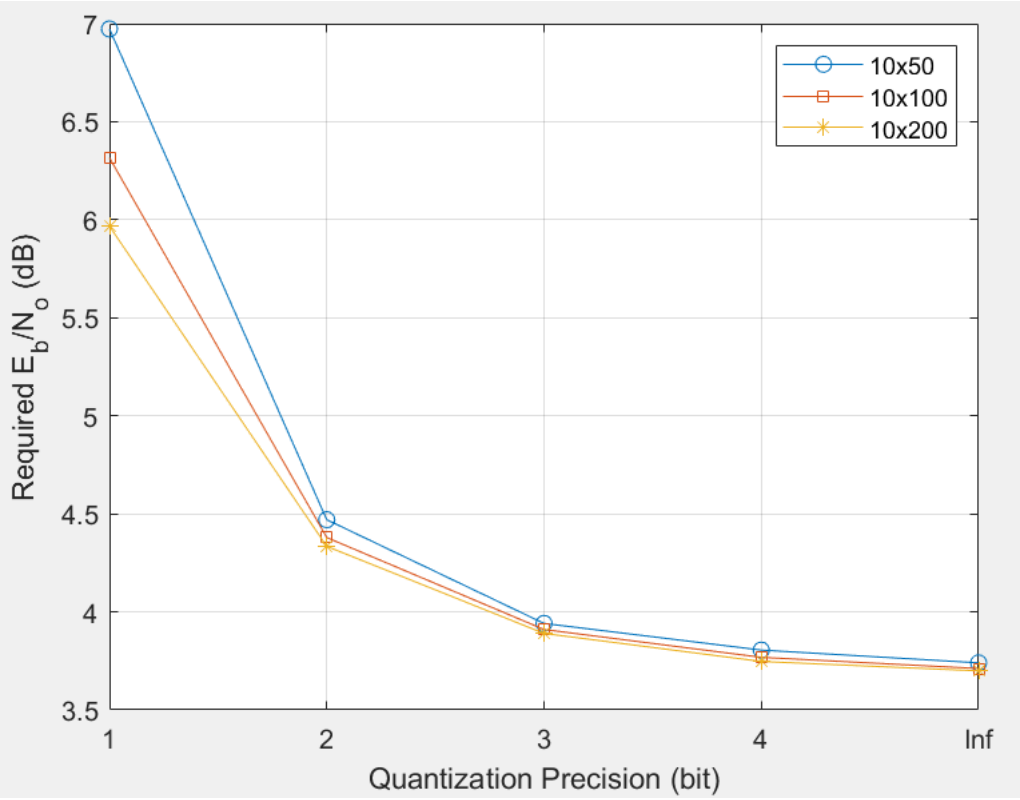

Figure 4.8: Required E_b/N_o for different quantization precision at BER=10^{-4}

4.3 EXIT Chart

An Extrinsic Information Transfer (EXIT) chart is a powerful tool to visualize the convergence of iterative algorithms [41]. It helps us examine and evaluate the performance of detectors and decoders by visual observation of information exchange within each iteration. A EXIT chart is a plot of the mutual information of the exchanging messages between two components. One component's mutual information coefficient

is plotted with its input on the x-axis and output on the y-axis, while the second component's mutual information coefficient is plotted with its input on the y-axis and output on the x-axis. As an iterative structure is implemented at the BS, we use this methodology to measure and analyze the information exchange between the detector and decoder in each iteration.

The EXIT chart was first used to analyze and observe the convergence behavior of iteratively decoding turbo codes [41]. What drives the iterative receiver is that its detector and decoder are both soft-input and soft-output modules that accept and deliver LLRs as values of reliability. The soft-output of the detector is passed to the decoder as a priori input. Similarly, the subtraction of the decoder's soft-output and detector's soft-output is accepted by the detector in a new iteration as the extrinsic information. Here, we will use the EXIT chart method to analyze the information flow between the CE, equalization, data detector system and the data decoder system.

4.3.1 Mutual Information

The EXIT chart is constructed using mutual information. The mutual information is measure at both input and output of the detector and decoder. For each component, an input-to-output mutual information curve can be obtained showing how that component can improve the mutual information of the data estimates with different input mutual information values. This allows the mutual information flow within the receiver to be described. Mutual information is a measure of the dependence between the two variables. Assume X and Y are two real-valued discrete random variables. The mutual information between X and Y is defined as [94]:

$$\mathbf{I}(X;Y) = \sum_{y\in\mathcal{Y}} \sum_{x\in\mathcal{X}} p_{X;Y}(x,y) \log\left(\frac{p_{X;Y}(x,y)}{p_X(x)p_Y(y)}\right) \tag{86}$$

where $p_{X;Y}$ is the joint probability density mass function of X and Y, and p_X and p_Y are respectively the marginal probability density mass functions of X and Y. In our system, a soft-decision detector and decoder produce LLR, we approximate the

continuous LLR variables as discrete variables so that (86) can be used to compute mutual information. We apply the histogram method from [95] to convert continuous variables to discrete variables by distributing a set of continuous variables into different bins. The optimized bin width is formulated based on LMMSE methods as [96]:

$$w = 3.49\sigma \mathbf{n}^{-\frac{1}{3}} = \frac{1}{2}(3.49\sqrt{\sigma^2_{\mathcal{L}_0}}\mathbf{n}_0^{-\frac{1}{3}} + 3.49\sqrt{\sigma^2_{\mathcal{L}_1}}\mathbf{n}_1^{-\frac{1}{3}}) \tag{87}$$

where w denotes the bin width, σ denotes the standard deviation of the LLR vector. ${\sigma_{\mathcal{L}_0}}^2$ and ${\sigma_{\mathcal{L}_1}}^2$ denote the variance of the LLR about 0 and 1 respectively, $\mathbf{n}$ denotes the size of data set. Therefore, the Probability Density Function (PDF) of the soft value $\mathcal{L}_{\hat{X}\pi}$ — $p(l_{\hat{x}\pi})$ is approximated as a histogram.

