DESIGN AND ANALYSIS OF A ROBUST CODEC FOR WIRELESS CHANNEL USING MACHINE LEARNING

A PROJECT REPORT

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE AWARD OF THE DEGREE

OF

MASTER OF TECHNOLOGY

IN

[SIGNAL PROCESSING AND DIGITAL DESIGN]

Submitted by:

MOGAHED ADAM MOHAMED SHARAFUDEN

Roll No (2K20/SPD/16)

Under the supervision of:

ASSISTANT PROF MR. VINAY KUMAR

(ECE)

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bwana Road, Delhi – 110042 (India)

May, 2022

ELECTRONICS AND COMMUNICATION DEPARTMENT

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

DECLARATION

I **MOGAHED ADAM MOHAMED SHARAFUDEN**, Roll No. 2K20/SPD/16 of M. Tech (Signal Processing and Digital Design, hereby declare that the Major Project II titled "DESIGN AND ANALYSIS OF A ROBUST CODEC FOR WIRELESS CHANNEL USING MACHINE LEARNING" which is submitted by me to the Department of Electronics and Communication Engineering, Delhi Technological University, Delhi is original and not copied from any source without proper citation. This work has not previously formed the basis for the award of any Degree, Diploma Associateship, Fellowship, or other similar title or recognition. **The work has been published** in indexed journal with the following details:

- TITLE OF THE PAPER: 1. "**Performance Analysis of Adaptive Modulation and Codec Selection Based Long-Short Term Memory Neural Network (LSTM-NN)**"
- AUTHOR NAMES: Mogahed. A. M. Sharafuden, Vinay Kumar, Mohammed E. Y. Abdalla.
- NAME OF JOURNAL: IOSR Journal of Electronics and Communication Engineering (IOSR-JECE).
- DATE OF PAPER SUBMISSION: 04-04-2020
- DATE OF PAPER ACCEPTANCE: 20-04-2020

Place: Delhi

Date: 30/05/2022

MOGAHED A. M. SHARAFUDEN

ROLL NO: 2K20/SPD/16

ECE DEPARTMENT

ELECTRONICS AND COMMUNICATION DEPARTMENT

DELHI TECHNOLGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

CERTIFICATE

This is to certify that the Major Project II titled "DESIGN AND ANALYSIS OF A ROBUST CODEC FOR WIRELESS CHANNEL USING MACHINE LEARNING" is a bonafide record of work done by MOGAHED A. M. SHARAFUDEN, Roll No. 2K20/SPD/16 at Delhi Technological University, New Delhi for the Major ProjectII. this project was carried out under my supervision and has not been submitted anywhere else, either in part or full, for the award of any other degree or diploma to the best of my knowledge and belief.

Place: Delhi

Date: 30/05/2022

Vinay Kumar 31/05/22

Mr. Vinay Kumar

SUPERVISOR

Assistant Professor

ECE Department

Delhi Technology University

Bawana Road, Delhi – 110042, Delhi, India

ABSTRACT

This work presents the design and analysis of adaptive modulation and coding scheme intended for a codec used in short-range wireless communication system; the study aims to investigate and analyse the existing channel prediction parameters for adapting the variation of wireless communication channels. Three modulation and coding schemes are used to develop the codec namely 8QAM, 16QAM with a 1/3 BCH encoder and 4QAM without error correction. MIMO antennas are introduced in the design to increase the spectral efficiency of the codec within the fading channel. To make channel prediction possible a linear discrete channel model is presented based on ARMA processes. The presented scheme is operated at a BER threshold of 10^{-3}, with a channel correlation of 0.96 and a noise variance of 3 dB.

The adaption parameter and the channel model provide the basis of the machine learning model. A conventional neural network (NN) and Long-short term memory (LSTM) predictors are used as prediction models for the adaptation with root means square error (RMSE) of 5.6121 and 3.9157, respectively. a perfect prediction model based on the BER, and adaption parameter is also presented with Zero RMSE as a standard predictor.

The result found that the adaptive modulation and codec parameters chosen under the LSTM-NN have a data rate of 15 Mbits/s and 10 Mbit/s for conventional NN predictor. The codec achieved the design requirements, and it can serve users operating within 2.4GHz. However, the design is not a sustainable one as it will soon be outdated due to increasing data rate demands. The data rate of the codec can be improved by using a more robust encoding such as turbo and LDPC codes and transmitting using more antennas to transmit a variety of different data types such as random packets and images.

ACKNOWLEDGEMENT

It gives me immense pleasure to turn my sincere thanks and appreciation to my respected guide **Mr. Vinay Kumar,** Assistant Professor (ECE, Department, Delhi Technological University), for his valuable guidance, help, encouragement, and valuable suggestions for this work.

I sincerely also, thank all faculties and the head of the Department of Electronics and Communication Engineering, Delhi Technological University, for providing me with a solid knowledge of signal processing and digital design and communication. In addition, my gratitude to friends and others who helped me directly or indirectly to complete this project.

Above all, I thank my **family in Sudan** for their unconditional love and motivation to come to study in India. Moreover, extending my thanks to the Faculty of the Engineering University of Al Fashir, Sudan, for nominating me for postgraduate studies.

MOGAHED A. M. SHARAFUDEN

Roll No: 2K20/SPD/16

M. Tech (Signal Processing & Digital Design)

Department of Electronics and Communication Engineering

Delhi Technological University

CONTENTS

CHAPTER 2 (TECHNICAL BACKGROUND)

CHAPTER 3 (DESIGN PROCEDURES AND CONSIDERATIONS)

CHAPTER 4 (Simulation Results and Analysis)

CHAPTER 5 (Conclusion & Future Recommendations)

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

3GPP	3rd Generation Partnership Project
AI	Artificial Intelligence
AMC	Adaptive Modulation and Coding
ARMA	Autoregressive moving average
ASK	Amplitude Shift Keying
AWGN	Additive White Gaussian Noise
BCH	Bose Chaudhuri Hocquenghem codes
BER	Bite Error Rate
BPSK	Binary-Shift Keying
CSI	Channel State Information
DL	Deep Learning
FEC	Forward Error Correction
FIR	Finite Impulse Response
FNN	Feed-Forward Neural Network
FSK	Frequency Shift Keying
ICT	Information Communication Technology
IEEE	Institute of Electrical and Electronics Engineers
IIR	Infinite Impulse Response
IoT	Internet of Things
ISI	Intersymbol Interference
ISO	International Organization for Standardization
LAN	Local Area Network
LDPC	Low Density Parity Check Code
LMS	Least Mean Square
LOS	Line of Sight
LSTM	Long Short-Term Memory
MCS	Modulation and Coding Schemes

ML	Machine Learning
MIMO	Multiple Input and Multiple Output
MMSE	Minimum Mean Square Error
MPSK	M-array Phase-shift keying
NLOS	Non-Line of Sight
NN	Neural Network
OFDM	Orthogonal Frequency Division Multiplexing
PDF	Probability Density Function
PSK	Phase-Shift Keying
QAM	Quadrature Amplitude Modulation
QoS	Quality of Services
QPSK	Quadrature Phase-Shift Keying
RMSE	Root Mean Square Error
RNN	Recurrent Neural Networks
RS	Reed Solomon codes
Rx	Receiver antenna
SNR	Signal to Noise Ration
SVM	Support Vector Machine
TCM	Turbo Codes Modulation
Tx	Transmitter antenna
WHO	World Health Organisation
Wi-Fi	Wireless Fidelity
Wi-Max	Worldwide Interoperability for Microwave Access
ZF	Zero Forcing Equaliser

CHAPTER 1
INTRODUCTION

Nowadays, the demand for wireless communication is rapidly increasing due to the vast intensification of digital devices that are connected wirelessly. Increasing digital devices became a starting point to determine and plan for coming wireless communication networks within the next ten years. The Internet of things (IoT), smart cities, smart transports, and beyond needed a cognitive wireless structure to handle the complexity of systems and increase the qualities of services (QoS). This increasing demand led to the complex wireless communication networks, which introduce many challenges and opportunities that need high computational and intelligent methods to solve such problems. Therefore, Machine Learning will be the best computational method for such tasks to build such a network. This thesis aims to provide a machine learning model for short range wireless communication systems such as wireless fidelity network (Wi-Fi) based on a codec working on a Raleigh fading channel. The codec model is accomplished by using an adaptive modulation and coding scheme (AMC) to accommodate varying signal strengths over the feeded channel. AMC enables the wireless network to dynamically select modulation and coding schemes based on the wireless channel environments to achieve specific performance requirements. Nevertheless, it might not accomplish all adaption capabilities because of differences in the channel conditions. MATLAB is used for design and simulation because it provides most implementations for the elements required in this project.

1.1. Overview

The wireless network became popular because its expendable, adaptable, and accessible. Today, most internet users access a wireless network such as a Wireless Fidelity network (Wi-Fi). The number of users is rising due to the increase of applications designed to access wirelessly, such as smart homes, robotics, and healthcare wearables devices. This means increasing data traffics and downtime [1] [2]. There are no license fees required to run Wi-Fi devices, yet there are protocols that prescript how such equipment should operate. IEEE establishes these rules under the 802.11n family. These protocols describe the frequency band and working data rates.

1.1.1. Challenges and Solutions

As more people share the wireless networks, the more challenging it becomes for communication systems to cope with the increasing demands. Network congestion and drop happen more frequently to the systems. This problem usually leaves many users frustrated, and businesses may lose customers because their services are unreliable due to not being able to connect in time with customers. It is essential to address these problems to avoid buffering videos, time-out file downloads, and freezing video calls. Network resources are limited by hardware and software advancements. Therefore, essential to try to free up some of these network resources by providing more options to users. The proposed engineering solution is developed to increase the end user's rate of receiving information. For example, in the Wi-Fi network, the present solution is Wi-Fi router, a device owned by many users. The solution is designed with sustainability and impact in mind. The solution present is able to meet current traffic demands and has comparable specifications to some commercial products. However, it falls short of being a viable solution for the future.

Generally, wireless communication provides an easy and cheap way of accessing networks. Powering some of the latest technologies such as robotics, IoT, and self-driving cars, business and individuals benefit positively from good internet coverage. Wireless infrastructure is under stress due to the significant number of users. Therefore, improving the underlying infrastructure is important to improve the user experience.

1.1.2. Purpose and Objectives

This work aims to design and analyze the AMC scheme selection used for the Wi-Fi router network as a short-range wireless communication system. The modulation scheme is intended to select codec techniques called quadratic amplitude modulation (QAM); these techniques are 4QAM, 8QAM and 16QAM for the wireless network and make the channel prediction possible assuming linearity of the channel model. Machine learning is required for adaptation and channel model parameters for this task. For adaptation, a neural network and long short-term memory (LSTM) will be the basis of the machine learning prediction model for this work. The objective of this work is to study previous work in this field of study and discuss the major contribution of this work and the previous one.

Furthermore, try to answer the questions how to predict channel model and index for future transmission in wireless channels. How accurate is this prediction, highlighting the future work that could be done using machine learning models for wireless communication to achieve user's demand, representing the high date rate of internet services.

1.2. Motivation for this work

AMC have been broadly applied in both wired and wireless systems to adjust the variations of the communication channel environments, particularly in wireless communication. The channel is continuously affected by shadowing, pass losses, and interference. Generally, in communication systems, wireless channels are intensely time-variant than wired communication channels. Therefore, presently AMC has been combined with orthogonal frequency division multiplexing and multiple input multiple output systems (MIMO), improving the consistency of systems such as Wi-Fi standards IEEE 802.11n and IEEE 802.11ac.

However, wireless systems protocols such as the IEEE 802.11n founded wireless network are time-varying. The complicated operation of wireless channels made the physical layer adaptation problematic to implement in practice. They are causing the AMC schemes not to be well adapted and yielding performance degradation. In addition, once AMC structures are combined with OFDM and MIMO, modulation and the convolutional coding process are highly challenging in practical wireless systems. The design of communication systems divided signal processing into many independent blocks.

Consequently, optimizing these combinational systems leads to computationally complex systems [1]. In contrast, machine learning (ML) has the potential to enhance those independent blokes and make the computation for such complex systems possible. Hence, ML assists as a batter contender for optimizing AMC helped wireless communication systems with explicit hardware configurations when communicating across wireless channels.

All these challenges motivate researchers in wireless communication, opting to maximize their efforts by proposing solutions to enhance spectrum efficiency. Consequently, the fundamental motivation of this work is to present the design and analysis of AMC schemes for the codec used in short-range wireless communications.

1.3. Approach and Methodology

In this work, AMC is the primary technique used to adapt varying signal strengths. For the system to meet requirements and achieve a high data rate, the assumptions made are the system channel's linearity and fading coefficients constant for the entire codeword. The method flows the general adaptive modulation and coding scheme depicted in Fig. 1.1. below. In section 2.2. an explanation of the figure has been given.

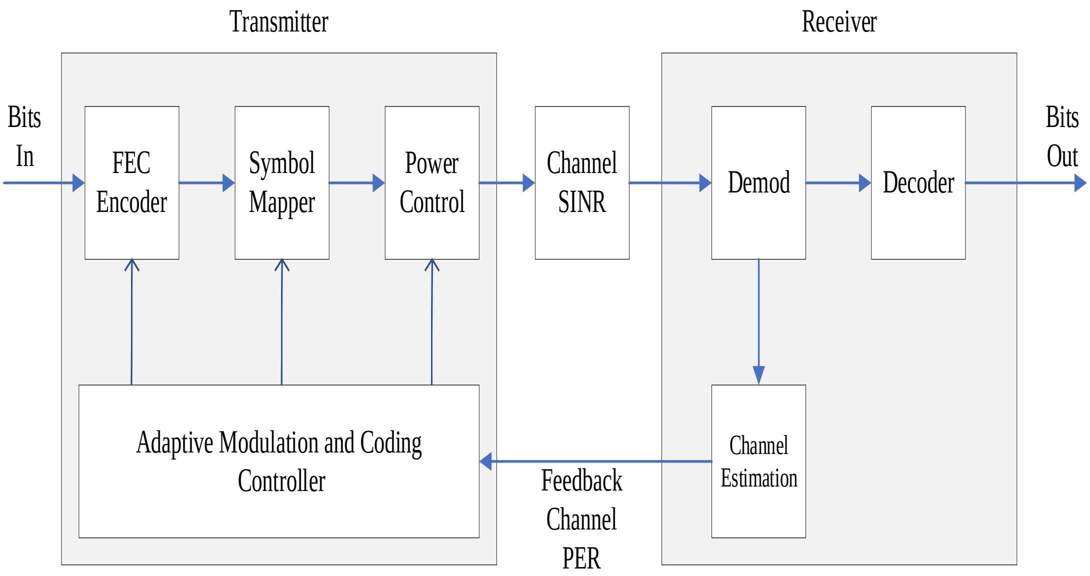

Fig.1.1. Adaptive modulation and coding block diagram [3].

Two models are proposed for designing the machine learning prediction model; one is based on conventional neural networks (NN), and the second proposed predictor uses a Long-Short Term Memory (LSTM). The input of the binary dada sequences is autogenerated and combined with the machine learning predictors to produce the Modulation Coding Scheme (MCS) selection decision; those algorithms are based on NN and LSTM. The processes used in LSTM are similar to those in a traditional NN, with slight variations to accommodate other computational units. This NN is used for regression in this design; a similar approach is used for the LSTM, elaborated in section 3.4. MATLAB is the main tool for designing and simulating the proposed system and prediction model design.

1.4. Limitations

This work is a robust codec for MCS selection for the short-range wireless communication system-based machine learning algorithmic method. Likewise, there are various drawbacks to the overall system that are considered before the ultimate decision is made. The current AMC is only simulated with fewer parameter variations due to the time limitation; as a result, we used a single type of data source, fixed MIMO structure, and the same FEC encoder. However, further experimentation needs to optimize the machine learning parameters and explore several neural network architectures highlighted in the future recommendation section of this thesis.

1.5. Impact of Wireless Networks

It is uncommon these days to meet someone without a mobile device. Most people do their jobs and activities easily with the power that the mobile device puts in our hands, making communication more powerful. Wireless communication networks solve all these problems. It gives the ability to text, and video calls your people across the world without being at a fixed location to do so.

Wireless communications mean your physical location no longer limits you; it also means that you are able to get news faster and make the decision much faster. The internet means that there are multiple sources of information, such as social media and blogs; it is hard to mislead people over the internet because it has become easy to check facts. Wireless communication provides straightforward access to the internet. The internet is a database of human experiences where everyone can append their own. Such a database means one does not have to experience situations first-hand to appreciate them. Each individual has no limit to what they can or cannot access in this database. Internet access increases people's productivity and enables them to access educational and develop resources at a low cost and some for free.