The probability of 0 and 1 at the receiver are close to $\frac{1}{2}$ in a real system, therefore we assume the coded user bit sequence X is uniformly and randomly distributed with probabilities $p_{(x=0)} = p_{(x=1)} = \frac{1}{2}$. Detected data at a receiver $\mathcal{L}_{\hat{X}}$ are soft values in terms of LLR due to the SISO detector. In addition, the posterior mutual information of a detector $\mathbf{I}_e DeT$ is also a prior mutual information of the decoder $\mathbf{I}_a Dec$, we can be obtain the mutual information through [89]:

$$\mathbf{I}_e Det = \mathbf{I}_a Dec = \mathbf{I}(x; \mathcal{L}_{\hat{X}}) = \mathbf{H}(X) - \mathbf{H}(\hat{X}|\mathcal{L}_{\hat{X}}) \tag{88}$$

where the $\mathbf{H}(\cdot)$ denotes the entropy [97]:

$$\mathbf{H}(X) = -\sum_{i=1}^{I} p(x_i)\log_2 p(x_i) \tag{89}$$

where I is the data length. The conditional entropy from (90) can be purposed as:

$$\mathbf{H}(X|\mathcal{L}_{\hat{X}}) = -\sum_{x\in\mathcal{X}, l\in\mathcal{L}_{\hat{X}}} p(x, l_{\hat{x}})\log_2\left(\frac{p(x, l_{\hat{x}})}{p(l_{\hat{x}})}\right) \tag{90}$$

By subtracting the LLR of coded bits after the decoder $\mathcal{L}_{\hat{X}_o}$, we define the extrinsic LLR $\mathcal{L}_{ex} = \mathcal{L}_{\hat{X}} - \mathcal{L}_{\hat{X}_o}$, then the corresponding extrinsic mutual information is $\mathbf{I}_e Dec$, which is also the prior mutual information passing into the detector in a new iteration $\mathbf{I}_a Det$, is given by [89]:

$$\mathbf{I}_e Dec = \mathbf{I}_a Det = \mathbf{I}(X; \mathcal{L}_{ex}) \tag{91}$$

4.3.2 Analysis of the EXIT Chart

Channel capacity can be defined as the maximal mutual information between the input and output of a channel [98].

$$Capacity = \max_{p_{\mathcal{X}}(x)} \mathbf{I}(\mathbf{X}_{transmit}; \mathbf{Y}_{receive}) \tag{92}$$

Theoretically, as E_b/N_o increases, channel capacity increases [98]. However, the performance of the iterative receiver is also restricted by the decoder, as in decoding algorithm, different ECCs and ECC properties such as block length and code rate. To make sure the feedback information is extrinsic, we can examine the performance of a detector and decoder from an EXIT chart. An EXIT chart contains plots of the input to output mutual information curves of the CE/detector sub-system as well as the input to output mutual information curves of the decoder subsystem. For the CE/detector system, input LLR values are generated with a controlled level of mutual information with random data sequences and a single iteration of CE and detection is performed and the output LLRs mutual information with the data computed. By averaging over several independent trials, a curve of output mutual information versus input mutual information is developed. The EXIT curve of the CE system is dependent on E_b/N_0 and channel conditions.

In Fig. 4.9, mutual information flow of detector in 0 to 6 dB is a given. The X-axis $I_aDet = IeDec$, shows the posterior information after the detector or a priori information of the decoder. The Y-axis $\mathbf{I}_aDec = \mathbf{I}_eDet$ represents the extrinsic information which is fed into the detector in the next iteration. A detector curve represents that given a prior information $\mathbf{I}_aDet$, information of $\mathbf{I}_eDet$ is provided by the current detector. The decoder curve represents given a prior information $\mathbf{I}_aDec$ from the detector, an extrinsic information of $\mathbf{I}_eDec$ is produced by the decoder. To evaluate the performance of an ECC, we locate a crossing point with the detector curve at the first knee of the decoder curve. For example, a rate $\frac{1}{2}$ convolutional code with block length 16378 has crossed a detector curve of 1 dB, which means that at 1 dB, the

extrinsic information is contaminated with detector error, therefore, some errors and distortions cannot be eliminated by iterative equalization due to error enhancement (discussed in Section 3.3.2).

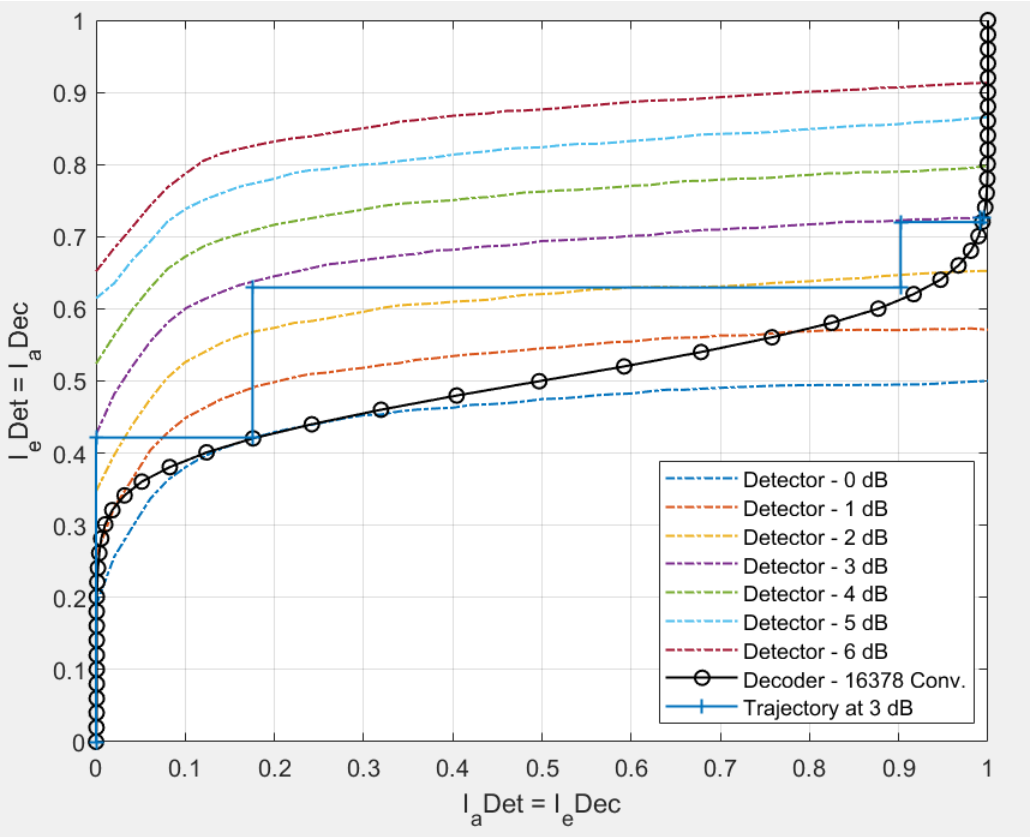

Figure 4.9: EXIT chart for R_c=1/2 IEEE 802.11 standard convolutional code with block length 16378 at 0 to 6 dB E_b/N_o with infinite precision quantization

In the same graph, a trajectory is plotted as the straight lines at $E_b/N_o = 3$ dB. In the first iteration, there is no extrinsic information available, therefore $\mathbf{I}_E Dec = 0$. while the detector is able to exploit mutual information from the detected signals and transmitted signals for ~ 0.42 bit. At this moment, the first vertical segment from coordinate $(0, 0)$ to $(0, 0.42)$ is plotted. With 0.42 bit mutual information from the detector, the extrinsic information is purposed after the decoder, which is ~ 0.18 bit. Thus the first horizontal segment from coordinate $(0, 0.42)$ to $(0.18, 0.42)$ is plotted. In the new iteration, the 0.18 bit extrinsic information will feed into the detector which will provide better knowledge for the decoder until the detector reaches its

limit of any given E_b/N_o. The number of required iterations can be obtained by plotting the trajectory over detector and decoder.

We also applies 5G-NR-LDPC code with block length 8192 associating with the iterative decision feedback receiver.

Fig. 4.10 gives the EXIT curves of LDPC and convolutional code with the same block length. Their corresponding BER performance is available in Fig. 4.11, where a sum-product algorithm [93] was utilized for LDPC code as a decoding algorithm. As we can see, at 0 and 1 dB E_b/N_o, the detector curves cross the decoder curves of both ECCs, this illustrates that the insufficient feeding back extrinsic information will not ignite the iterative system. However, for instance, at $E_b/N_o = 3$ dB, the detector curve does not cross the decoder curve until the very end ($\mathbf{I}_A Det = \mathbf{I}_E Dec = 1$), which means 3 dB E_b/N_o is enough in our simulation to ensure the feedback being extrinsic. In Fig. 4.11, we can see the 8192 LDPC BER curve starts to drop significantly at ~ 2.5 dB in a Multipath Fading Channel (MFC).

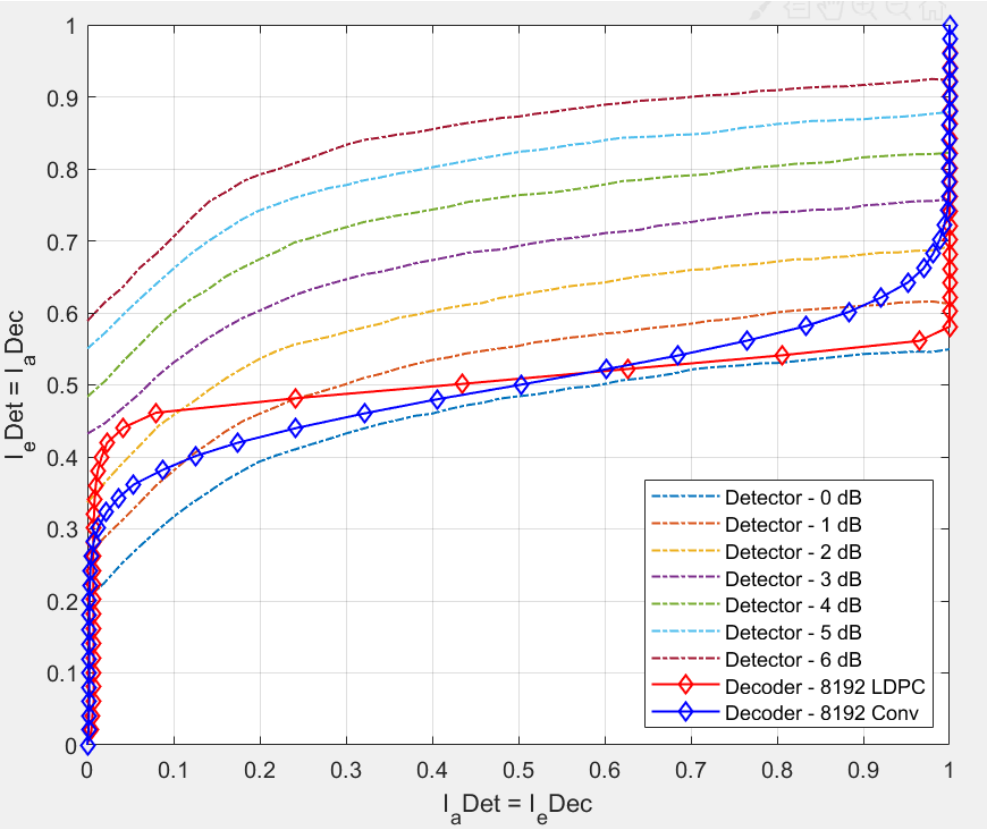

Figure 4.10: EXIT chart for R_c=1/2 5G-NR-LDPC code with block length 8192 at 0 to 6 dB E_b/N_o with infinite precision quantization