Wireless communication makes technologies such as wearable devices, IoT devices, and self-driving cars possible, and robots help people perform physical tasks remotely. Technological advances in health care are looking into removing surgical operations to increase health care reach. Wearables devices such as smart watches are also a technology that can track your health more accurately and can be used to give more detailed medical records. Moreover, responsive equipment can operate remotely through wireless channels and protocols to perform a particular function depending on the client's state. In businesses also, wireless communication is a source of increasing incomes. For instance, Wi-Fi availability is usually at the top of people's list when visiting a place such as malls or restaurants. Because Wi-Fi is becoming a source of enjoyment that make customers spend more time in restaurants and malls, which means more income. It can be a source of entertainment, and you can stream, and download content from extra via a wireless network. The biggest benefactors of wireless networks are rural communities and schools; allowing people in such areas to access information would be challenging otherwise because of the considerable distance one would have to travel. It can be used as an alternative source of complementary information.

It is necessary to identify and evaluate the influence of an engineering solution to ensure that it is a holistic design and that unintended consequences of the invention are recognized and addressed. The design is evaluated across the fields discussed below.

1.5.1. Environmental

For instance, Wi-Fi routers are mainly constructed from plastic and other recyclable components; therefore, their e-waste contribution is reduced. However, the manufacturing process used to produce plastic usually has negative environmental impacts. The process of producing plastic is energy demanding, and most manufacturing plants burn coal as an energy source. Furthermore, not all kinds of plastic are recyclable, and oil spills are a huge environmental hazard. Communication systems also reduce the carbon footprint since they provide remote communication.

1.5.2. Social

Wireless communication made it easier and cheaper to access a global audience. People are now able to keep in contact even when far apart. Many developing countries access the internet through the wireless network. This is because Wi-Fi routers are potable and cheaper than fiber optic cables. Communication is one of the defining constructs of human societies. Therefore, Wi-Fi as a communication method can alter existing social structures. The social impact of the communication system can be seen during the COVD19 pandemic, where most social interaction occurs remotely due to social distancing.

WHO indicates that there have not been any medical cases due to exposure to Wi-Fi signals as of date. A study by [30] supports this claim by examining previous Wi-Fi studies, and they conclude that most of them were poorly conducted and thus presented flawed results. Wireless communication makes it easy to consume multimedia as a source of entertainment. The side effect of this is that it may result in health conditions associated with a lack of exercise or movement and eyesight problems due to prolonged screen time.

1.5.3. Economic

Remote communication means businesses can coordinate across different time zones. This means businesses can conduct more business, and individuals can access more opportunities by working remotely. Wireless communication systems remove the cost barrier to acquiring new information; many institutions and sites provide open access to educational resources. Individuals and small companies can now compete with large corporations because they can conduct business without having physical offices. Wireless communication can cut business operation costs by reducing the budget spent on travel costs. This kind of communication encourages a competitive market since clients have access to many options. A study by [4] shows that a one percent increase in information communication technology (ICT) infrastructure for broadband subscribers can increase between 0.0767% to 0.396%.

1.6. Sustainability in Engineering Solutions

A sustainable engineering solution meets current requirements while accounting for future demands. The design should not use limited resources, be operational for a reasonable duration, and have a positive impact. Wi-Fi is expected to expand to the 6GHz spectrum. This is a move to prove faster, reliable, and broader network coverage. This is in response to the growing number of Wi-Fi devices. Typical Wi-Fi networks use the 2.4GHz and some at the 5GHz band. The router limits the frequency at which devices network operates [4]. The 2GHz band tends to have more comprehensive coverage and is better at passing through obstructions than 5GHz. However, 5G is faster than the 2.4GHz band. There are more interferences in the 2.4GHz band due to many devices utilizing it, making it more unreliable and congested.

The proposed solution is a software-based codec with only the antennas as physical components. This means that the system can easily be updated without additional costs to the user. The limiting factor is the hardware, which goes obsolete when a new ad improved one is developed. The energy consumption of the proposed solution is minimized by using a 2x2 antenna configuration, which has the least number of antennas needed for MIMO. It usually takes a couple of years for device manufacturers to design devices that meet the latest standards; this means that the designed codec could be deployed for years before being absolute. Essentially the presented solution will not be able to compete with routers operating in higher bands when devices migrate to newer standards. However, the given solution will be able to serve most users in the meanwhile operation at the 2.4GHz band.

1.7. Thesis Parts

The rest of this thesis comes in as flow; chapter two presents a technical background of the wireless communication systems, followed by chapter three, which introduces the design procedure of this work. And the design analysis comes in chapter four. Conclusion and future recommendations were stated in chapter five of this work.

CHAPTER 2

TECHNICAL BACKGROUND

In this chapter, general concepts and techniques of wireless communication systems have been explained. These concepts are also applied in short-range wireless communication such as Wi-Fi and Wi-Max, and long-range wireless communication systems as cellular network systems to transmit data between the transmitter and receiver devices.

2.1. Digital communication Model

Wireless networks take advantage of a digital communication model. Digital communication systems generally outperform their analog counterparts. Digital systems are much cheaper to implement and well suited for long-distance transmission.

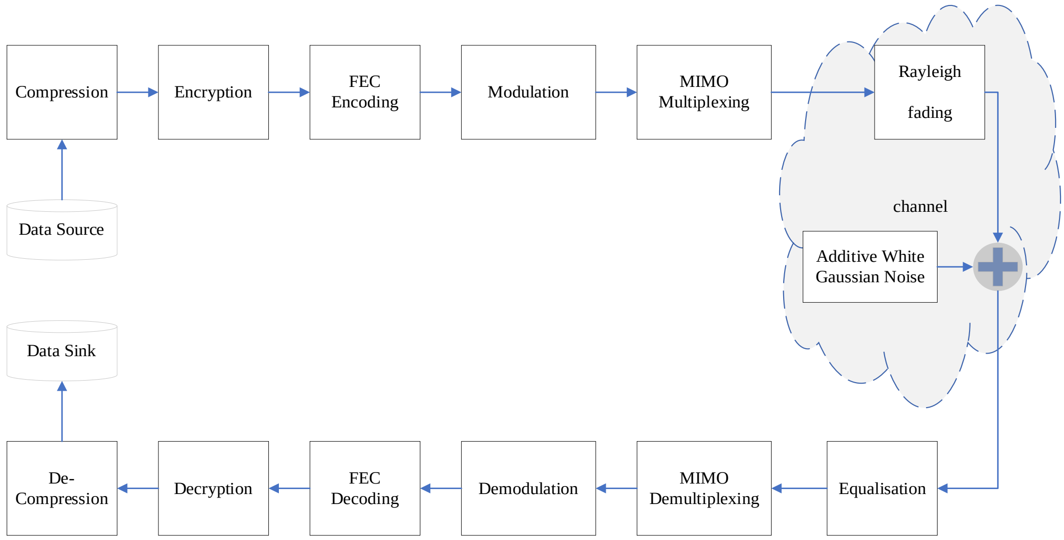

Fig.2.1. Digital communication model.

This project adapts the communication model depicted in figure 2.1. compression or source coding is the mapping from one sequence to a shorter sequence while minimizing information entropy. Compression is needed to reduce file size without significant loss of information. Encryption is introduced in communication systems to make it hard for eavesdroppers to decipher transmitted information. Most digital communication systems use advanced encryption standards to introduce confusion and diffusion to a message.

2.1.1. Forward error correction (FEC)

FEC is a means of implementing error control capability in communication systems. The transmitter introduces redundant information to the information to be transmitted. The receiver then automatically performs error correction of the received data. Given a message containing k bits, the encoder appends r parity bits to add redundancy. The encoder then maps the message into a unique n bitstream, such as FEC codes said to have a code rate of k/n. Commonly used FEC codes are classified in Fig.2.2. below and this work is used a BCH codes encoder.

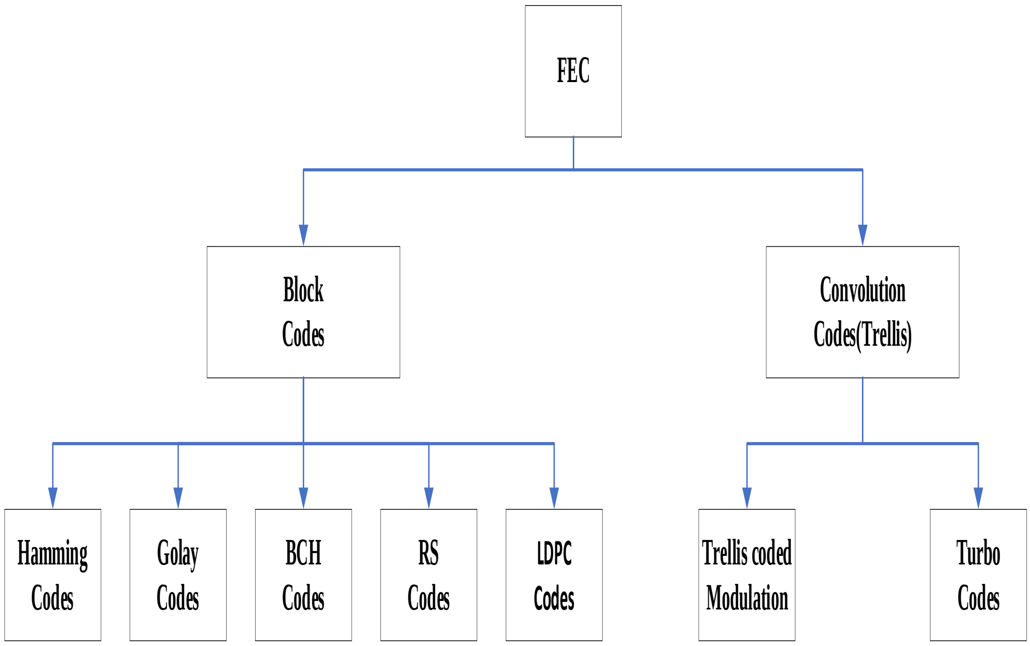

Fig.2.2. Classification of FEC codes [5].

Basically, FEC codes are divided into two main code groups: block codes and convolution codes (also known as Trellis codes).

a. Block Codes

The codes under this category are include, Hamming codes, RS codes, LDPC codes, and Golay code in addition to BCH codes.

Hamming codes is a procedure established by R. W. Hamming. The purpose of this codes is to identify the bits error among the information bits stream; this code has the ability to determine the exact location of the bits error for information bits of every length by using the correlation among the information bits and noise bits. The first error-correcting codes developed by hamming is in 1950, known as (7, 4) hamming codes.

Golay codes is a linear binary block code. It's a great N-error rectifying code capable of correcting every bit within the N code word. And the codes has d_{min} = 2N+1 distance to exactly one codeword. In this code, a maximum of t-errors can be corrected by the t-error correcting code, and every word is a perfect t-error correcting code is within t of exactly one codeword. Compared to hamming codes, the time complexity is 0(n2) for the Golay code and 0(n) for the hamming that calculates the error.

BCH code is known as cyclic codes because its continuously correcting error passed on polynomial function generating codes. in this work, we used BCH binary code, which has a particular ability to control several errors caused by the encoder. The codeword in this code generates by polynomial generator g(s), which can correct multiple random errors. Error correction in BCH has n = 2m-1 block length and numbers of n-k parity check bits and distance of d_{min} = 2N+1 as the smallest amount of length [5]. g(s) generates a simple binary BCH code which is a continuous polynomial with i^{th} degree over $GF(2^m)$ has a to a^{2t} as roots, a is the first component in $GF(2^m)$. based on this roots g(s) have $(s+a)(s+a^2).....(s+a^{2t})$ factor. This factor leads to the general polynomial generator equation:

$$g(x) = LCM[\sigma_1(s)\ \sigma_2(s)\ldots\ \sigma_i(s)] \tag{2.1}$$

Where σ from $\sigma_1(s)$ to $\sigma_i(s)$ representing the marginal polynomials has a $(s+a)(s+a^2).....(s+a^{2t})$ factor.

Reed Solomon codes (RS), this codec technique has an extensive application in digital communication systems particularly in telecommunications cellular systems, satellite communication, and high-speed modems. RS encoder adds extra redundant bits to the digital block data. The RS decoder handles every data block in the decoder side and tries to fix errors and extract the original information. RS codes is denoted as RS*(n,k)*, which means that the encoder brings k symbols including parity symbols to make n codeword. RS detector has the ability to correct N-symbols that containing errors into the codeword.

Low-Density Parity-Check code (LDPC) is considered as the best error correction code applied in recent technology, such as 5G communication systems. R. Gallager first introduced it in his Ph.D. dissertation in 1960. Nevertheless, it was not implemented due to the limited technological advancement until revived in the mid-90s by D. Mackay and R. Neal at the University of Cambridge. This code is defined for N bit long codes in positions of M-number of parity check equivalences, and these equivalences can be defined by an MxN parity check matrix H.

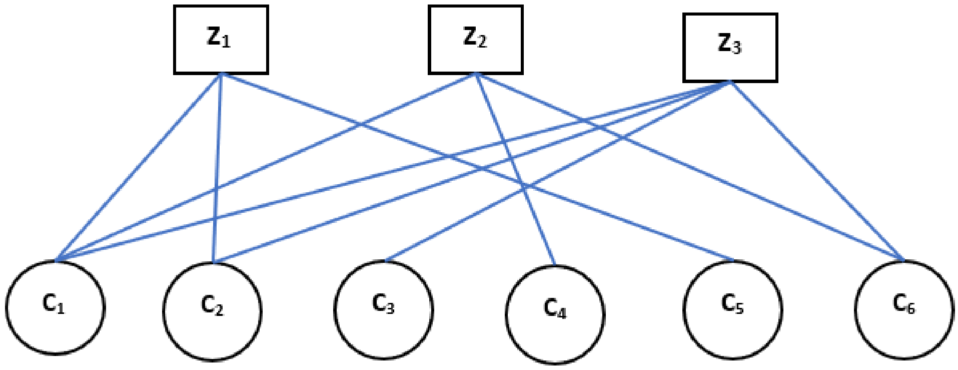

Fig.2.3. LDPC codes representation [5].

The parity check matrix H characterized graphically by the Tanner graph. The Tanner graph connects the M-number of check nodes to N-number of variable nodes making the parity check matrix H of m to n diminution element and rows respectively. for example, a 6-bits long codeword in the form $c = [c_1, \; c_2, c_3 \,, c_4, c_5, c_6]$ Which satisfies three parity check equations can be written as:

$$c_1 \oplus c_2 \oplus c_5 = 0 \tag{2.2}$$

$$c_1 \oplus c_4 \oplus c_6 = 0 \tag{2.3}$$

$$c_1 \oplus c_2 \oplus c_5 \oplus c_6 = 0 \tag{2.4}$$

We can now define the 3x6 parity check matrix as,

$$c_1 \oplus c_2 \oplus c_5 \tag{2.5}$$

$$H = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \end{bmatrix} \tag{2.6}$$

From the tanner graph representation in Fig.2.3. can draw the marked path, for example $z_2 \rightarrow \; C_1 \rightarrow z_3 \rightarrow C_6 \rightarrow z_2$ is the cycle of 4 which means the time required to return to the initial location. This process is known by girth of code [5].

b. Convolution codes (CC)

CC is the second type of FEC codes, with differs from the block codes. The difference is that the convolutional code encoder has memory. Its n output at any time is not just dependent upon k numbers of input; instead, it's also subjected to m previous input blocks. CC codes represented as an (n, k, m) which means that the encoder brings k symbols including parity symbols to make n codeword, and m representing a memory circuit as a linear sequential. Normally, n and k are integers number [5]. The decoding under CC codes is proposed by Wozencraft as an efficient decoding scheme for the convolution codes and performed many investigational on the same. Convolution code has simple methods for encoding and decoding, and kind of a simple generalization of linear codes and has encoding as a cyclic code. An (n, k) convolution codes is specified by a k X n generator matrix of over F_2 polynomials. For example, the flowing two equations are generator matrix under k and n.

$$G_1 = [x^2 + 1, x^2 + x + 1] \tag{2.7}$$

Equation (2.7) is representation of a (2,1) convolution codes matrix.

$$G_2 = \begin{bmatrix} 1 + x & 0 & x + 1 \\ 0 & 1 & x \end{bmatrix} \tag{2.8}$$

And equation (2.3) is representation of a (3,2) convolution codes matrix.