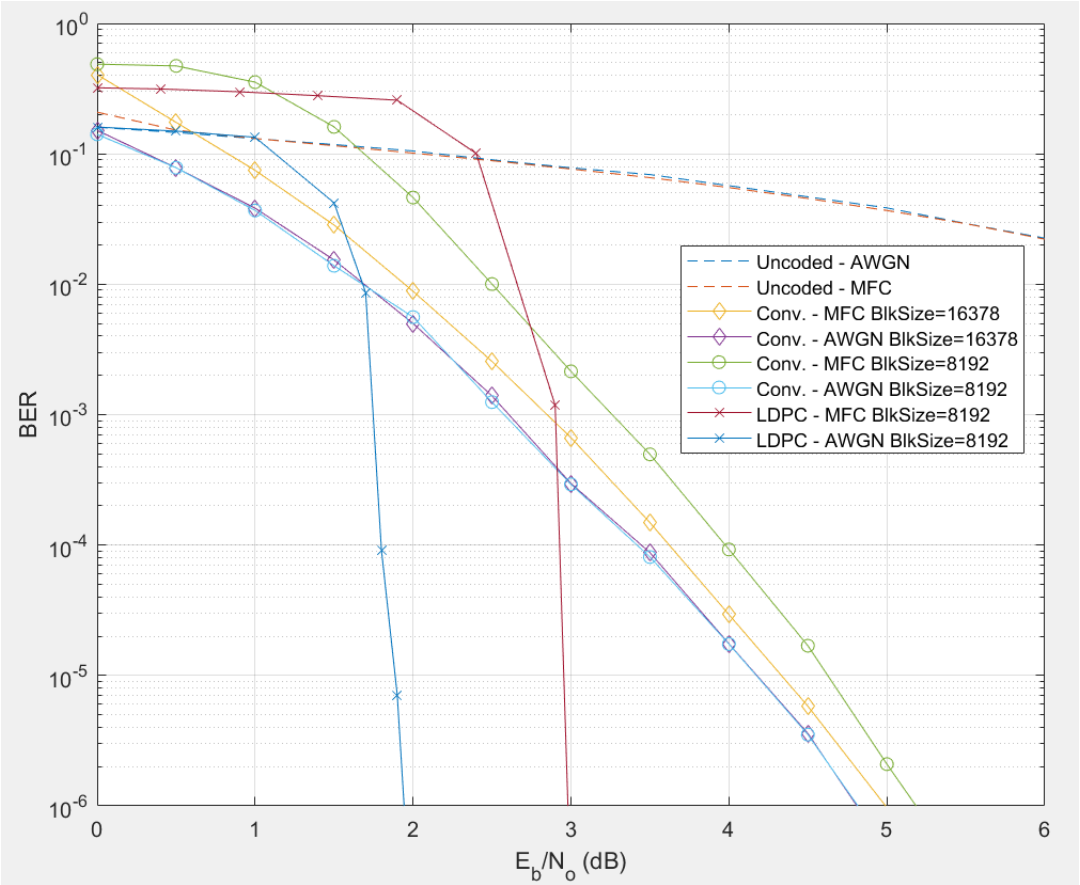

Figure 4.11: BER comparison of IEEE 802.11 R_c=1/2 convolutional code (block length = 16378 and 8192) and a R_c=1/2 5G-NR-LDPC Code (block Length = 8192) in AWGN and MFC with infinite precision quantization

As for a system with coarse quantization, Fig. 4.12 (A) and 4.12 (B) show the EXIT chart for 1-bit to 4-bit and infinite precision quantization at 2 and 5 dB respectively. At 2 dB E_b/N_o, the quantization is not good enough for the system to produce a satisfactory BER of 10^{-6}, because even in the system with infinite precision quantization, the detector curve still crosses the decoder curve before it reaches the maximum mutual information. However, in Fig. 4.12 (B), 2-bit quantization is enough for the iterative receiver at 5 dB E_b/N_o to converge to a high mutual information, indicating a low output BER. In addition, it is obvious that the detector curves of different numbers of quantization bits in both Fig. 4.12 (A) and (B) tend to get close to the infinite precision quantization curve. The increment between 1-bit to 2-bit quantization is relatively huge comparing to the higher-order precision. With 4-bit quantization, the gap with infinite precision is small. This is further evidence that supports Fig. 4.8. As the quantization precision increases, the improvement becomes less significant compared to infinite precision. As a result, only coarse quantization is required in our system.

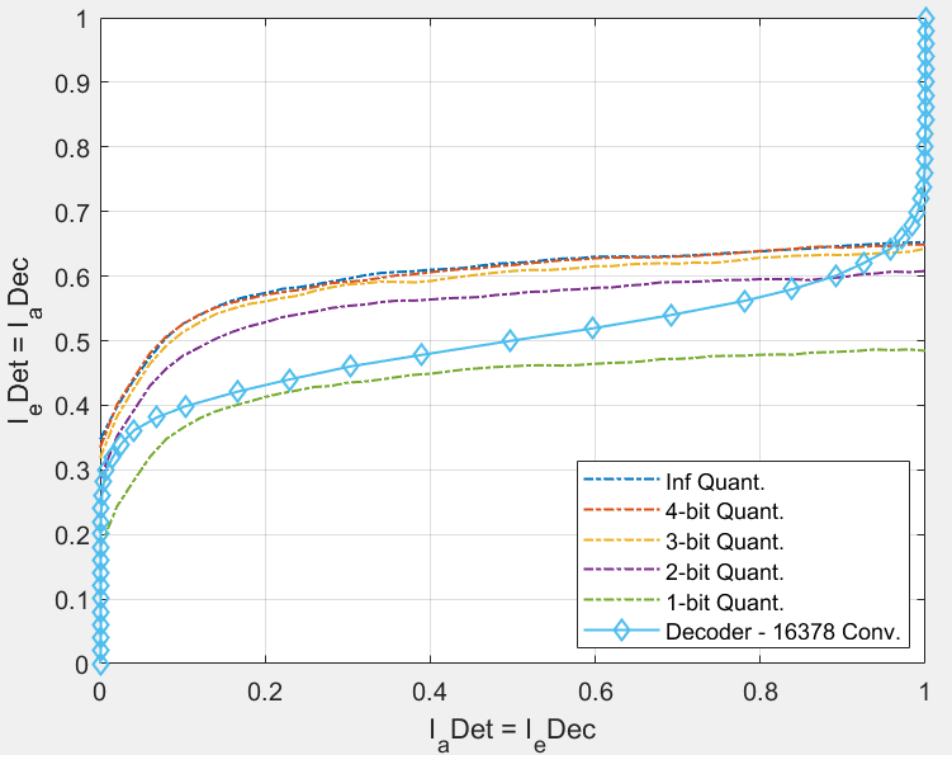

Figure 4.12 (A): EXIT chart for 1-bit to 4-bit and infinite precision quantization at E_b/N_0=2 dB

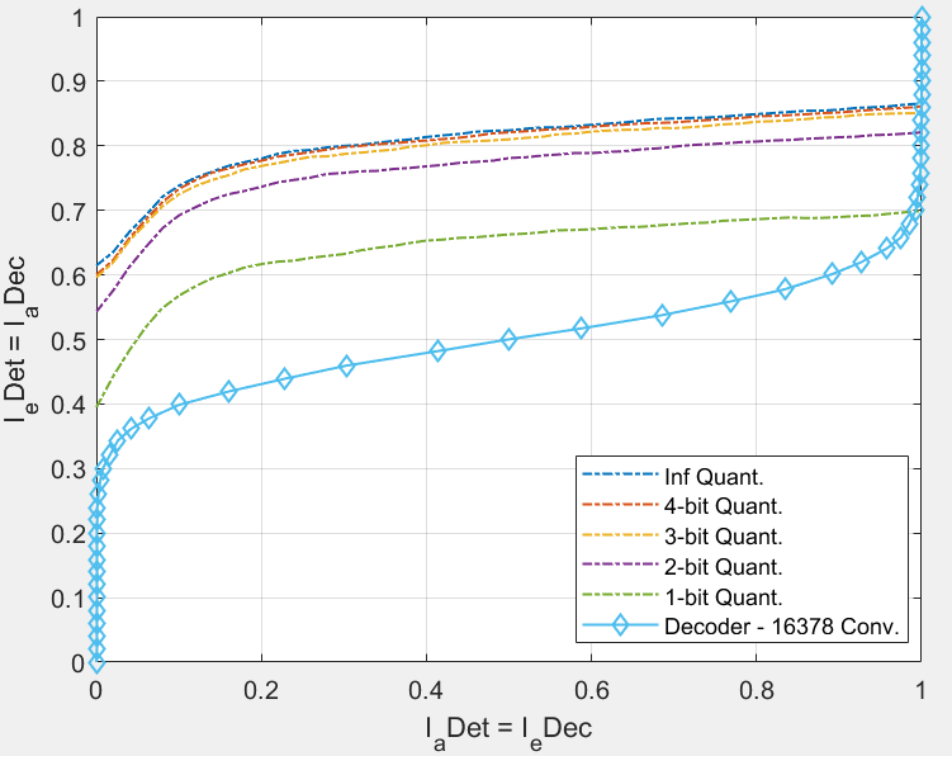

Figure 4.12 (B): EXIT chart for 1-bit to 4-bit and infinite precision quantization at E_b/N_0=5 dB

5 Conclusion and Future Work

5.1 Conclusion

This thesis examined an Iterative Decision Feedback Receiver for a Massive MU-MIMO uplink system. For modern wireless communication, RF quantization and computation require the most energy, which always leads to a cost-efficiency trade-off between these two modules in implementation. Conventional Massive MU-MIMO integrates the most power-consuming processes at the BS so that mobile users can be released from expensive and battery-draining devices. The BS equipped with a large number of antennas which requires a great amount of energy for high complexity computations for channel estimation, data detection, and data decoding. The higher requirements for precision, the higher the requirements for more complicated and expensive RF front-ends at the receiving antennas. Our iterative receiver for a Massive MU-MIMO uplink system lowers the requirement for quantization precision, so that cheap and simple ADCs can be applied without significant performance loss in terms of BER at the receiver.

Through our simulations, we found that iteratively feeding back extrinsic information improves the accuracy of CE and data detection. For CE, only a small number of pilot blocks is required for initial CE, because CE results can be improved iteratively by using the data symbols from user as additional "pilots", which indirectly mitigated the pilot contamination effect without lowering the data rate. Essentially, as the random user data sequence are quite long, the probability that two such data sequences are close to being orthogonal is very high. Accurate CSI leads to a more accurate data detection process. Coarse quantization generally gives rise to heavy distortion in terms of information loss that is present in cheaper radio receivers. As our simulation suggested, with low transmission power, the iterative receiver provides good BER performance, even though a small number of pilots is used. This iterative structure will allow increased data rates which low cost hardware and low power consumption.

5.2 Future Work

In the future, our first suggestion for future work is to investigate different CE schemes, such as the blind and/or semi-blind CE. In this thesis, a linear MMSE estimator provides the initial estimations of the channel. Non-linear estimators, such as the Least Square Estimation algorithm, will commonly give a closer estimation with extra computational complexity without consuming extra data rate for pilot transmission. A better initial estimation might result in a lower requirement for iteration, where the efficiency can be improved.

Secondly, a Massive MU-MIMO downlink precoding can be considered in future work. Precoding techniques at the BS can simplify the CE and data detection at user terminals. With the assistance of iterative structure at the BS, the performance and complexity relation of precoding methods worth for further discussion.

Lastly, a wireless system with relay BS in both an uplink and a downlink system can be examined in the future, where the relay acts as a connection between the BS and user terminals who suffer heavy multipath propagation.

Bibliography

[1] C. Boano. Hot packets: A systematic evaluation of the effect of temperature on low power wireless transceivers. *in 5th Extreme Conference on Communication, Thorsmork, Iceland*, August 2013.

[2] B. M. Adnan and F. M. Alamgir. Performance simulation and comparision in High Altitude Platforms (HAPs) communications systems under PSK, DPSK, QAM FSK modulation schemes and AWGN, Rician and Rayleigh communication channels. In *2016 IEEE 7th Annual Information Technology, Electronics and Mobile Communication Conference*, pages 1–11, October 2016.

[3] B. Sklar. *Digital Communications: Fundamentals and Applications (2nd Edition)*. Prentice Hall, December 2006.