Generally, convolutional codes are divided into two code groups: Turbo codes and trellis-coded modulation (TCM). And the flowing paragraphs explain the general idea for those categories.

Turbo codes is suggested by Berrou and Glavieux since 1993. These codes have such an impressive power efficiency in Additive white Gaussian Noise and flat fading channels for the low bit error rate and are commonly applied in the handover occurs in cellular systems. Nevertheless, turbo codes have lengthy invisibility and less functioning at lower BER, making the channel evaluation and tracing a crucial issue. Turbo codes is suitable for applications such as deep space and satellite communication. The channel capacity under turbo codes with bandwidth B Hz can be calculated using equation (2.9) [5].

$$C = B\, log_2 \left[1 + \frac{P}{N_0 B}\right] \quad bits/second \tag{2.9}$$

Where, average power at transmitter side with R_b rate and E_b energy per bit is $P = E_b.R_b.$

Trellis coded modulation (TCM): the idea of TCM is combining both encoding and modulation to decrease the error probability by adding more code bits using Euclidean distance only. This type of codes has a large coding increase without compromising bandwidth and enables for simultaneous switching across nodes. TCM codes have an excellent ability to feed channels and are broadly used in modems. Figure 2. Shows four state Trellis coded modulation.

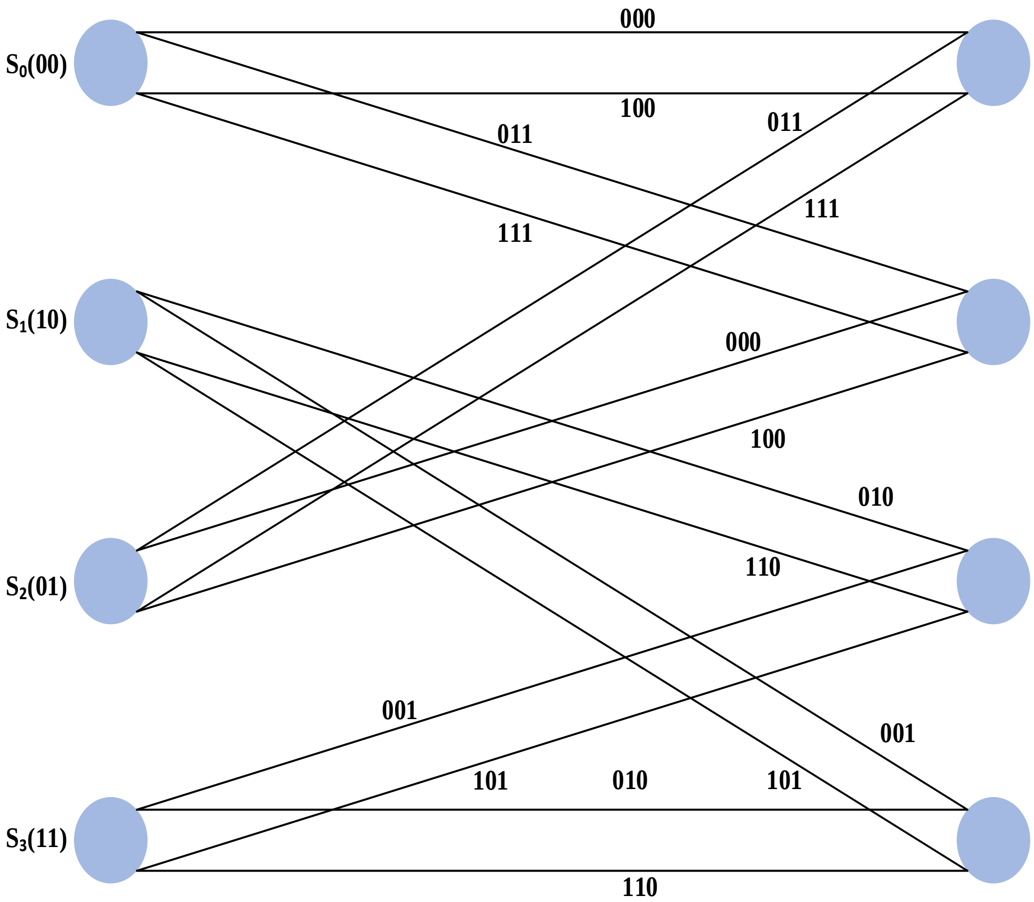

Fig.2.4. Encoder for four state Trellis TCM [5].

Usually, the throughput (b/s/Hz) in an uncoded communication system increases [6], and the performance of BER declines due to the decrease of the minimum distance d_{min}. However, when TCM is used, increasing the constellation range. TCM accomplishes better gain by using both distances of code properties and constellation.

2.1.2. Digital Modulation

Modulation performs bitstream maps to signal waveform; this process is completed by using a single or combination of parameters of the carrier signal. The wildly used modulation techniques in digital communications are Frequency-shift keying (FSK), Phase-Shift Keying (PSK), Amplitude-Shift Keying (ASK), and Quadrature Amplitude Modulation (QAM), which is a combination of PSK and ASK. An M-array modulation system with b represents code bits capable of transmitting (2^{b}) unique waveforms at a given sample. The coming is a brief discussion of the commonly used modulation techniques.

FSK is a method of changing the frequency features of the information signal corresponding to the carrier signal. In this category the information bits transmitted by changing the frequency of the carrier. FSK modulates information signals using two carrier signals due to the fact that FSK is a representation of the information signal to two different frequencies.

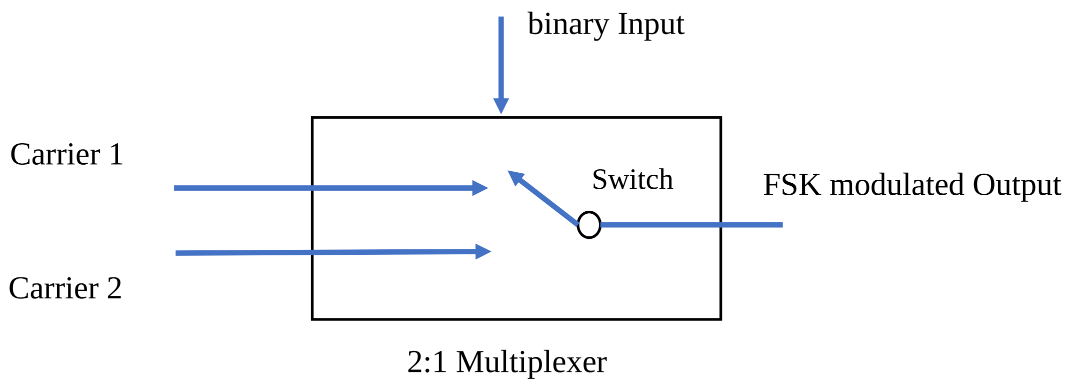

Figure 2.5. FSK block-diagram.

The difference between the two carrier signals input to FSK multiplexer is that one carrier has more frequency than the other carrier frequency, and the switch(s) connected to carrier input 2 for logic 0’s of the input binary sequence. Therefore, the modulated FSK output waveforms have mark and space frequencies.

PSK is a technique used to transmit information signal by modulating the carrier signal phase. This type of modulation is mostly used in data communications because it sends the information data more efficiently than the other modulation techniques. The

PSK technique can be represented in the constellation diagram, and the constellation points can be represented in a uniform angular circle. Constellation point arrangement is allowed to transmit bits by similar energy; this method can eliminate noises and provides superior power for the signal interruption. PSK has vast application and demand than analog modulation. PSK is classified into two main types include Binary phase-shift keying (BPSK) and Quadrature phase-shift keying (QPSK).

BPSK is a type of phase-shifting that uses two phases; each phase is split by 180 degrees. This technique takes the maximum noise level, but it modulates at single bit only for each symbol. This technique is not a suitable choice for modulating information at high rate.

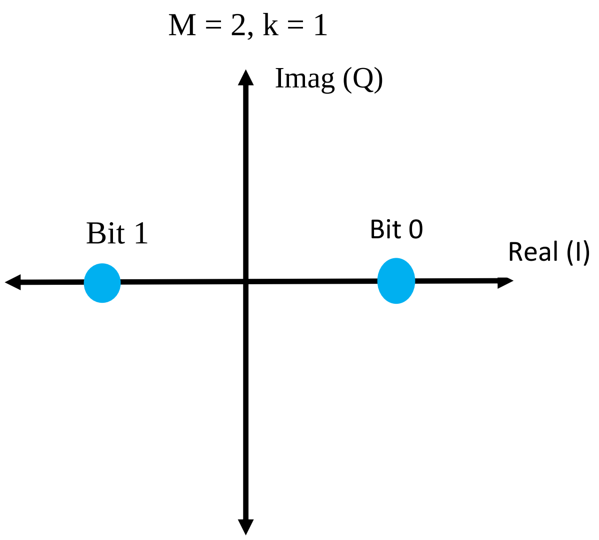

Figure 2. 6. BPSK constellation diagram.

QPSK is the second type of PSK that combining more than one bit on a single-phase section. In this modulation type, the bitstreams are parallelized so that every two bits can be separated up in the same PSK carrier rate. The carrier rate phase can be moved by 90 degrees as of other quadrature phases. Then multiplexer of two-phase shift keying signals is used to add that separated phase to create one of four signal elements.

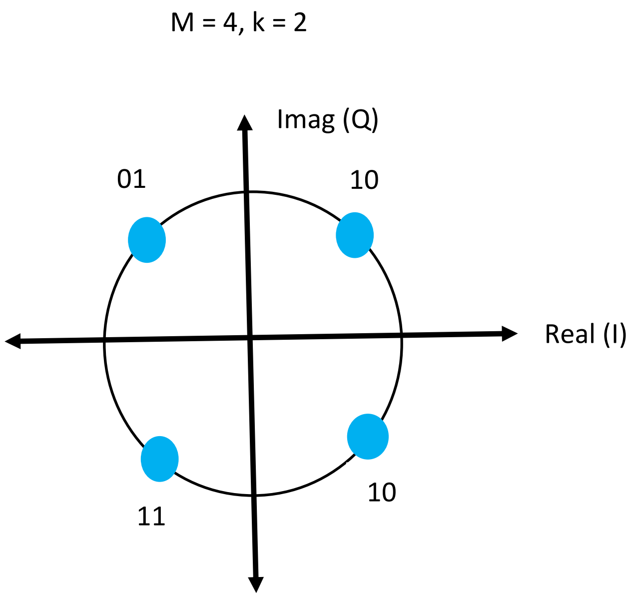

Figure 2. 7. QPSK constellation diagram.

Other PSK forms: in addition to the PSK as mentioned above modulation types, other types have been developed based on the above-mentioned. The idea is to add more constellation points to permit extreme data rates to transfer over the certain bandwidth. However, once the information rates are begun to increase, that means causing the S/N ratio to increase. Those other types are Offset-Quadrature Phase shift keying (O-QPSK), 8-point Phase shift keying (8-PSK), and 16-Point Phase shift keying (16-PSK).

Those are the primary PSK modulation types commonly used in radio communication applications, and each types contains advantages and disadvantages. The variety of the suitable PSK modulation scheme differs upon the communication system's design requirements and conditions. Other typical applications of PSK are in wireless LAN networks such as Bluetooth and RFID, Multiple channel WDM, and more.

Finally, PSK may be defined as a digital modulation method applied widely in today's technologies to securely and efficiently transmit data by changing the phase of a signal frequency known as carrier. This technique is generally used in the field of biometric, wireless LAN.

Quadrature Amplitude Modulation (QAM) is a combination of ASK and PSK to increase the bit rate. This technique uses two sinusoidal waves with the similar frequency in a different phase of 90^o; these two waves in a quadrature phase correlation.

This combination usually refers to one wave as an In-phase wave or (I-wave), and the other wave is known as a Quadrature wave (Q-wave). I-wave is a sine function, and Q-wave is a cosine function. These functions are orthogonal to each other. These orthogonality helps to demodulate the signal at the receiver side without any interference between the two signals.

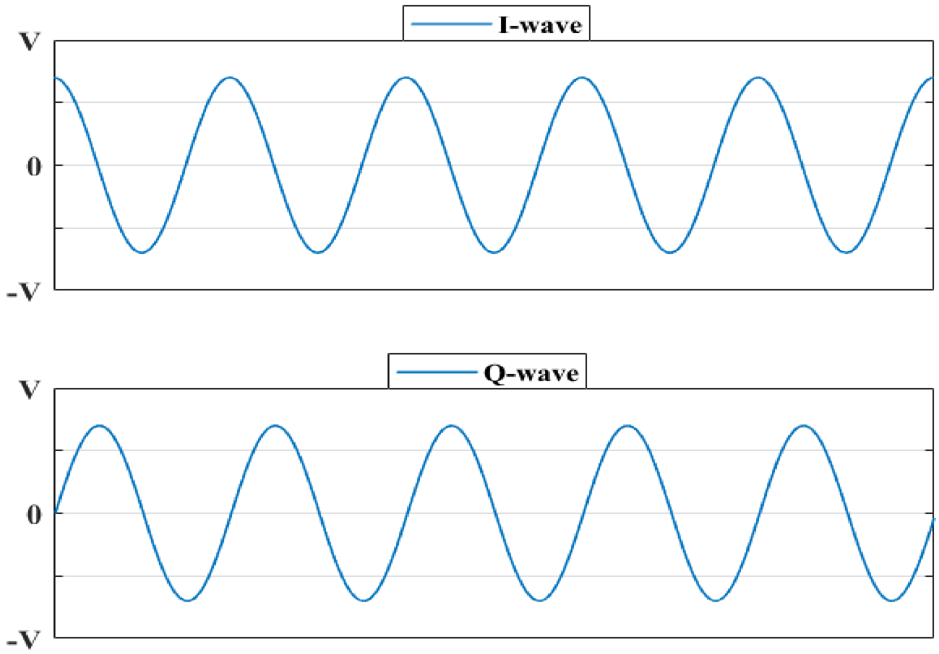

Fig. 2. 8. QAM Orthogonal I-wave and Q-wave

Frown the fig. 2.8. above the orthogonal I-wave to Q-wave is as sine and cosine function, since they are orthogonal, it remains same both are inverted. QAM symbols can be characterized on the signal constellation chart, as demonstrated in fig. 2.9. the plot shows I-wave and Q-waves. In each phase, one of the dots will transmit, for instance, to understand the QAM constellation, fig. 2.9 shows the constellation of 16-QAM; each dot in the diagram represents a symbol.

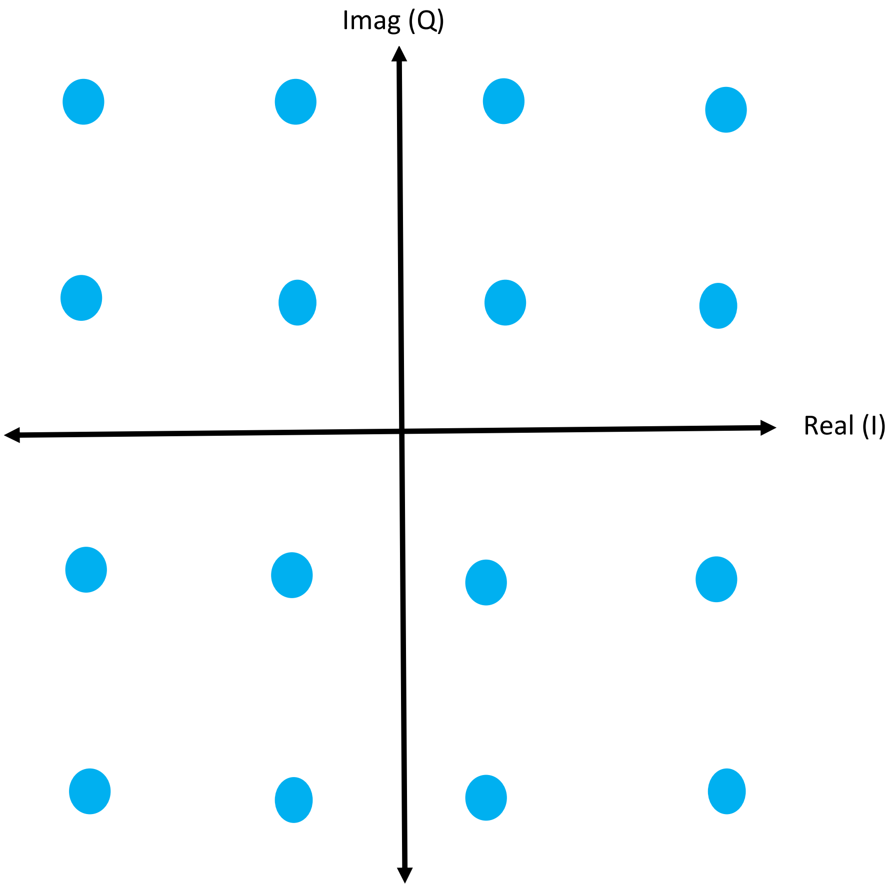

Fig. 2.9. arrangement diagram for 16-QAM.