[4] M. Biguesh and A. B. Gershman. Training-based MIMO channel estimation: a study of estimator trade-offs and optimal training signals. *IEEE Transactions on Signal Processing*, 54(3):884–893, March 2006.

[5] T. L. Marzetta. Massive MIMO: An introduction. *Bell Labs Technical Journal*, 20:11–22, March 2015.

[6] A. Chockalingam and B. Sundar Rajan. *Large MIMO Systems*. Cambridge, Feburary 2014.

[7] T. L. Marzetta. Noncooperative cellular wireless with unlimited numbers of base station antennas. *IEEE Transactions on Wireless Communications*, 9(11):3590–3600, November 2010.

[8] H. Lipfert. *MIMO OFDM Space Time Coding – Spatial Multiplexing, Increasing Performance and Spectral Efficiency in Wireless Systems, Part 1 Technical Basis (Technical report)*. Institut für Rundfunktechnik, August 2007.

[9] T. L. Marzetta, T. G. Larsson, H. Yang, and H. Ngo. *Fundamentals of Massive MIMO*. Cambridge University Press, December 2016.

[10] C. Masouros, C. Jianling, K. Tong, M. Sellathurai, T. Ratnarajah, and W. Junhong. Large scale antenna arrays with increasing antennas in limited physical space. *China Communications*, 11(11):7–15, November 2014.

[11] Y. Peng and J. Jung. *Guidelines for evaluation of radio interface technologies for IMT-2020. Revision 1 to Document 5/57-E. [accessed 2017 December 24].* October 2017.

[12] R. Chakraborty, N. Kumari, M. Mousam, and A. Mukherjee. The future of 5G and millimeter waves. In *2018 Second International Conference on Electronics, Communication and Aerospace Technology*, pages 1679–1683, March 2018.

[13] S. Yang and L. Hanzo. Fifty years of MIMO detection: The road to large-scale MIMOs. *IEEE Communications Surveys Tutorials*, 17(4):1941–1988, September 2015.

[14] I. Telatar. Capacity of multiple antenna gaussian channels. *Eur. Trans. Telecommun.*, 10(6):585–595, September 2008.

[15] H. Rutishauser. Handbook series linear algebra: The Jacobi method for real symmetric matrices. *Numerische Mathematik*, 9(1), November 1966.

[16] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson. Scaling up MIMO: Opportunities and challenges with very large arrays. *IEEE Signal Processing Magazine*, 30(1):40–60, January 2013.

[17] L. Tan and J. Jiang. *Digital Signal Processing (Second Edition).* Elsevier India, 2013.

[18] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta. Massive MIMO for next generation wireless systems. *IEEE Communications Magazine*, 52(2):186–195, February 2014.

[19] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang. What Will 5G Be? *IEEE Journal on Selected Areas in Communications*, 32(6):1065–1082, June 2014.

[20] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang. An overview of Massive MIMO: Benefits and challenges. *IEEE Journal of Selected Topics in Signal Processing*, 8(5):742–758, October 2014.

[21] C. Han, T. Harrold, S. Armour, I. Krikidis, S. Videv, P. M. Grant, H. Haas, J. S. Thompson, I. Ku, C. Wang, T. A. Le, M. R. Nakhai, J. Zhang, and L. Hanzo. Green radio: radio techniques to enable energy-efficient wireless networks. *IEEE Communications Magazine*, 49(6):46–54, June 2011.

[22] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta. Energy and spectral efficiency of very large multiuser MIMO systems. *IEEE Transactions on Communications*, 61(4):1436–1449, April 2013.

[23] M. McGuire and P. Wan. Analysis of Joint Channel Estimation and Equalization using a Kalman Filter. *2006 Canadian Conference on Electrical and Computer Engineering, Ottawa*, pages 912–915, May 2006.

[24] W. Hardjawana, R. Li, B. Vucetic, and Y. Li. A New Iterative Channel Estimation for High Mobility MIMO-OFDM Systems. *2010 IEEE 71st Vehicular Technology Conference, Taipei*, pages 1–5, May 2010.

[25] M. Rana and M. Hosain. Adaptive Channel Estimation Techniques for MIMO-OFDM Systems. *International Journal of Advanced Computer Science and Applications*, 1(6):134–138, December 2010.

[26] N. Wang, G. Gui, Z. Zhang, T. Tang, and J. Jiang. A Novel Sparse Channel Estimation Method for Multipath MIMO-OFDM Systems. *2011 IEEE Vehicular Technology Conference (VTC Fall), San Francisco, CA*, pages 1–5, September 2011.

[27] Y. Liu and S. Sezginer. Iterative compensated MMSE channel estimation in LTE systems. *2012 IEEE International Conference on Communications, Ottawa, ON*, pages 4862–4866, June 2012.

[28] L. Tong and S. Perreau. Multichannel blind identification: from subspace to maximum likelihood methods. *Proceedings of the IEEE*, 86(10):1951–1968, October 1998.

[29] H. Xie, F. Gao, and S. Jin. An overview of low-rank channel estimation for Massive MIMO systems. *IEEE Access*, 4:7313–7321, November 2016.

[30] F. Ling and J. Proakis. Nonstationary learning characteristics of least squares adaptive estimation algorithms. In *ICASSP '84. IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 9, pages 118–121, March 1984.

[31] R. A. Ziegler and J. M. Cioffi. A comparison of least squares and gradient adaptive equalization for multipath fading in wideband digital mobile radio. In *1989 IEEE Global Telecommunications Conference and Exhibition 'Communications Technology for the 1990s and Beyond'*, pages 102–106 vol.1, November 1989.

[32] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela. V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel. In *1998 URSI International Symposium on Signals, Systems, and Electronics. Conference Proceedings*, pages 295–300, October 1998.

[33] C. Li, C. Meng F. Jiang, and Z. Gong. A New Turbo Equalizer Conditioned on Estimated Channel for MIMO MMSE Receiver. *IEEE Communications Letters*, 21(4):4862–4866, April 2017.