More data bits be able to transmit per symbol in QAM modulation as long as the number of symbols, known as M-QAM modulation M, represents the number of characters. For example, QAM can constellate in M = [4, 8, 16, 64,256, …], which means the scheme with symbols 4, 8, 16, 256, …, etc. An M-array modulation system with b represents code bits capable of transmitting (2^{b}) unique waveforms at a given sample.

2.1.3. Multiple-Input Multiple-Output Systems

MIMO are used in wireless networks to improve channel capacity by decomposing a waveform, improving errors that occur during the coding process, and transmitting codded information using Tx independent channels. Generally, MIMO systems can work at lower energy and achieve higher spectral efficiency.

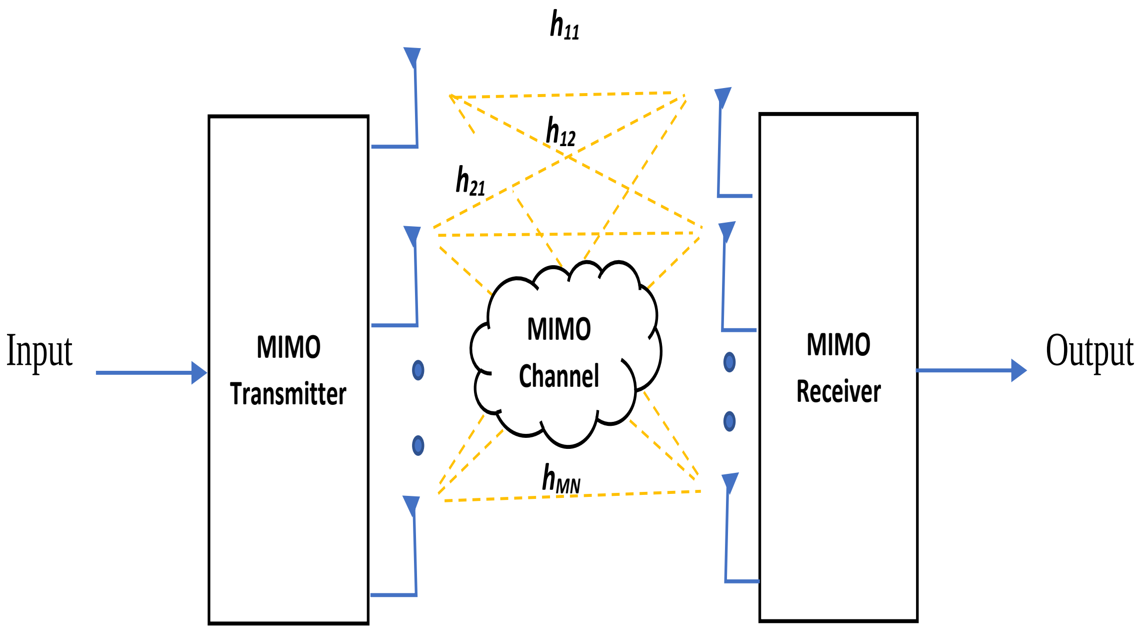

Fig. 2.10. Structure of a MIMO systems with M transmitters and N receiver antennas.

As shown in Fig. 2.10. MIMO uses multiple transmit and receive antennas to model a narrowband channel in the form of matrix H [6]. These matrix elements represent amplitude and phase shift introduced by the channel and denoted by hij, representing MIMO channel connection between i^{th} receiver and j^{th} transmitter antennas. The antenna output signal (y) comes from groups of the input signal (x) via input antennas, where $y \epsilon$ C^{Nx1} and x ϵ C^{Nx1} come from the assumption that H ϵ C^{NxN} can compute from the equation (2.10). where n ϵ C^{Nx1} $\tilde{n}(0, \sigma^2)$ is the channel noise vector, which is modelled as an independent complex Additive White Gaussian Noise (AWGN) random variable with zero means.

$$y = Hx + n \tag{2.10}$$

In [7], the capacity and spectral enhancements by MIMO depend on channel State Information (CSI) available between both Tx and Rx gain of the antennas. Additionally, [8] define the channel capacity of MIMO antennas using an equation (2.11) under the assumption that CSI is available at uncorrelated MIMO Tx and Rx antennas.

$$C_{MIMO} = log_2\left(det\left[I + \frac{\gamma}{Tx}(H^H H)^{-1}\right]\right) \tag{2.11}$$

Where; γ is the SNR at the receiver, and *I* is the unit matrix with the exact dimensions as *H*. Use of MIMO antenna in a digital communication system provides more spatial dimension for the system and gains a degree of freedom, spatially multiplexing. The spatial multiplexing led to several data streams onto the MIMO channel, increasing channel capacity.

To take advantage of the enhanced throughput, MIMO wireless systems employ a matrix technique. In MIMO systems multiple antennas give possibility to transmit data from various path that will reduce the path loss for the channel. Transmitted data is sent in form called data streams which is necessary to enable the receiver to differentiate between the various data packages. These are denoted by the properties h_{12}, which means transmit information from antenna number one in transmitter side to the antenna number two in receiver side. A matrix can be set up in this manner for a three signal to transmit, and receive via antenna in MIMO system:

$$y_1 = h_{11}x_1 + h_{21}x_2 + h_{31}x_3 \tag{2.12}$$

$$y_2 = h_{12}x_1 + h_{22}x_2 + h_{32}x_3 \tag{2.13}$$

$$y_3 = h_{13}x_1 + h_{23}x_2 + h_{33}x_3 \tag{2.14}$$

At the received antenna side, a significant extent of signal processing is needed to recover the transmitted information signal stream at the receiver side. The MIMO system must first approximate the individual transfer characteristic of the channel h_{ij} to calculate the channel transfer matrix H. After estimating all of this, H is formed, and the transmitted data streams reconstructed by reproducing the received vector by the inverse of the transfer matrix.

2.1.4. Receiver Equalisation

Receiver equalization is a method of restoring high-frequency signal components usually attenuated by channels. Generally, signal transmission over band-limited wireless channels is susceptible to record diffusive effects of the channel. Which can cause inter-symbol interference (ISI). There are three important types of linear equalizers used to solve the effect of ISI, specifically, the zero-forcing (ZF) equalizer, the minimum mean squared error (MMSE) equalizer, and the least mean square (LMS) adaptive equalizer.

a. Zero-forcing equalizer (ZF)

ZF equalizer forces the residual ISI in the equalizer output, say *d*[k] to zero. This equalizer forces the remaining ISI at the sampling instants kT (k don't equal *zero).* In practical applications, the perfect ZF equalizer counterpart W(z) behaves as an IIR filter, and the transfer function of the channel H(z) is approximated to a finite response (FIR) filter [9]. The approximation is represented in the equation (2.15).

$$W(z) = \frac{1}{H(z)} \tag{2.15}$$

The transfer function of ZF denoted by Q(z), is presented by

$$Q(z) = H(z)W(z) = 1 \tag{2.16}$$

This thesis uses the idea of the zero-forcing equalizer (ZF) to reject intercarrier, and inter-symbol interference by applying the reverse operation of the channel frequency response to the obtained signal [9]. This equalizer W is given by the equation below.

$$W = (H^H H)^{-1} H^H \tag{2.17}$$

Where the equalizer W is a complex matrix in the exact dimensions as H. the drawback of the equalizer is that it amplifies noise caused by linearly dependent columns [6]. ZF is also unfunctional if $H^H H$ is singular. Since the response of ZF is inverse of the channel response ZF equalizer suffers from noise enhancement, resulted for the large gain will amplifies the AGWN at specific frequencies.

b. Minimum mean square error (MMSE) equalizer

ZF suffers from noise enhancement. An improvement has come to the equalization (ZF) detector in the form of the MMSE detector, which introduces a regularization term $\lambda = \sigma^2$ to the main equalizer equation given above in (2.17), which enhanced the system to be less sensitive to the channel conditions [6].

$$W = (H^H H + \lambda I)^{-1} H^H \tag{2.18}$$

MMSE is applicable to work in both cases where Tx > Rx or Tx < Rx [6]. And the remaining part of the receiver operates in the inverse operation of their respective transmitter parts. The goal of the MMSE equalizer is design is to reduce error variance and bias, for example, the design of FIR MMSE equalizer of size N their transfer function can determine by:

$$W(z) = \sum_{k=}^{N-1} w[k] z^{-k} \tag{2.19}$$

So, the MMSE equalizer seeks to minimize the bias, and variance.

c. least mean square (LMS) adaptive equalizer

the idea of this equalizer is to use the LMS algorithm to adapt the FIR equalizer. The LMS algorithm is developed from the gradient descent algorithm to look for the least amount of error at the output of the equalizer filter. Generally, the adaptive equalizer contains two main parts: linear filter, FIR filter, and adaptive algorithm, which is LMS. LMS algorithm is companied to minimize the noise and intersymbol interference in this equalizer by adjusting the filter coefficients. The brief expiation of LMS algorithm to build an FIR filter of N length is given below.

To design an adaptive equalizer for the channel model with Wn current filter coefficient at instant H transpose and r received complex sequence, the LSM algorithm in training mode is given by equations (2.20) and (2.21). And algorithm 1: shows the summary of LMS [10].

$$e = a[k] - w_n^H r \tag{2.20}$$

$$W_{n+1} = w_n + \mu e r \tag{2.21}$$

Where, $w_n = [w_0, w_1, \dots, w_{N-1}]^H$, and $r = [R_k, R_{k-1}, \dots, R_{k-N-1}]^T$.

Algorithm 1: Least mean square (LMS) algorithm [11]

```
Input: N, µ, R, a
Output: W, E
Booting: W_n = zeros(N,1)
for k = N: length(r) do
        calculate R = [R_k, R_k-1, …,R_k-n-1]^T
        calculate error E = a[k] - W_n^H r
        calculate filter update wn+1 = w_n+µer
Return w
```

2.1.5. Rayleigh Fading Channel

In this work, we simulated our model under the Rayleigh fading model to see the robustness of the codec in scenarios where the signal is considered scattered in the channel between the transmitter and the receiver. In general, Rayleigh fading is a model that can specify the shape of fading that occurs in the wireless channel when multi-path propagation happens. In the multi-path propagation, the objects may cause refract, reflect or scatter the transmit signal. In wireless communication, the transmitters and receivers are frequently not fixed in one palace; this movement causes the path length to change overtimes.

In Rayleigh fading channel model, the elements of H matrixes are assumed to be zero-mean symmetric complex Gaussian random variables with unit variance. It's a more appropriate model for wireless channels for the reason that there is no direct path between the transmitter and receiver. Rayleigh distribution has been used to model scenarios where there is no dominant line-of-sight transmission [9]. Generally, the propagation in any wireless channel is either a line of sight (LOS) propagation or a non-line of sight (NLSO) propagation. Received signal in LOS propagation is under probability density function (PDF) environment which obeys Rician distribution. However, in NLOS propagation environment, the distribution obeys Rayleigh channel distribution. Rayleigh channel components can represent in the complex form to shape the elements matrix H for the entire channel path by the expression as follows [11].

$$h(t) = \sum_{n=1}^{N} a_n e^{j(2\pi f_n t + \varphi_n)} \quad (2.22)$$

where N represents multipath channel and a_n is the amplitude of the n^{th} path. Frequency and phase shifts are represented by f_n and φ_n Respectively. In this thesis, to design a machine learning prediction model, the channel components are separated into real and imaginary parts. So, the model predicts both the real and imaginary parts of the channel state information and stores it separately. This procedure can be represented in equation (2.23).

$$h(KT_s) = h_i(kT_s) + jh_Q(kT_s) \quad (2.23)$$

Once, the state information is stored machine learning model captures channel information $h_i(kT_s)$ and $jh_Q(kT_s)$. Based on the channel information captured, the channel predicts later information based on learning from stored on. More explanation in the coming chapters.

2.2. Adaptive Modulation and Coding

As mentioned in section (1.3), AMC is the main technique used in the system design of this work to adapt the varying signal strengths under the assumptions that the system is linear and fading coefficients to be constant for the entire codeword. Generally, AMC has been applied broadly in wired and wireless communication systems, especially in telecommunication, with several existing solutions to adapt to the variations of the communication channel conditions.

Wi-Fi, for instance, as a short-range communication system, uses adaptive modulation and coding to adapt the channel fluctuations. It flows the basic AMC block diagram depicted in figure (1.1). AMC acts like a controller to control signal transmitting according to the channel conditions. When the channel conditions are good, it trays to transmit at high a data rate as possible and transmits at a low data rate when the channel is poor to avoid having too many dropped packets. The low data rate adjusts under the small constellation such as PSK, and the high data rate adjusts at large constellations such as 16QAM and 64QAM [3].

To explain the functionality of the AMC depicted in figure (1.1), let us assume a single user trying to send a package of data instantly through the channel with a varying signal to interference pulse noise ratio (SINR). The role of AMC is to manage three quantities simultaneously, which is quite challenging duty. These quantities are transmitted power, constellation, and coding rate. In this task, the transmitter needs to know the SINR (γ) which is received SINR (γ_r) divided by the transmit power P_t.

Theoretically, the study of AM (Adaptive Modulation) can develop a sensible guideline for the engineer to develop algorithms based vast simulations. Adaptation algorithms help to increase the throughput of the systems working in a time-varying channel. The coming sub-section is a literature review of the previous work that contributed to this area of research using adaptive modulation and coding techniques.

2.2.1. AMC Previous works

Most standards for wireless communications apply AMC to improving spectrum efficiency, and reliability; among those standards are IEEE 802.11n and IEEE 802.11ac for Wi-Fi networks. in [12], Use SNR based adaption using MPSK and RS codes makes switching between schemes using a channel predictor given by $SNR_1 = \alpha^2(SNR)$ where SNR_1 is the SNR value determined by the transmitter is possible with the exists of the fading amplitude α. In [13] they present a design that considers the medium access control performance requested by an application. They used M-QAM with convolutional code bounds described in [14], and they defined the prediction model under the uses of MMSE as $\widehat{H} = H_{current} + \frac{L_{current}}{L_{Prev}} \Delta H$, where $L_{current}$ and L_{prev} represents the length of current and previous packet length and ΔH is the difference between the current and previous channel matrices. [13] show that the output decoder performance can be modelled using code rate and pre-decoder with uncoded BER. [15] introduced practical considerations such as hardware specifications and pulse shaping, their model is based on power and BER constraints using trellis codes. A more robust algorithm is presented in [16], which considers IEEE 802.11n guard interval sizes. The paper derived the optimal BER value for switching between the MSCs and concluded that their algorithm provides high throughput whilst being reliable. The backbones for many AMC systems are SNR estimation for varying channel conditions whereby receiver requires CSI for a successful AMC.

In [17] the authors used a fully connected neural network (NN) and a convolutional neural network; they achieved a better result than traditional AMC designs for 64QAM using MMSE. And [18] provided the AMC model by using supervised learning to reduce complexities and a deep neural network for channel estimation using channel matrices as training data. Each channel matrix is mapped to its respective transmitter antenna in this design. Results from [18] show that such a design outperforms conventional designs while being less complex. [19] deployed a model that achieved an accuracy of 50% using trace logs as training data, and by using a classification NN with a single hidden and the type of MCS neurons in the output layer. Another contribution is [20] using reinforcement learning-based AMC for underwater communication; their design reduces BER with less energy consumption. [11] explores a NN channel prediction model under different channel models, namely Jakes, Clarke/Gan's, and 3GPP spatial channel model. Their model performs better when using the Jakes model because of Jakes's strong time domain correlation over different SNR values. In order to improve performance and system reliability, [21] suggests an online learning algorithm for AMC that keeps a database of post performances. This approach allows the training data to be generated whilst the system is in use. The advantage is that the model is updated as the channel conditions change as a result accuracy and reliability are also improved.

Recently, AMC has been combined with MIMO and OFDM techniques to improve reliability. it helped wireless systems solve overhead and latency. And meet the upper layer demand for real-time transmissions. Nevertheless, the complicated operational circumstances and the time-varying nature of wireless communication media made the configuration for the physical layer quite challenging in practice. These scenarios lead to two major issues. Firstly, there are various limitations in an applied system, such as nonlinear distortion of power amplifiers and quantization error. Secondly, it is a convention to divide signal processing into blocks. Those two drawbacks cause the computations more expensive [1].

2.3. Machine Learning

This section introduces basic machine learning concepts (ML), from definition to operation. In addition, some important algorithms will highlight to differentiate between machine learning categories.

2.3.1. Concept of ML

ML is a field that incorporates ideas from statistics, computer science, perceptual science, optimization theory, and various other arithmetic and science areas. In his published paper, the Machine Learning concept was introduced in the early 1950 by Alan Turing, highlighting and proposing basic concepts of Artificial Intelligence (AI) and machine learning. Arthur Samuel defined machine learning as a field of study that gives computer the ability to learn without being explicitly programmed; this definition was introduced in 1959. And Tom M. Mitchell widely mentioned a more formal description for Machine Learning said its a program that enables the computer to learn from experience E based on the class of tasks T, and its performance P, and this performance P will improve with experience E.

This software algorithm is used to predict an outcome based on input data. This software does not follow a set of rules to detect or predict the outcome. However, it iteratively learns from data. Generally, ML is classified into dual broad groups, which is Supervised and Unsupervised machine learning. ML uses statistical methods to do so from different scenarios in both supervised and unsupervised for the machine to learn. Deep learning is one of the major subsets of ML, where the algorithms can learn with experience without any external interaction. AI, ML, and deep learning relationships are illustrated in figure 2.11.

Unsupervised ML make decisions from datasets containing of data without considering responses. In other words, it enables computers to learn on their own. In contrast, supervised learning needs to teach the model using training data, which is known as samples, from the considered dataset. So supervised learning challenges is to the fine obtainable relationship between self-regulating input variables and a goal feature. Furthermore, supervised techniques are classified into two main categories: regression and classification. Coming subsection will give a brief introduction and explanation of regression and classification in supervised learning.

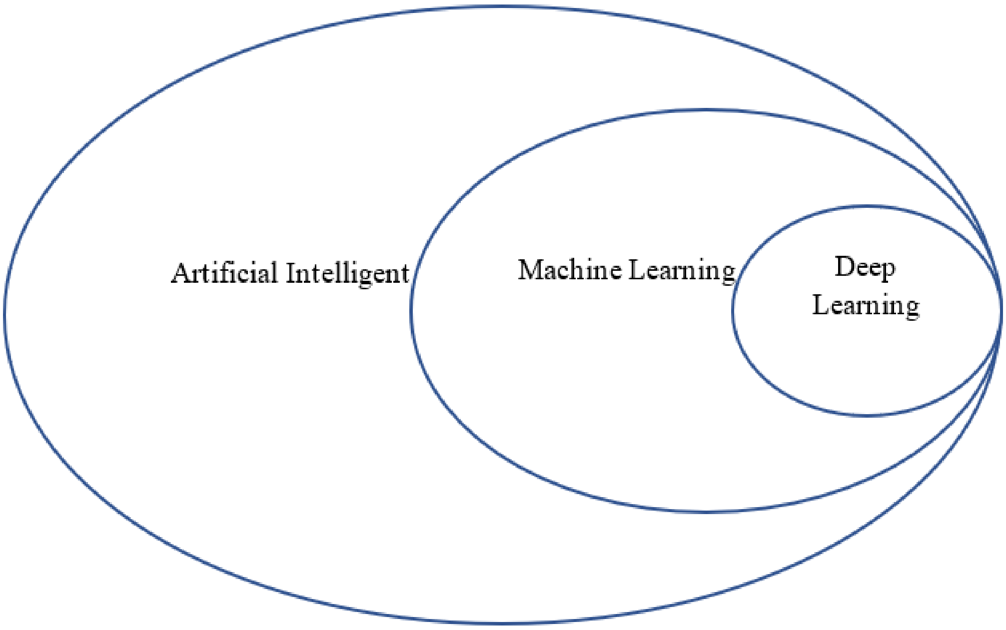

Fig. 2.11. the relationship between AI, ML, and DL [22].

Generally, ML can make predictions to future events based on data collected in real-time in the supervised learning method. And in Figure 2.12. illustrated the overview of the ML is depicted [23].

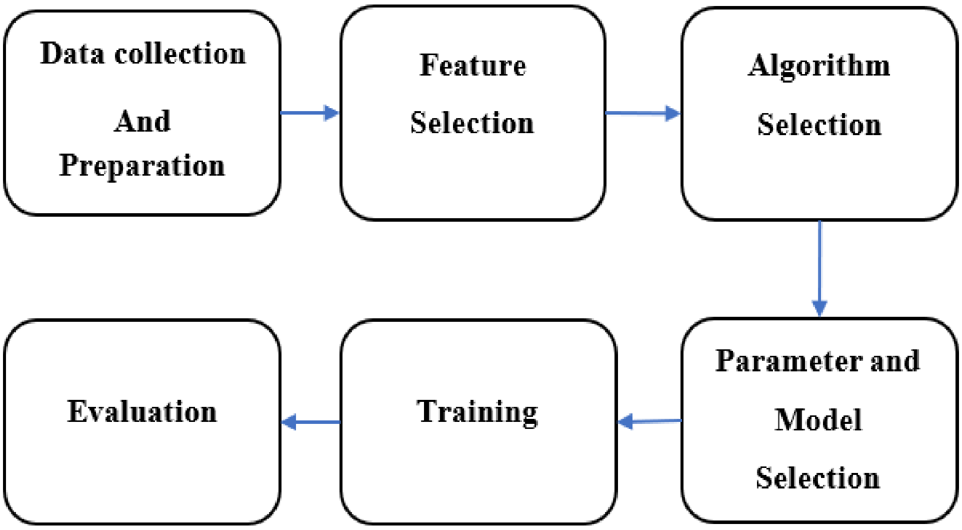

Fig. 2.12. Process of Selection and Evaluation of the Algorithms in Machine learning

In Fig. 2.12 dataset is an essential thing in ML; it is collected from a significant amount of data which is preferable to be without noise to label it in the form of a column. Collected data labeled in the dataset has features which is helpful to predict a future decision, that every application demands are different to fulfill the correct output. The next step is to choose a suitable logarithm to finalize the model.

In ML, some algorithms need a specific set of parameters in order for the model to be selected. The dataset is used to train and predict the result for the training process by splitting the dataset into two parts based on the size of the training set. The test size controls samples meant to be used for training and testing; for this, it's considered an important parameter for every ML algorithm intending to compensate for the effects of overfitting laid out of sample accuracies. In order to increase model accuracy, the test part should be included in the training part. Therefore, the test set is extremely reliant on the training set. In the evaluation step, algorithms output are the predicted values compared to the test set equivalent, resulting in the evaluation of two accuracy parameters: training accuracy and out-of-sample accuracy rate.

- The training accuracy is the percentage of correct assumptions using the same dataset. A high training accuracy does not correspond to the typical requirements resulting in an overfit model.
- The out-of-sample rate accuracy is the percentage of correct prediction from outside the dataset. This rate should be high enough to generalize the model for real-time scenarios.

Machine learning is going to make a difference in the technologies of today. Therefore, the coming subsection focuses on studying machine learning algorithms that have enormous contributions to the various applications applied in different fields.

2.3.2. ML Algorithms

Various algorithms for several applications exist in the ML libraries. So, according to the applications, the algorithm should be carefully chosen to acquire the best solution with minimum complexity. Flowing is highlighted on ML algorithms.

2.3.2.1. Supervised Learning

In Supervised learning, the algorithms are commonly used for regression or classification tasks.

a. Regression

The unsupervised algorithms for regression tasks predict continuous values depending on the past dataset. In regression, the dataset features a curve with two sets of dependent and independent variables. This process is called data set curve fitting. Regression is based on linear or nonlinear variables. According to those variables, the operation can be simple or multiple regression. In simple regression, the predicted output is based on a two-dimensional curve, curveting sample variables. However, in multiple regression, the curvetting occurs in more than two dimensions. Regression will be complicated with a multiple-dimensional curve.

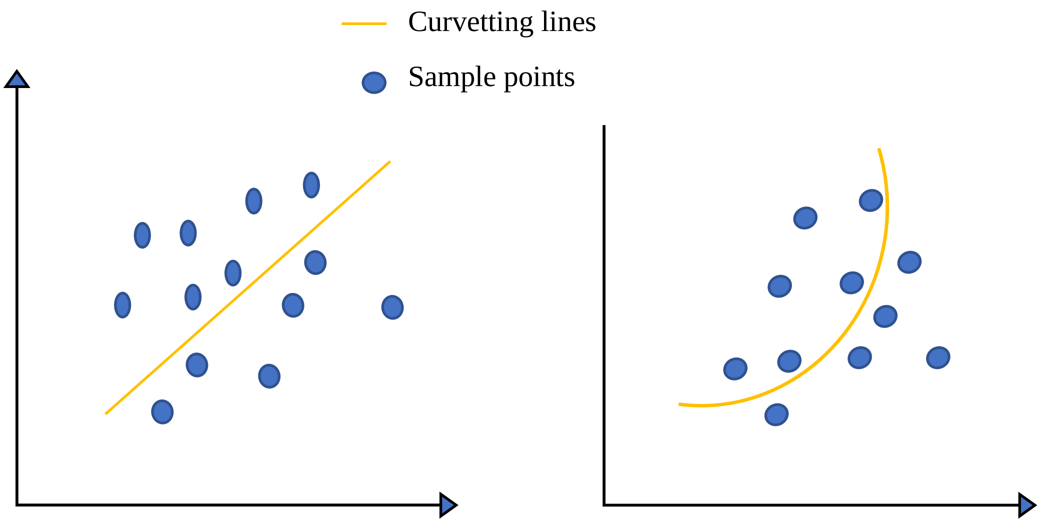

Fig. 2.13. the linear and nonlinear curvetting of samples [23].

The regression is said to be linear if and only if a linear function defines the relationship between the output-input parameters. However, it will be non-linear when non-linear functions define the relationship between the input and output parameters as quadratics, exponentials, and polynomials.

b. Classification

In ML, classification algorithms are means to classify unknown variables into a classes in discrete form. Those unsupervised classification algorithms try to learn the connection between a number of feature variables and a targeted variable [24]. In classification, the input data is separated into dual or further classes, and the learning requirement yields a model that allocates in the hidden inputs to one or more classes.

Data classification is widely applied and has various applications in the modern manufacturing categories. In classification, many issues that developers are trying to solve using classification algorithms can be stated as the association of a feature and a target variable; this association gives a vast range of solutions to various scenarios. For instance, human activity prediction is a classification application, wherever the inputs are accelerometer signals, and output is the classes that represent activities such as laying, standing, walking, climbing stairs, and sitting. The most used classification algorithms in machine learning are Naive Bayes, decision trees, K-Nearest Neighbours (K-NN), logistic regression, support vector machines, and Neural Networks [24].

- Naïve Bayes is a classification logarithm that considers the assumption of Maximum A Posteriori (MAP) for solving classification problems.
- Decision Trees splits the training data into separate nodes. And each node has a relation to the node above it.
- K-Nearest Neighbours (K-NN) is a non-parametric algorithm that calculates the alien input's distance from every other training sample to make decisions.
- Logistic Regression is a statistical method established upon the values of the input fields.
- Support Vector Machine (SVM) is projected as one of the appropriate procedures for resolving issues correlated to classifying a new unknown data set by learning and forecasting. The algorithm is named after its execution of the data by a hyperplane in vast dimensional astronomical, which is known as a data vector. The basic idea behind classifying the data using SVM is to produce a function to divide the data points into the matching labels with the smallest probable number of faults or through the huge probable boundary [24].
- Neural Networks (NN) are based upon the log-likelihood function. This algorithm represents the training data as a column vector in the input layer. The parameters associated with a NN are usually real and positive. In NN and activation function use to decide which neurons should be active during forwarding propagation. A neural in layer says l will be active if the output layer exceeds the threshold value. More elaboration on NN is introduced in chapter three. These algorithms mentioned above can be used both for regression and classification. Those logarithms are considered one of the uppermost used to resolve issues related to classifying or predicting new unknown data by learning and forecasting.

2.3.2.2. Unsupervised Learning

Unlike supervised learning, no labels are prearranged to the learning algorithm in unsupervised learning. Unsupervised algorithms discover the structure of the input data on their own. In other words, the algorithms do not have detailed information about the inputs' nature or status.

The goal of the unsupervised method centered on finding clusters or similarities in the input data and accordingly categorizing the unlabelled values in related classes. In addition, this method focuses on determining the allocation of data within the input dataset; this process is known as density estimation [23]. Generally, the model does not need to supervise the model in unsupervised learning. Unsupervised algorithms can broadly be categorized into two types of tasks, as in figure 2.14.

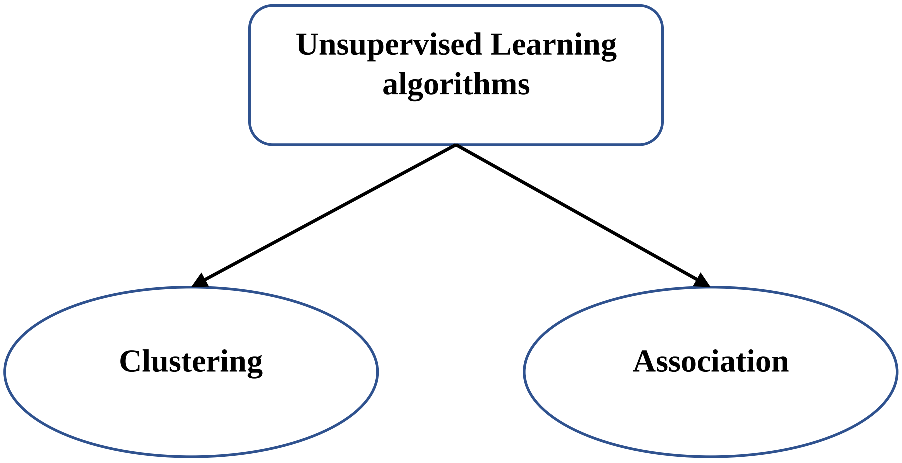

Fig. 2.14. unsupervised learning algorithms categories.

Clustering is a technique of grouping the object into constellations so that objects with the most similarities remain in one group and have little or similarities with objects in another group. Cluster analysis discovers commonalities between data objects and categorizes them based on the presence or absence of those commonalities.

Association in the other hand, it is an unsupervised technique that focuses on finding the relationships between variables in large datasets. Association techniques are widely used to predict marketing strategy and market analysis.

The most know unsupervised algorithms are K-means clustering, neural networks, k nearest neighbors (kNN), anomaly detection, and more [23].

CHAPTER 3

DESIGN PROCEDURES AND CONSIDERATION

In order for the proposed design to be practical and feasible to implement, certain design decisions are taken based on the following components.

3.1. Design Requirements

The main goal of the project is to design a model for a Wi-Fi router to meet the following requirements:

- The codec must achieve a rate of at least 2 bits/Hz.
- Rayleigh fading channel with a bandwidth of 4MHz.
- The data rate must be at least 10 Mbits/s.

3.2. Design Assumptions

To reduce the complexity and scope of this work, the following assumptions are made in developing the AMC simulation.

- The channel is quasistatic, which means that the fading coefficients stay constant for the entire codeword.
- Both the transmitter and receiver are stationary.
- CSI is known at the receiver antenna; as such, there is no feedback channel needed.
- The antennas are independent of each other. That is, the channel has unique coefficients for each antenna.
- The data to be transmitted has been compressed and encrypted.
- Hardware exists to support the proposed codec.

In general, the project's success is mainly dependent on designing a simulation model that meets the requirements. And the type of MCS, MIMO architecture, and machine learning model should be justified the design process and parameter choices.

3.3. Codec Adaptive Modulation and Coding Model

The design runs at a frame level, with each frame consisting of a sequence of randomly generated bits. This work uses BCH codes with symbols from binary filed $GF(2^m)$ where m = 4. The encoder used is a BCH with message length k = 5 and codeword length that is n = 2m – 1, which is capable of correcting up to 3 errors, given by $t \geq \frac{n-k}{m}$. As a result, the stream is encoded at a rate of 1/3. BCH codes are chosen over other simple codes, such as RS, because of better performance over fading channels [24]. The modulation scheme used in this project is QAM. QAM utilizes both ASK and PSK; it provides better spectrum efficiency than any of them separate.