[34] A. Mezghani, M. Rouatbi, and J. A. Nossek. An iterative receiver for quantized MIMO systems. In *2012 16th IEEE Mediterranean Electrotechnical Conference*, pages 1049–1052, March 2012.

[35] S. Lee, A. C. Singer, and N. R. Shanbhag. Linear turbo equalization analysis via BER transfer and EXIT charts. *IEEE Transactions on Signal Processing*, 53(8):2883–2897, August 2005.

[36] A. Mezghani, M. Rouatbi, and J. A. Nossek. An iterative receiver for quantized MIMO systems. In *2012 16th IEEE Mediterranean Electrotechnical Conference*, pages 1049–1052, March 2012.

[37] C. Studer and G. Durisi. Quantized Massive MU-MIMO-OFDM Uplink. *IEEE Transactions on Communications*, 64(10):2387–2399, June 2016.

[38] E. R. Berlekamp. The technology of error-correcting codes. *Proceedings of the IEEE*, 68(5):564–593, May 1980.

[39] S. Lin and D. J. Costello, Jr. *Error Control Coding (2nd Edition)*. Pearson, May 2004.

[40] J. Hokfelt, O. Edfors, and T. Maseng. A turbo code interleaver design criterion based on the performance of iterative decoding. *IEEE Communications Letters*, 5(2):52–54, February 2001.

[41] S. Brink. Convergence of iterative decoding. *Electronics Letters*, 35(10):806–808, May 1999.

[42] A. Viterbi. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Transactions on Information Theory*, 13(2):260–269, April 1967.

[43] R. Gallager. Low-density parity-check codes. *IRE Transactions on Information Theory*, 8(1):21–28, January 1962.

[44] O. Iscan, D. Lentner, and W. Xu. A comparison of channel coding schemes for 5g short message transmission. In *2016 IEEE Globecom Workshops*, pages 1–6, December 2016.

[45] IEEE. Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications; amendment 4: Enhancements for higher throughput. September 2007.

[46] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv. Optimal decoding of linear codes for minimizing symbol error rate (corresp.). *IEEE Transactions on Information Theory*, 20(2):284–287, March 1974.

[47] D. J. C. MacKay and R. M. Neal. Near shannon limit performance of low density parity check codes. *Electronics Letters*, 33(6):457–458, March 1997.

[48] C. Hsu and A. Anastasopoulos. Capacity achieving LDPC codes through puncturing. *IEEE Transactions on Information Theory*, 54(10):4698–4706, October 2008.

[49] IEEE committee. *The IEEE Standard Dictionary of Electrical and Electronics Terms (1st Edition)*. Hoboken, New Jersey, U.S.A.: Wiley-Interscience, January 1972.

[50] S. Haykin. *Digital Communication*. Wiley, 1988.

[51] M. S. Gast. *802.11 Wireless Networks: The Definitive Guide: The Definitive Guide, 2nd Edition*. O'Reilly Media, April 2005.

[52] H. Otsuka, R. Tian, and Senda K. Transmission performance of an ofdm-based higher-order modulation scheme in multipath fading channels. *Journal of Sensor and Actuator Networks*, 8:19, March 2019.

[53] S. B. Weinstein. The history of orthogonal frequency-division multiplexing [history of communications]. *IEEE Communications Magazine*, 47(11):26–35, November 2009.

[54] E. O. Brigham and R. E. Morrow. The fast Fourier transform. *IEEE Spectrum*, 4(12):63–70, December 1967.

[55] J. T. E. McDonnell and T. A. Wilkinson. Comparison of computational complexity of adaptive equalization and ofdm for indoor wireless networks. In *Proceedings of PIMRC '96 - 7th International Symposium on Personal, Indoor, and Mobile Communications*, volume 3, pages 1088–1091 vol.3, October 1996.

[56] D. Falconer, S. L. Ariyavisitakul, A. Benyamin-Seeyar, and B. Eidson. Frequency domain equalization for single-carrier broadband wireless systems. *IEEE Communications Magazine*, 40(4):58–66, April 2002.

[57] T. Jiang and Y. Wu. An Overview: Peak-to-Average Power Ratio Reduction Techniques for OFDM Signals. *IEEE Transactions on Broadcasting*, 54(2):257–268, June 2008.

[58] S. Han and J. Lee. An overview of peak-to-average power ratio reduction techniques for multicarrier transmission. *IEEE Wireless Communications*, 12(2):56–65, April 2005.

[59] H. Sari, G. Karam, and I. Jeanclaud. Frequency-domain equalization of mobile radio and terrestrial broadcast channels. In *1994 IEEE GLOBECOM. Communications: The Global Bridge*, volume 1, pages 1–5, November 1994.

[60] C. Ciochina, D. Castelain, D. Mottier, and H. Sari. Single-carrier space-frequency block coding: Performance evaluation. In *2007 IEEE 66th Vehicular Technology Conference*, pages 715–719, September 2007.

[61] G. Berardinelli, L. A. M. Ruiz de Temino, S. Frattasi, M. I. Rahman, and P. Mogensen. OFDMA vs. SC-FDMA: performance comparison in local area IMT-A scenarios. *IEEE Wireless Communications*, 15(5):64–72, October 2008.

[62] M. S. Obaidat, A. Anpalagan, and I. Woungang. *Handbook of Green Information and Communication Systems*. Academic Press, Inc., USA, 1st edition, 2012.

[63] E. Björnson, E. G. Larsson, and T. L. Marzetta. Massive MIMO: ten myths and one critical question. *IEEE Communications Magazine*, 54(2):114–123, February 2016.

[64] D. Tse and P. Viswanath. *Fundamentals of Wireless Communication*. Cambridge University Press New York, NY, USA, 2005.

[65] N. C. Beaulieu. Introduction to "certain topics in telegraph transmission theory". *Proceedings of the IEEE*, 90(2):276–279, February 2002.

[66] P. L. Dragotti and M. Gastpar. *Distributed Source Coding 1st Edition*. Academic Press, January 2009.

[67] P. Duhamel and M. Kieffer. *Joint Source-Channel Decoding: A Cross-Layer Perspective with Applications in Video Broadcasting (1st)*. Academic Press, Inc. Orlando, FL, USA, 2010.

[68] P. Handel. Properties of the ieee-std-1057 four-parameter sine wave fit algorithm. *IEEE Transactions on Instrumentation and Measurement*, 49(6):1189–1193, December 2000.

[69] J. Ma and L. Ping. Data-aided channel estimation in large antenna systems. *IEEE Transactions on Signal Processing*, 62(12):3111–3124, June 2014.

[70] O. Elijah, C. Y. Leow, T. A. Rahman, S. Nunoo, and S. Z. Iliya. A comprehensive survey of pilot contamination in Massive MIMO—5G system. *IEEE Communications Surveys Tutorials*, 18(2):905–923, November 2016.

[71] J. Zhao, S. Ni, Y. Gong, and Q. Zhang. Pilot contamination reduction in TDD-based Massive MIMO systems. *IET Communications*, 13(10):1425–1432, June 2019.

[72] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath. Pilot contamination and precoding in multi-cell TDD systems. *IEEE Transactions on Wireless Communications*, 10(8):2640–2651, August 2011.

[73] G. Fodor, P. Di Marco, and M. Telek. Performance analysis of block and comb type channel estimation for Massive MIMO systems. In *1st International Conference on 5G for Ubiquitous Connectivity*, pages 62–69, November 2014.

[74] H. O. Kunz. On the Equivalence Between One-Dimensional Discrete Walsh-Hadamard and Multidimensional Discrete Fourier Transforms. *IEEE Transactions on Computers*, C-28(3):267–268, March 1979.

[75] J. R. Barry, D. G. Messerschmitt, and E. A. Lee. *Digital Communication: Third Edition*. Kluwer Academic Publishers Norwell, MA, USA, 2003.

[76] S. M. Kay. *Fundamentals of Statistical Signal Processing - Estimation Theory (Vol. 1)*. Prentice Hall PTR, January 2010.

[77] H. V. Henderson and S. R. Searle. On deriving the inverse of a sum of matrices. *SIAM Review*, 23(1):53–60, 1981.

[78] R. W. Lucky. The adaptive equalizer. *IEEE Signal Processing Magazine*, 23(3):104–107, May 2006.

[79] B.P. Lathi and Z. Ding. *Modern Digital and Analog Communication Systems (Fourth Edition)*. Oxford University Press, Inc., 2009.

[80] S. Chen, G. Dai, and T. Yen. Zero-Forcing equalization for OFDM systems over doubly-selective fading channels using frequency domain redundancy. *IEEE Transactions on Consumer Electronics*, 50(4):1004–1008, November 2004.

[81] M. Burger, B. Kaltenbacher, and A. Neubauer. *Iterative Solution Methods*. New York, NY, USA: Springer, 2011.

[82] J. Wang, M. Li, Y. Zhang, and Q. Zhou. Effect of channel estimation error on the mutual information of mimo fading channels. In *2008 4th International Conference on Wireless Communications, Networking and Mobile Computing*, pages 1–4, October 2008.

[83] M. A. Woodbury. Inverting modified matrices, memorandum rept. 42. *Statistical Research Group, Princeton University, Princeton, NJ*, 4, 1950.

[84] G. H. Golub and C. F. Van Loan. *Matrix Computations 3rd ed.* Baltimore, MD: Johns Hopkins, 1996.

[85] A. Neubauer, J. Freudenberger, K. Volker, J Freudenberger, and V. Kuhn. *Coding Theory : Algorithms, Architectures and Applications*. New York: John Wiley & Sons, Incorporated, 2007.

[86] J. R. Barry and E. A. Lee and D. G. Messerschmitt. *Digital Communication*. Springer, third edition, 2004.

[87] J. Ren, S. Yoon, and Y. Byun. A new iterative receiver scheme for multiple-input multiple-output (MIMO) systems. In *2011 First ACIS/JNU International Conference on Computers, Networks, Systems and Industrial Engineering*, pages 137–140, May 2011.

[88] C. Douillard, M. Jezequel, C. Berrou, A. Picart, and P. Didier. Iterative correction of intersymbol interference: turbo-equalization. *European Transactions on Telecommunications, Wiley*, 6(5):507–512, September 1995.

[89] J. Hagenauer. The EXIT chart - Introduction to extrinsic information transfer in iterative processing. In *2004 European Signal Processing Conference*, pages 1541–1548, September 2004.

[90] S. Sen, N. Santhapuri, R. R. Choudhury, and S. Nelakuditi. Successive interference cancellation: A back-of-the-envelope perspective. In *Proceedings of the 9th ACM SIGCOMM Workshop on Hot Topics in Networks*, Hotnets-IX, pages 17:1–17:6, New York, NY, USA, 2010. ACM.

[91] K. Zhai, Z. Ma, and X. Lei. Accurate performance analysis of coded large-scale multiuser MIMO systems with mmse receivers. *Sensors (Basel)*, 19(13):2884, July 2019.

[92] V. Crâşmariu, M. Arvinte, A. Enescu, and S. Ciochină. Performance analysis of the singular value decomposition with block-diagonalization precoding in multi-user massive mimo systems. In *2016 IEEE International Symposium on Electronics and Telecommunications (ISETC)*, pages 71–74, October 2016.

[93] F. R. Kschischang, B. J. Frey, and H. Loeliger. Factor graphs and the sum-product algorithm. *IEEE Transactions on Information Theory*, 47(2):498–519, February 2001.

[94] M. S. Pinsker. *Information and Information Stability of Random Variables and Processes*. Holden-Day, San Francisco, USA, english translation by a. feinstein in 1964 edition, 1960.

[95] M. P. Wand. Data-based choice of histogram bin width. *The American Statistician*, 51(1):59–64, 1997.

[96] D. W. Scott. On optimal and data-based histograms. *Biometrika*, 66:605–610, December 1979.

[97] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, July 1948.

[98] N. Chiurtu, B. Rimoldi, and E. Telatar. On the capacity of multi-antenna gaussian channels. In *Proceedings. 2001 IEEE International Symposium on Information Theory (IEEE Cat. No.01CH37252)*, page 53, June 2001.