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ \varkappa_2 \end{bmatrix} + \begin{bmatrix} n1 \\ n2 \end{bmatrix} \tag{3.1}$$

In order to increase channel capacity and spectral efficiency, the proposed AMC uses MIMO multiplexing. The AMC makes use of a 2x2 MIMO; this is Tx = 2 and Rx = 2. Equation (2.1) can be rewritten as Equation (3.1) for simulation purposes. Equation (3.1) shows the exact structure adopted in the code.

Three different MCS are used in the simulation. These subsystems are combined to make up the AMC, and Table 1 shows the structure for each MCS. MCS1 is the only one that does not make use of FEC.

Table 3.1: Modulation and coding schemes for the AMC

MCS	Code rate	QAM order	MIMO	SE	Es
MCS1	1	4	2x2	4	2
MCS2	1/3	8	2x2	2	6
MCS3	1/3	16	2x2	2.67	10

The MCS parameters are chosen such that each of them meets the requirements defined in Section 3.1. Equation (3.2) shows that the selected MCS meets the 2 bits/Hz rate. Using a pulse shaping coefficient of α_P=1, MIMO coefficient $\alpha_m = 2$, M is the modulation order, and R represents the FEC code rate. MCS1, MCS2, and MCS3 have spectral efficiencies of 4 bits/Hz, 2 bits/Hz, and 2.67 bits/Hz, respectively.

$$SE = \frac{R\,log_2(M)\alpha_m}{\alpha_P} \tag{3.2}$$

The AWGN is generated at the beginning of each frame. The noise variance $\boldsymbol{\sigma_n^2}$ dependent on SNR $(\boldsymbol{\gamma})$ and the MCS used. The noise variance for each MCS can be determined using Equation (3.3).

$$\sigma_n^2 = E_s \frac{n}{k\,log_2(M)} \cdot 10^{\left(-\frac{\gamma}{10}\right)} \tag{3.3}$$

Where Es is the average energy of the symbols after modulation, the energy can be computed by taking the mean of the distances squared of the constellations. The average energies for 4QAM, 8QAM, and 16QAM are 2, 6, and 10.

$$C_{n,k} = \frac{n}{k} F \tag{3.4}$$

$$s_M = \frac{c_{n,k}}{log_2(M)} \tag{3.5}$$

$$s_{MIMO} = \frac{s_M}{Tx} \tag{3.6}$$

The frame structure of the simulated packets is described using equation (3.4 – 3.6). Where F is the frame size, $C_{n,k}$ is the number of coded by bits from an (*n*, *k*) encoder, S_M is the symbols from an M-array modulator, and S_{MIMO} is symbols from each antenna. The simulation uses a frame size of 2000 bits for each MCS. Table 2 shows the frame structure for each of the MCS.

We use an MMSE equalizer based on the reasons already discussed in Section 2.3 to remove errors introduced by the channel. The output signal X' from the MMSE equalizer is defined using Equation (3.9). In order to recover the transmitted signal x, the respective inverse operations are performed on X' as illustrated in Figure 2.10.

$$X' = wy \tag{3.7}$$

Table 3.2: Modulation and coding schemes frame structure.

MCS	Coded bits	QAM symbols	Antenna symbols
MCS1	2000	1000	500
MCS2	6000	2000	1000
MCS3	6000	1500	750

In order for the proposed model to be adaptive to channel conditions, a recursive model is provided in Equation (3.8) that relates the current channel $\mathbf{H_f}$ to the previous frame channel $\mathbf{H_{f-1}}$. Equation (3.9) models a discrete channel update; it follows that the channel update can be written as an ARMA process that can model the correlation in a discrete linear process.

$$H_f = \alpha H_{f-1} + (1-\alpha)G_f \tag{3.8}$$

The extent to which the channel is correlated is controlled by $\alpha \in (0,1)$ which is the fading steady-state coefficient. A large α means a high channel correlation, and a lower means a low channel correlation. The parameter $\mathbf{G}_f$ introduces random channel errors using conditions stated in section 2.4. the channel starts to fade from the second frame as such $\mathbf{H}_f$ is initialized to an identity matrix for the first frame. $\mathbf{H}_f$ and $\mathbf{G}_f$ are determined at the end of the frame.

We define the adaption parameter ξ, which measures the effect of the channel and equalizer to the SNR. The adaption parameter specifies the range at which a given MCS is active. The adaption parameter is given a measure of the effect of the channel and equalizer **Z = HW** to **SNR.**

$$\xi = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\left|H_{ij}\right|\right)^2}{\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\left|W_{ij}\right|\right)^2} \tag{3.9}$$

A translation is done from the SNR vs BER curves to a set of ranges in ξ. The translation rule is the most spectral effective MCS should be active for the majority of the frames. Using the model in Equation (3.8) it can be shown that ξ has a variation centered at the initial value of $\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\left|H_{ij}\right|\right)^2$.

3.4. Machine Learning for (AMC) Modelling

Machine learning has the ability to model complex functions without explicit programming. A closed-form AMC predictor will have superior performance than a machine learning-based predictor. However, it does not generalize well to unknown channels.

3.4.1. Structure of conventional NN

The designed AMC predictor is adopted to flow the unsupervised machine learning class in this work. The machine learning model is expected to use training data in this class to build a function mapping inputs to outputs. The neural network takes the training data x a column vector in the input layer. The parameters associated with a NN are usually real and positive. The hidden layer takes the outputs al-1 from the previous layer process and then passes them to a successive layer during forwarding propagation. This process continues until the output layer. Figure 2.3 represents the neural network layers.

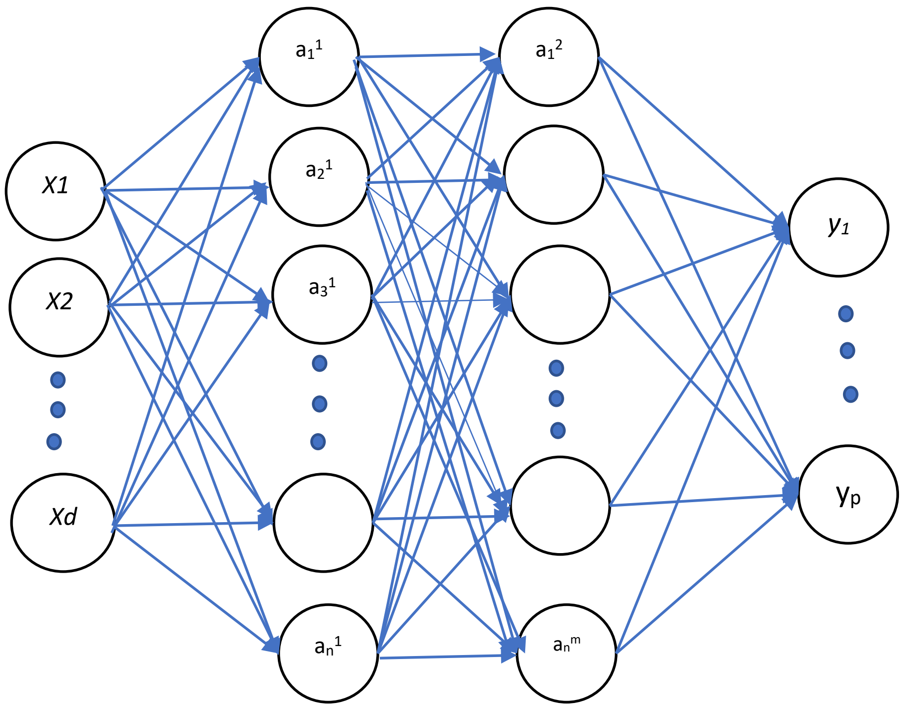

Fig. 3.1. Neural network structure with two hidden layers.

In Neural Network, an activation function is used to decide which neurons should be active during forwarding propagation. A neural in layer ***l*** will be involved in a is greater than some threshold value. Equation (3.10) determines the action for neurons in layer hidden layer ***l***.

$$a^L = \sigma(b^L + w^l . a^{L-1}) \tag{3.10}$$

The activation function σ is a non-linear function, where $\mathbf{b}^{l}$ is the bias row vector for neurons in layer l, and $\boldsymbol{W}^{\boldsymbol{l}}$ is the weight matrix for weights between layer ***l*** and ***l*-1**.

The operator (**.**) denotes the dot product between two matrices. The Sigmoid, SoftMax, and ReLu are the most commonly used activation functions. The weight is proportional to how much contribution the neuron in the previous layer has to the neuron in the next layer. The accuracy of the neural network is determined at the end of each forward propagation epoch or iteration. The cost or loss function *L* determines the difference between the predicted value ^**y** and the expected value ***y***.

Forward propagation is not enough since it is a one-way process used to get prediction values. The actual training of the NN is done through a process known as Backpropagation. Backpropagation is used to optimize ***L*** such that the error is minimum. Backpropagation using gradient descent aims to find the global minima of L, representing the state when the NN has a minimum error. Gradient descent computes the derivative of the cost function with respect to the parameters of interest in the weights and biases.

$$\omega' = \omega - \eta \frac{\partial \mathcal{L}}{\partial \omega} \tag{3.11}$$

$$b' = b - \eta \frac{\partial \mathcal{L}}{\partial b} \tag{3.12}$$

Equations (3.11) and (3.12) provide the update rule for adjusting the weights and biases for neurons in backpropagation. The learning rate η is the step size between points at which the gradient is computed. The learning rate should be set such that the model converges to a local minimum at a reasonable rate. Generally, there is a trade-off between accuracy and generalization of a model. The model should capture the fundamental relationship without capturing outliers of a dataset. The model should be able to perform at comparable accuracy when given unseen data.

The second proposed predictor uses a Short-Term Long Memory (LSTM), a Recurrent Neural Networks (RNN) variant. RNN allows information to persist in the neurons. This is in contrast to a simple NN, which only depends on the current input. LSTM enables a RNN to long-term dependencies in data. LSTM contains structures called memory cells, which modify the information in the network. Gates controls the information flow in a memory cell. The forget gate is used to discard information no longer needed by the memory cell. The input gate dictates when new information is added to the memory cell, and the output is responsible for determining the output based on input data. The operations used in an LSTM are similar to those in a traditional NN, with slight variations to accommodate additional computational units. Detailed LSTM operations are described in [25].

3.4.2. Structure of LSTM-NN

Since the conventional feed-forward neural network (FNN) does not have sequences or loops, its ability for sequential modeling is relatively poor. In contrast, RNN is a type of neural network capable of dealing with time-varied data. The previous input sequences can specify its internal memory and the feedback chain structure that makes prediction output. For channel prediction purposes, RNN is enforceable to accomplish better performance than FNN. Nevertheless, in RNN, the back-propagated error speedily vanishes [26]. Therefore, RNN is not suitable for solving problems requiring long-term temporal learning dependencies. To overcome such an issue, the LSTM-NN is proposed [27].

Contrasted to RNN, the LSTM-NN controls the passing data through three gates over the sequence, and its internal loops can store the long-term dependencies previously learned from input data. LSTM-NN basic architecture is shown in figure 3.2.(a). during each time, the state of hidden layer $\mathbf{h_L}$ updates with input data $\mathbf{x_i}$ at present and the hidden state $\mathbf{h_L}$ of the last moment:

$$h_L = \sigma_h(w_{ih} + w_{hh}h_{L-1} + b_h) \tag{3.13}$$

Where σ_h is the activation function, w_{ih} weight matrixes between the input to the hidden layer and w_{hh} are between two consecutive hidden states, and b_h representing the biases function.

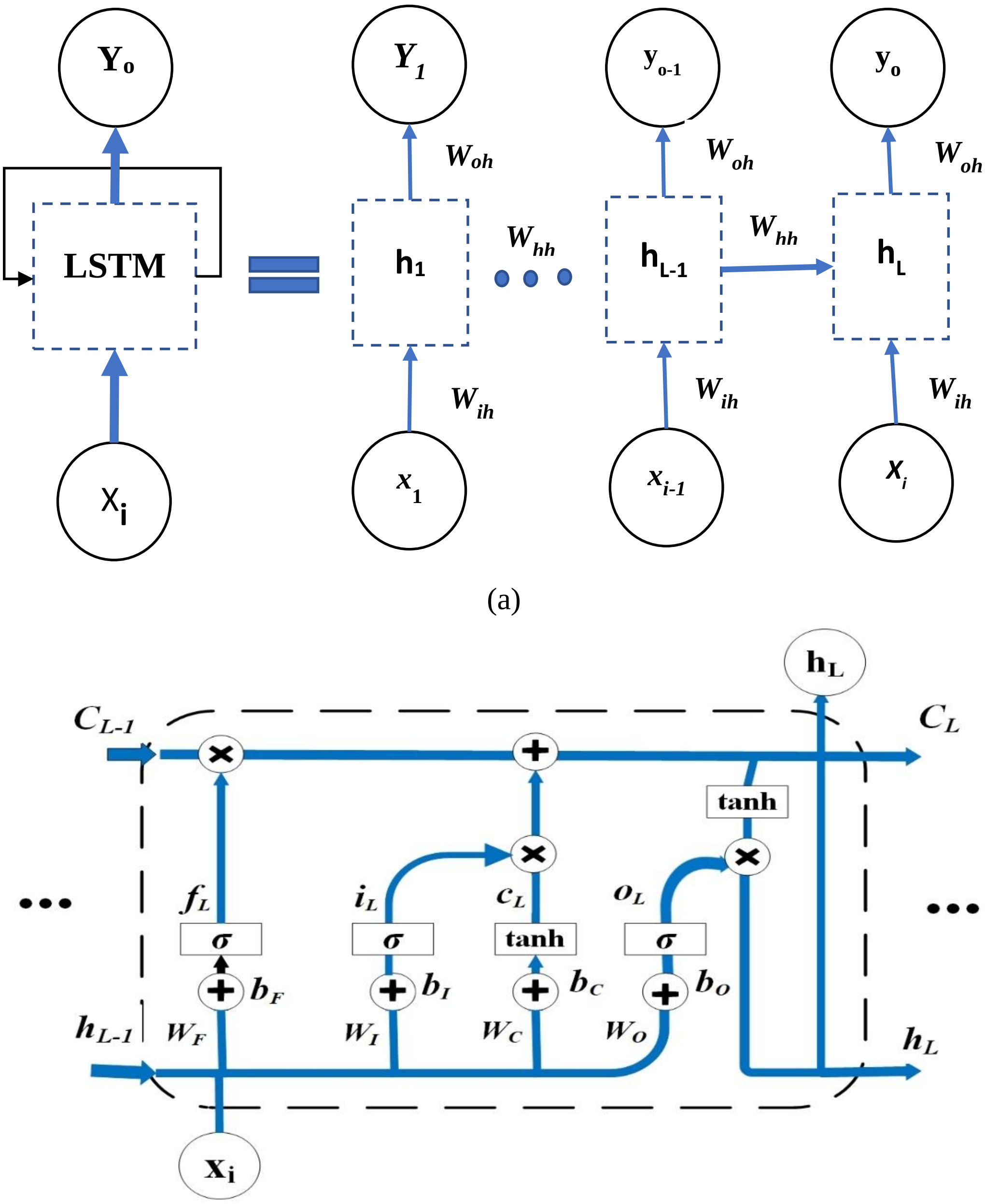

Fig. 3.2. (a) Basic LSTM-NN structure. (b) LSTM cell structure [28].

Fig. 3.2. (a) shows the unfolded structure of LSTM-NN with Xi and Y_o as the input and output vector, respectively; where h_L represents the hidden state; W_{ih} and Woh are the weight matrix between the input-to-hidden layer and hidden-to-output layer, respectively. LSTM cell structure is shown in the fig. 3.2. (b), LSTM has three gates: forget, input, and output gates. The forget gate f_L determines what state information at time L-1 (C_{L-1}) is to be forgotten. Then the input gate $\boldsymbol{i_L}$ adopts which information is updated, and the cell input state $\tilde{C}_L$ A tanh neural network layer will produce it. After that, the cell state at this time C_L It can be calculated. Lastly, the output gate $O_L$will filter the output h_L to determine

the cell state C_L. During the reiteration of the all-time steps from 1 to L, the output sequence $y_0 = [h_1, h_2, \ldots, h_L]$ Is computed.

$$\boldsymbol{f}_L = \sigma_g(W_F[h_{L-1}, x_i] + b_F) \tag{3.14}$$

$$\boldsymbol{i}_L = \sigma_g(W_I[h_{L-1}, x_i] + b_I) \tag{3.15}$$

$$\boldsymbol{o}_L = \sigma_g(W_O[h_{L-1}, x_i] + b_O) \tag{3.16}$$

$$\tilde{\boldsymbol{c}}_L = tanh(W_c[h_{L-1}, X_i] + b_c) \tag{3.17}$$

$$\boldsymbol{c}_L = f_L * C_{L-1} + i_L * \tilde{C}_L \tag{3.18}$$

$$\boldsymbol{h}_L = O_L * tanh(c_L) \tag{3.19}$$

Where $\sigma_g$is the activation function of the gates, W_F, W_I, W_O, and W_c Represent the weights matrices between the hidden layer to these three gates. While b_F, b_I, b_O, and b_C are representing bias functions. LSTM-NN can learn from the previous input data because its cell structure can keep data for long-term dependencies. Therefore, it is suitable for the time-varied sequence prediction [28]. Because LSTM-NN cell structure has the ability to store the input information for log-term sequential dependencies learned from prior input data. In this work, the designed AMC predictor is adapted to flow the supervised machine at a specific time and can be precisely predicted from the signal randomly generated from the bits sequence at this moment and previous signal sequences using the LSTM-NN cell.

3.5. Predictor Design

The NN used in this project is for regression. The NN has trained a train with a window size of 4. This means that each four ξ points are needed to predict the next value of ξ. The value of the expected ξ will be used to select a specific MCS used in the next frame. A similar approach is used for LSTM. The final parameters are chosen such that the output from the loss function is minimum. Both predictors use Root Mean Square Error (RMSE) as a loss function, calculated using Equation (3.20) over n data points. MCS classification is done using the ξ values determined by the predictor using the range of ξ values previously determined in Section (3.3).

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y} - y)^2}{n}} \tag{3.20}$$

CHAPTER 4

SIMULATION RESULT AND ANALYSIS

The first task carried out in this work is determining the modulation orders for the different MCSs. Figure A.1 and Figure A.2 – A.4 in the appendix are used to decide the QAM levels used in the simulation. The plots show the SNR vs BER curves for the MCSs in a non-fading channel without AWGN, we observe that spectral efficiency is inversely proportional to modulation. We expect the spectral performance to be similar when fading is included. Initially, the simulation is run in order to determine the effective ξ and SNR regions for each MCS.

Using Equation 3.8. We saw that if $\boldsymbol{\alpha}$ is closed to a value of 1, then the channel is almost perfectly correlated since the noise term $\boldsymbol{G_f}$ will have a negligible effect on the channel conditions in the current frame. Figure B.3 shows a high correlation, and we observe an almost linear relationship between 10log ξ and the frame progression. A small α means that the channel is uncorrelated, and the noise term significantly affects the channel. This can be seen in Figure B.4, where ξ has a large variation between successive frames. The α value is chosen such that it strikes a balance between the two extremes. Ultimately, we set on $\alpha = 0.96$ and $\sigma_p^2 = 3.$ The channel variance σ_g^2 is kept constant in order to make comparison simpler. The specific value of σ_g^2 was experimentally determined through an iteration process. Figure 4.1. shows ξ with the chosen channel model parameters settles on.

Using figure 4.1. the MCS regions are defined to be MCS1 for ξ ϵ [0, -15), MCS2 for ξ ϵ (0, -15) and MCS3 for ξ ϵ (-∞, -15]. Furthermore, we observe that the value of 10log (ξ) are centred around 10log ($\Sigma_{i=1}^{N} \sum_{n=1}^{N}(|H_{ij}|)^2$) = 0, which is more clearly in Figure 4.4 with most points saturated around the horizontal axis.

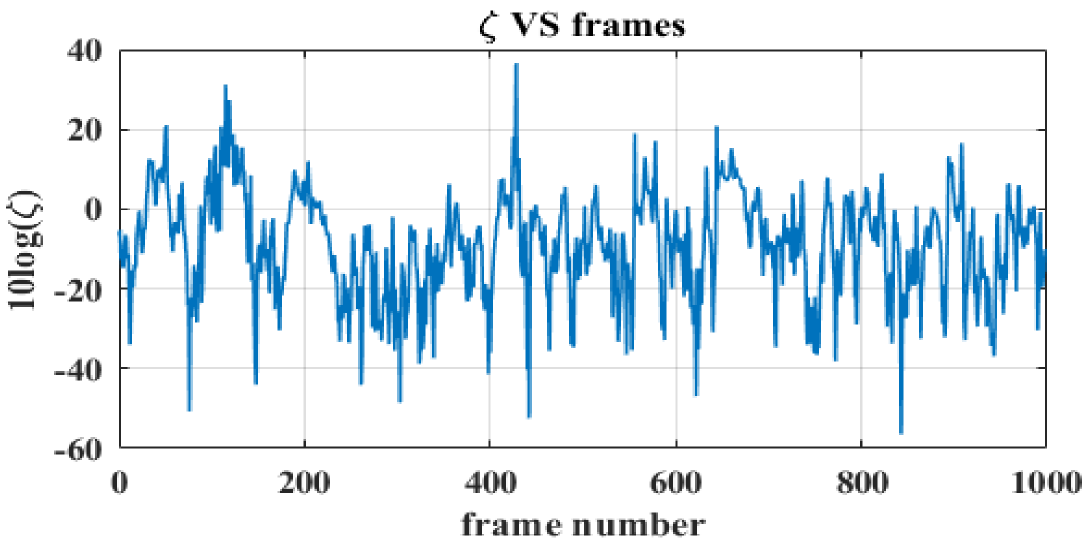

Fig. 4.1. ξ vs frames at low correlation $\alpha\ = 0.96$ and $\sigma_p^2\ = 3$.

The table below shows the operating ranges for each MCS system. These ranges are used to decide which MCS to employ at a given frame.

Table 4.1: MCS operational range

MCS	**ξ range**	**SNR range**
MCS1	[0, ∞)	(-∞, 32]
MCS2	(0, -15)	(32,37)
MCS3	(- ∞, -15,]	[37, ∞)

The SNR at which the received signal experiences significant BER curves are shifted to the right by a factor of 10log ($\Sigma_{i=1}^{N} \sum_{n=1}^{N}(|W_{ij}|)^2$) due to the fact that steady-state coefficient fading is less than 1. Therefore, the curves generated in this figure are used to determine the BER threshold for the AMC. We select a BER threshold of 10^{-3} which means 100 error bits for every 100,000 transmitted bits. The effective AMC SNR ranges can then determine using the BER threshold and MCS ξ regions as scaling conditions. The ξ = -11, which is determined through the simulation, is translated to a SNR value of 30 dB, which is the point at which the left most curve meets the BER threshold. The SNR regions are then defined to span a proportional region to ξ. The regions are defined to be MCS1 for SNR ϵ (-∞, 32], MCS2 for SNR ϵ (32,37) and MCS3 for SNR ϵ [37, ∞). The perfect predictor is developed to be the most optimal AMC. It predicts the channel conditions by calculating the ξ for the next frame and then selecting the appropriate MCS based on that value.

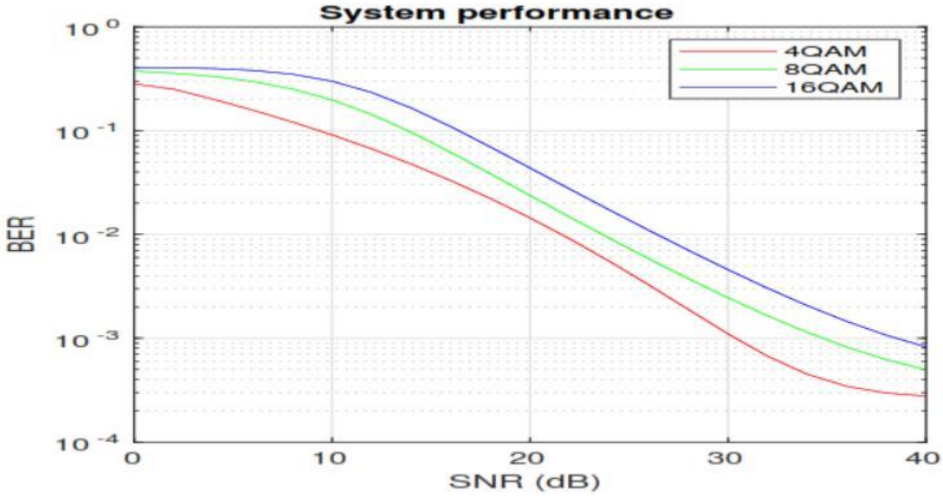

Fig. 4.2. SNR vs BER curves for the selected MCS at $\alpha = 0.96$ and $\sigma_p^2 = 3$.

As such, the perfect predictions are able to optimal MCS for a given frame. The perfect predictor is bounded by the 4QAM and 16QAM as expected with a shape that maximizes the effective data rate.

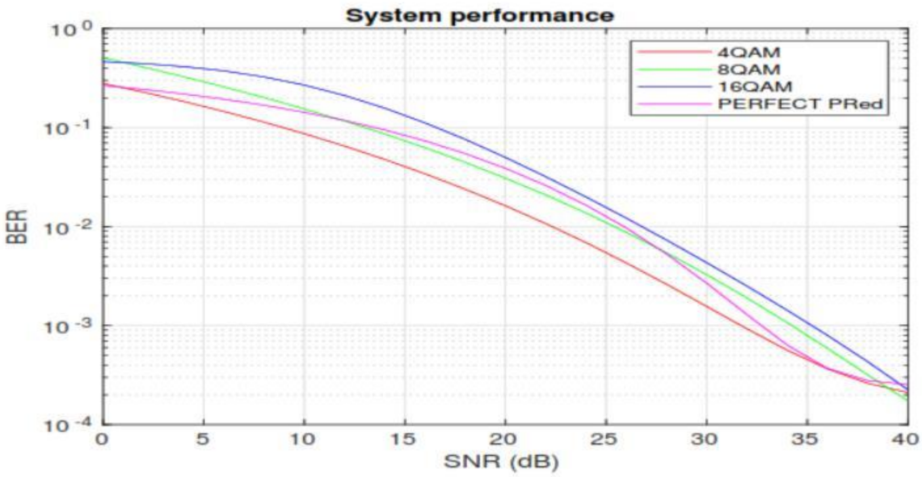

Fig. 4.3. Perfect AMC predictor performance.

The machine learning data set is generated by obtaining Figure 4.1. and Figure 4.2. the data set is slit such that 90% is used for training and the rest as testing data. Each of the predictors also gives the MCS classifications for each ξ. The specific parameters used for the neural network and LSTM were empirically determined.

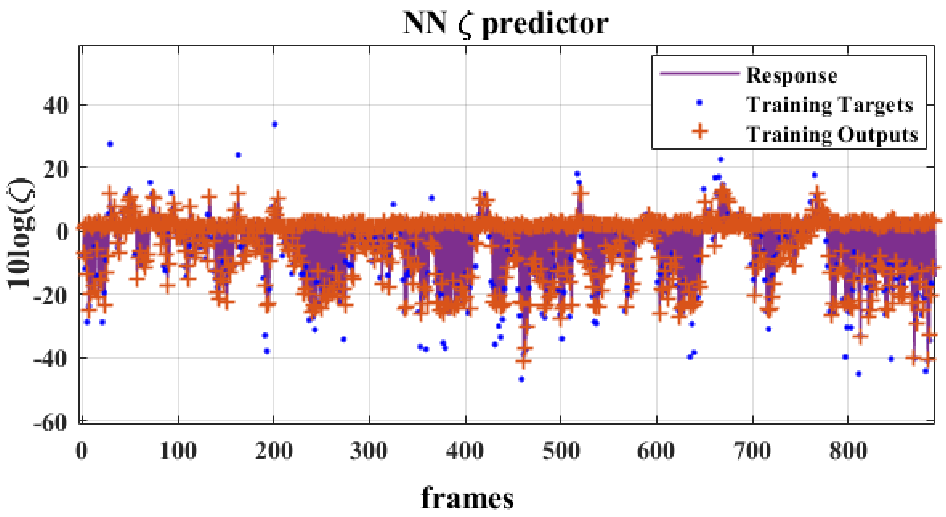

Fig. 4.4. Neural network predictor Response.

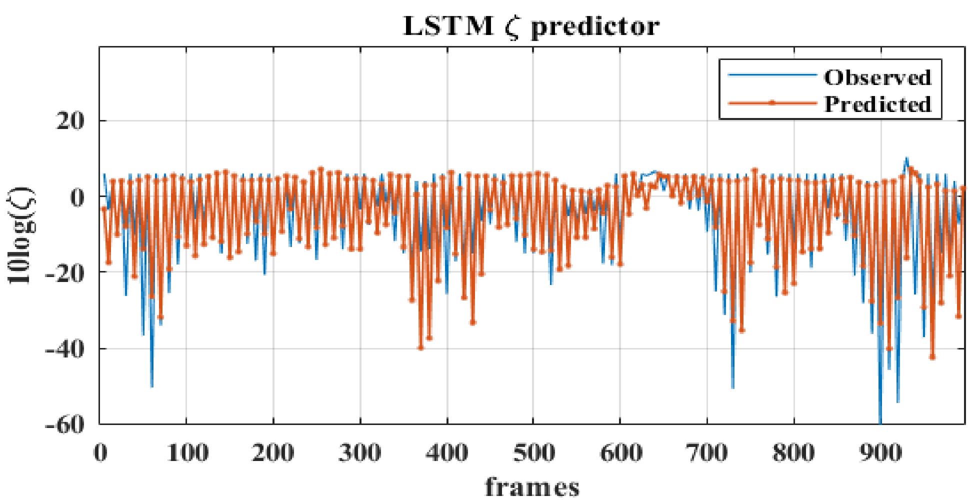

Fig. 4.5. Long Short-Term Memory Predictor Response.

The predictor's AMC performance is determined by computing the effective data rate of each system. The data rate is calculated as $R_b = T_x(BER_t)\Sigma(R)(FM)\, log_2 M$. Where M is the modulation order of the MCS, R is the code rate, FM is the number of frames that activate MCS, and (BER_t) is the BER threshold for AMC. The data rates for the AMC predictors are shown in the table below.

We concluded by considering the RMSE for each predictor since both predictors were trained using the same generated data. The RMSE for the perfect predictor is 0, whereas the LSTM-NN and NN predictors have an RMSE value of 3.9157 and 5.6121, respectively. We can conclude that the perfect predictor is more accurate, followed by the LSTM-NN predictor and the conventional NN predictor.

CHAPTER 5

CONCLUSION AND RECOMMENDATIONS

5.1. Conclusion

This report presented a machine learning-based adaptive codec design and implementation; the system is simulated under a Rayleigh fading channel. The designed codec used 8QAM and 16QAM with a 1/3 BCH encoder and 4QAM without FEC. MIMO is combined into the design in order to increase spectral efficiency. Three different predictors are implemented: a conventional neural network (NN) predictor, long short-term memory neural network (LSTM-NN) predictor, and a perfect predictor as a standard prediction model assumed under perfect channel conditions. Those predictors have RMSE of 5.6121, 3.9157, and 0 for NN, LSTM-NN, and the perfect predictor. The project has been deemed a success as it meets the specified design requirements. And the AMC design can be further improved by using more robust encoding and modulation techniques. The codec can service users operating within 2.4GHz and reduce the computation complexity. The AMC codec can improve by using more robust codecs such as turbo and LDPC codes.

5.2. Future Recommendations

The current AMC is only simulated with fewer parameters; we use a single type of data source, a fixed MIMO structure, and the same FEC encoder. The AMC can improve by making use of more robust FEC codes such as Turbo and LDPC. Further experimentation is needed to optimize the machine learning parameters and explore several neural network architectures. The simulation made that could be time-dependent in order to get more representative performance measurements. These recommendations can be used to design a codec with a higher data rate which would be a sustainable design solution compared to the current one.

APPENDIXS

The plot below is used to determine the 3QAM modulating levels for the MCSs. Performances drops as the modulation order is increased. The performance comparison is done in AWGN but we expect similar result in Rayleigh.

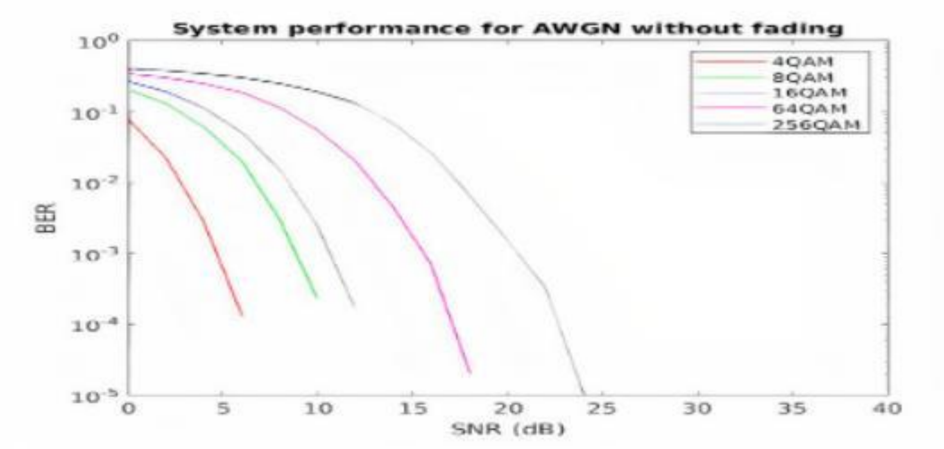

Figure. A.1. QAM SNR vs BER curves in AWGN [14].

Below is the SNR vs BER curves for the chosen MCS at high correlation. We see that high correlation means the channel is close to AWGN like shown in figure A.2.

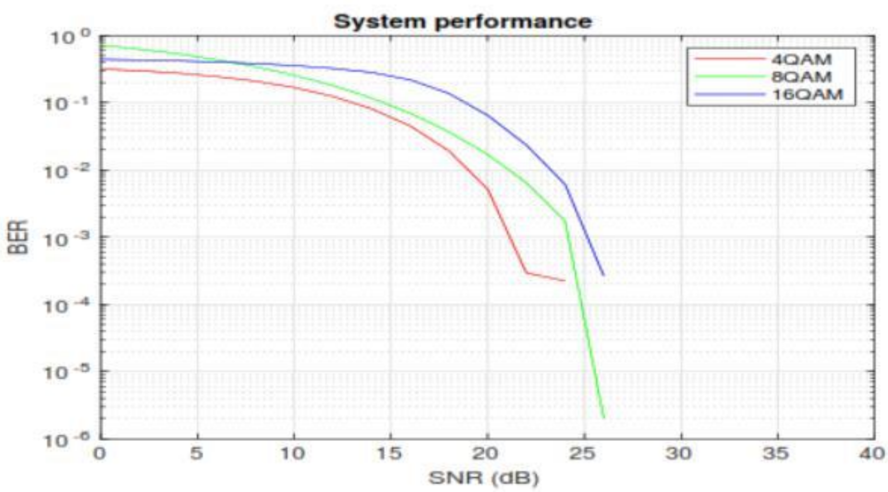

Figure. A. 2. High correlation at $\alpha = 0.99$ and $\sigma_p^2 = 3$, SNR vs BER.

The MCS experiences extreme fading when the channel has low correlation. We see that the curves are linear meaning they have a constant BER drop.

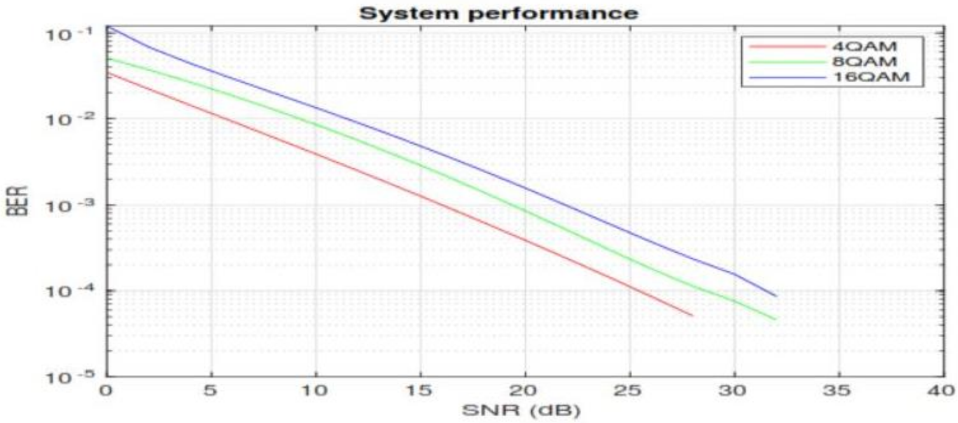

Figure. A. 3. Low correlation at $\alpha = 0.1$ and $\sigma_p^2 = 3$, SNR vs BER

Figure. B.1. is the corresponding ξ plot. The plot is not be perfectly linear due to $\boldsymbol{G_f}$ still being present. It is even more clear to observe the high correlation of the channel in this plot.

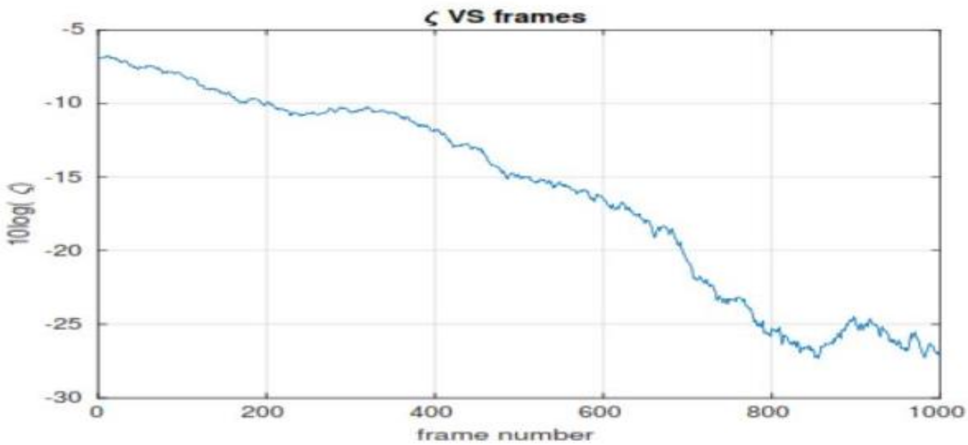

Figure. B.1. High correlation $\alpha = 0.99$ and $\sigma_p^2 = 3, \xi$ plot.

In contrast this figure shows a high channel variation as a result of low channel correlation.

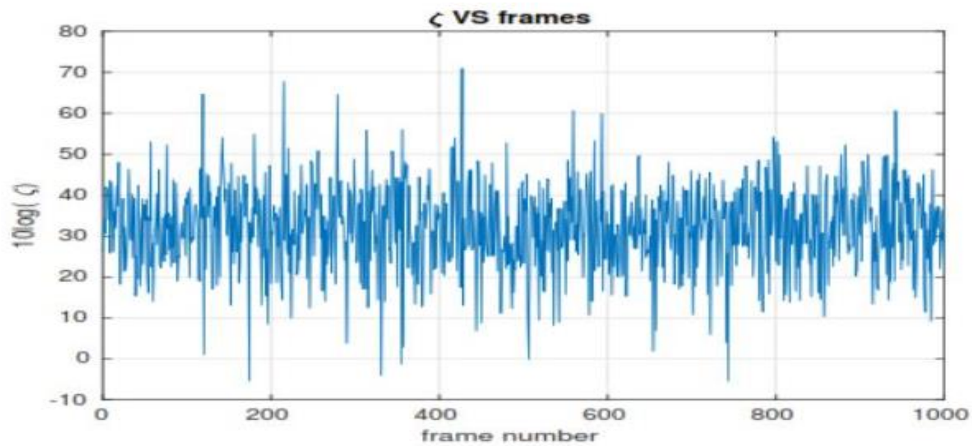

Figure. B.2. Low correlation $\alpha = 1$ and $\sigma_p^2 = 3, \xi$ plot.

REFERENCES

[1] L. Zhang and Z. Wu, “Machine Learning–Based Adaptive Modulation and Coding Design,” *Mach. Learn. Futur. Wirel. Commun.*, pp. 157–180, 2020, doi: 10.1002/9781119562306.ch9.

[2] J. Chen, L. Zhang, Y. C. Liang, X. Kang, and R. Zhang, “Resource allocation for wireless-powered iot networks with short packet communication,” *IEEE Trans. Wirel. Commun.*, vol. 18, no. 2, pp. 1447–1461, 2019, doi: 10.1109/TWC.2019.2893335.

[3] I. Adhicandra, R. G. Garroppo, and S. Giordano, “Optimizing system capacity and application delays in WiMAX networks,” *Proc. 2009 6th Int. Symp. Wirel. Commun. Syst. ISWCS’09*, no. January 2009, pp. 540–544, 2009, doi: 10.1109/ISWCS.2009.5285247.

[4] E. Toader, B. N. Firtescu, A. Roman, and S. G. Anton, “Impact of information and communication technology infrastructure on economic growth: An empirical assessment for the EU countries,” *Sustain.*, vol. 10, no. 10, pp. 1–22, 2018, doi: 10.3390/su10103750.

[5] C. Panem, V. Gad, and R. S. Gad, “Polynomials in Error Detection and Correction in Data Communication System,” *Coding Theory*, pp. 1–22, 2020, doi: 10.5772/intechopen.86160.

[6] K. Rundstedt, “Measurements and Channel Modelling of Microwave Line-of-Sight MIMO Links,” *MS thesis, Chalmers Univ. Technol.*, 2015.

[7] Z. Luo, H. Gao, Y. Liu, and J. Gao, “Capacity limits of time-varying MIMO channels,” *IEEE Int. Conf. Commun.*, vol. 2, no. 5, pp. 795–799, 2005, doi: 10.1109/icc.2005.1494461.

[8] S. Ahmadi, *New Radio Access Physical Layer Aspects (Part 2)*, no. Part 2. 2019. doi: 10.1016/b978-0-08-102267-2.00021-x.

[9] F. C. Vilar, “Implementation Of Zero Forcing And Mmse Equalization Techniques In OFDM,” *Univ. Fortaleza*, 2014.

[10] G. Malik and A. Singh, “Adaptive Equalization Algorithms: An Overview,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 2, no. 3, pp. 62–67, 2011, doi: 10.14569/ijacsa.2011.020311.

[11] R. F. Liao, H. Wen, J. Wu, H. Song, F. Pan, and L. Dong, “The Rayleigh Fading Channel Prediction via Deep Learning,” *Wirel. Commun. Mob. Comput.*, vol. 2018, 2018, doi: 10.1155/2018/6497340.

[12] K. M. S. Soyjaudah and B. Rajkumarsingh, “Adaptive coding and modulation using Reed Solomon codes for Rayleigh fading channels,” *EUROCON 2001 - Int. Conf. Trends Commun. Proc.*, pp. 50–53, 2001, doi: 10.1109/EURCON.2001.937761.

[13] F. Peng, J. Zhang, and W. E. Ryan, “Adaptive modulation and coding for IEEE 802.11n,” *IEEE Wirel. Commun. Netw. Conf. WCNC*, no. 1, pp. 657–662, 2007, doi: 10.1109/WCNC.2007.126.

[14] S. Hong and S. Kim, “Machine Learning – Based Nonlinear MIMO Detector 10 . 2 A Multihop MIMO Channel Model,” no. Mld, pp. 181–195, 2020.

[15] D. L. Goeckel, “Adaptive coded modulation for transmission over fading channels,” *Adapt. Cross Layer Des. Wirel. Networks*, vol. 2, no. 5, pp. 71–94, 2018, doi: 10.1201/9781420046021.ch3.

[16] P. Patil, M. Patil, S. Itraj, and U. Bombale, “IEEE 802.11n: Joint modulation-coding and guard interval adaptation scheme for throughput enhancement,” *Int. J. Commun. Syst.*, vol. 33, no. 8, pp. 3–5, 2020, doi: 10.1002/dac.4347.

[17] B. Zhu, J. Wang, L. He, and J. Song, “Joint Transceiver Optimization for Wireless Communication PHY Using Neural Network,” *IEEE J. Sel. Areas Commun.*, vol. 37, no. 6, pp. 1364–1373, 2019, doi: 10.1109/JSAC.2019.2904361.

[18] P. Yang, Y. Xiao, M. Xiao, Y. L. Guan, S. Li, and W. Xiang, “Adaptive Spatial Modulation MIMO Based on Machine Learning,” *IEEE J. Sel. Areas Commun.*, vol. 37, no. 9, pp. 2117–2131, 2019, doi: 10.1109/JSAC.2019.2929404.

[19] I. Technology, “Machine Learning Based Modulation and Coding Scheme Selection,” no. June, 2019.

[20] W. Su, J. Lin, K. Chen, L. Xiao, and C. En, “Reinforcement Learning-Based Adaptive Modulation and Coding for Efficient Underwater Communications,” *IEEE Access*, vol. 7, pp. 67539–67550, 2019, doi: 10.1109/ACCESS.2019.2918506.

[21] R. C. Daniels and J. Andrews, “Machine Learning for Link Adaptation in Wireless Networks,” *LiLi - Zeitschrift für Lit. und Linguist.*, 2010.

[22] M. Stephen, *Machine Learning An Algorithmic Perspective Second Edition*. 2014. [Online]. Available: https://b-ok.cc/book/2543746/ef80cb

[23] R. C. Daniels, C. M. Caramanis, and R. W. Heath, “Adaptation in convolutionally coded MIMO-OFDM wireless systems through supervised learning and SNR ordering,” *IEEE Trans. Veh. Technol.*, vol. 59, no. 1, pp. 114–126, 2010, doi: 10.1109/TVT.2009.2029693.

[24] F. R. Lone, A. Puri, and S. Kumar, “Performance Comparison of Reed Solomon Code and BCH Code over Rayleigh Fading Channel Faisal Rasheed Lone,” vol. 71, no. 20, pp. 23–26, 2013, doi: 10.5120/12603-9397.

[25] J. Brownlee, “Long Short-Term Memory Networks With Python,” *Mach. Learn. Mastery With Python*, vol. 1, no. 1, p. 228, 2017.

[26] F. A. Gers and F. Cummins, “A critique of neoclassical macroeconomics,” *Choice Rev. Online*, vol. 27, no. 09, pp. 27-5238-27–5238, 1990, doi: 10.5860/choice.27-5238.

[27] D. N. T. How, K. S. M. Sahari, H. Yuhuang, and L. C. Kiong, “Multiple sequence behavior recognition on humanoid robot using long short-term memory (LSTM),” *2014 IEEE Int. Symp. Robot. Manuf. Autom. IEEE-ROMA2014*, pp. 109–114, 2015, doi: 10.1109/ROMA.2014.7295871.

[28] C. Wang, S. Fu, Z. Xiao, M. Tang, and D. Liu, “Long Short-Term Memory Neural Network (LSTM-NN) Enabled Accurate Optical Signal-To-Noise Ratio (OSNR)

Monitoring," *J. Light. Technol.*, vol. 37, no. 16, pp. 4140–4146, 2019, doi: 10.1109/JLT.2019.2904263.

[29] M. L. Pall, " Wi-Fi is an Important threat to Human Helth", Enivironmental Research, vol, 164, pp. 405-415, 2018. [Online]. available: http://www.sciencedirect.com/science/article/pii/S0013935118399355.

[30] T. S. Raad and M. Yakan, " Implementation of a Turbo Codes test bed in the simulink environment," Proc. -8th Int. Symp. Signal Process. its Appl. ISSPA 2005, vol. 2, pp. 847-850,2005, doi: 10.1109/ISSPA.2005.1581071.

LIST OF PUBLICATIONS OF THE CANDIDATE'S WORK

[1] Mogahed A. M. Sharafuden, Vinay Kumar, Mohammed E. Y. Abdalla, "**Performance Analysis of Adaptive Modulation and Codec Selection Based Long-Short Term Memory Neural Network (LSTM-NN)**", IOSR Journal of Electronics and Communication Engineering (IOSR-JECE) e-ISSN: 2278-2834, p-ISSN: 2278-8735.Volume 17, Issue 2, Ser. I (Mar. – Apr. 2022), PP 34-42.

turnitin

Similarity Report ID: oid:27535:17907266

PAPER NAME

Chapters (2).docx

WORD COUNT	CHARACTER COUNT
13644 Words	**74887 Characters**
PAGE COUNT	FILE SIZE
62 Pages	**1.0MB**
SUBMISSION DATE	REPORT DATE
May 30, 2022 6:42 PM GMT+5:30	**May 30, 2022 6:43 PM GMT+5:30**

13% Overall Similarity

Vijay Kumar
31/05/22

The combined total of all matches, including overlapping sources, for each database.

- 7% Internet database
- Crossref database
- 9% Submitted Works database
- 5% Publications database
- Crossref Posted Content database

Excluded from Similarity Report

- Manually excluded text blocks