Performance Evaluation of Machine Learning Based Channel Equalization Techniques

NOSHABA TARIQ

01-242182-007

A thesis submitted in fulfillment of the requirements for the award of the degree of Master of Science (Computer Engineering)

Department of Computer Engineering

BAHRIA UNIVERSIY ISLAMABAD

NOVEMBER 2020

APPROVAL FOR EXAMINATION

Scholar's Name: **Noshaba Tariq,** Registration No. **6382,**

Programme of Study: **MS. (Computer Engineering),**

Thesis Title: **Performance Evaluation of Machine Learning Based Channel Equalization Techniques.**

It is to certify that the above scholar's thesis has been completed to my satisfaction and, to my belief, its standard is appropriate for submission for examination. I have also conducted plagiarism test of this thesis using HEC prescribed software and found similarity index **8**% that is within the permissible limit set by the HEC for the MS degree thesis. I have also found the thesis in a format recognized by the BU for the MS thesis.

Principal Supervisor's Signature: ________________________________

Date: ________________________

Name: **Dr. Shahzad Hassan**

DECLARATION

I, **Noshaba Tariq** hereby state that my MS. thesis titled "**Performance Evaluation of Machine Learning Based Channel Equalization Techniques**" is my own work and has not been submitted previously by me for taking any degree from **Bahria University** or anywhere else in the country/world.

At any time if my statement is found to be incorrect even after my graduation, the University has the right to withdraw/cancel my MS degree.

Name of scholar: **Noshaba Tariq**

Date: **August 31, 2020**

PLAGIARISM UNDERTAKING

I, solemnly declare that research work presented in the thesis titled "**Performance Evaluation of Machine Learning Based Channel Equalization Techniques**" is solely my research work with no significant contribution from any other person. Small contribution/help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero tolerance policy of the HEC and Bahria University towards plagiarism. Therefore, I as an Author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly referred/cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of MS degree, the university reserves the right to withdraw/revoke my MS degree and that HEC and the University has the right to publish my name on the HEC/University website on which names of scholars are placed who submitted plagiarized thesis.

Scholar/Author's Sign: __________

Name of the Scholar: **Noshaba Tariq**

ACKNOWLEDGEMENTS

All praises to Almighty Allah who bestowed me with everything and always helped me at every step of my life.

I would like to express my sincere gratitude towards my supervisor, Dr. Shahzad Hassan for giving me an opportunity to work on this challenging topic.

I am heartily thankful to my senior colleague Dr. Muhammad Asad, whose encouragement, guidance and support from the initial to the final level enabled me to develop an understanding of the subject. He has provided continuous feedback during the course of this work. This accomplishment would not have been possible without his support. I really appreciate his valuable comments on this thesis because it helps a lot to improve my thesis writing.

I am also thankful to my friends, well wishers and closed ones who helped me during this work directly or indirectly. I am grateful to my husband, Farhan Malik, for his support during this study course.

Lastly, I would like to dedicate this work to my daughters, Anaya Malik and Anabia Malik, who are so supportive throughout this journey.

ABSTRACT

Wireless communication systems are evolving with the passage of time and resulted in the advanced communication systems such as mobile and Adhoc communication systems, sensor networks etc. Two fundamental parameters of their performance are higher data rates and better spectral efficiency. In order to achieve high data rate and robust communication over wireless the most important task to be performed at the receiver side is the channel equalization. The transmitted data symbols when pass through the wireless channel suffer various types of impairments, such as fading, Doppler shifts and Inter Symbol Interference (ISI), degrade the overall performance of the communication system. In order to mitigate the channel related impairments many channel equalization algorithms have been proposed in the communication systems domain. These algorithms are based on either the least squares methods or minimum mean square estimation methods. Frequency domain equalization methods are also used for this purpose. The channel equalization problem can also be solved as a classification problem using Machine Learning (ML) methods. Many researchers have addressed this problem however there is a need to compare the performance of the ML based equalizers by using the already established communication systems criterion. In this thesis the channel equalization has been performed using ML techniques and their Bit Error Rate (BER) performance has been compared. Radial Basis Functions (RBF), Multi Layer Perceptrons (MLP), Support Vector Machines (SVM), and Polynomial based NN's have been used for channel equalization. The work is also extended to the multi-carrier systems where the channel equalization of OFDM system is carried out using Long Short Term Memory (LSTM) method. The simulation results show improved BER when compared with the LMS based traditional channel equalization method. Further the computational complexity of all the used ML algorithms has also been formulated theoretically and has been verified with the help of simulations.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ANN	Artificial Neural Network
RNN	Recurrent Neural Network
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
MLP	Multi Layer Perceptron
RBFNN	Radial Basis Function Neural Network
FLANN	Functional Link ANN
LSTM	Long Short Term Memory
SVM	Support Vector Machine
FFT	Fast Fourier Transform
IFFT	Inverse Fast Fourier Transform
ISI	Inter Symbol Interference
LMS	Least Mean Square
LS	Least Square
RLS	Recursive Least Square
MLSE	Maximum Likelihood Sequence Estimator
MMSE	Minimum Mean Square Error
MSE	Mean Square Error
NLMS	Normalized Least Mean Square

PAPR	Peak Average Power Ratio
SGD	Stochastic Gradient Descent
OFDM	Orthogonal Frequency Division Multiplexing.
FDD	Frequency Division Duplexing
QAM	Quadrature Amplitude Modulation
PSK	Phase Shift Keying
PAM	Pulse Amplitude Modulation
FSK	Frequency Shift Keying
CV	Computer Vision
NLP	Natural Language Processing
A/D	Analog to Digital Conversion
D/A	Digital to Analog Conversion
QPSK	Quadrature Phase Shift Keying
ANN	Artificial Neural Network
NN	Neural Network
ML	Machine Learning

CHAPTER 1

INTRODUCTION

Wireless communication is an evergreen research field. New requirements for communication require newer algorithms for robust performance in the presence of challenging environments. Digital wireless communication systems transmit digital data by modulating the data bits using PAM, QAM or FSK modulations. It is the goal of every digital communication system to transfer correct message signals from transmitter to receiver. The transmitted signal passes through the air gap or space between the receiver and transmitter. The signal as received by the receiving antenna is the accumulation of multiple signals that were transmitted by a single transmitter. The effect of multiple paths results in degradation of the signal in many aspects. This results in the phenomenon known as ISI. Hence, the communication signal at the receiver results in either constructive interference or destructive interference, thus making the communication less reliable. The situation worsens if the channel exhibits delay spread longer than the symbol time of the signal. In order to recover the transmitted data ISI must be compensated at the receiver. This process of reducing channel induced distortion is called channel equalization.

This chapter is organized as follows. Section 1.1 describes the objectives of the thesis. Section 1.2 describes the problem statement. Section 1.3 describes the motivation of the work. In the end, section 1.4 presents the thesis layout.

1.1 Objectives

The objectives of this research work are as follows.

- To identify the performance metrics for the existing channel estimation and equalization techniques.
- To identify an improved channel equalization technique for selected wireless channel.
- To critically assess the performance of the various channel equalization techniques by performing simulations.

1.2 Problem Statement

In wireless communication systems the performance may be severely degraded by the wireless channel. The transmitted signal passing through the channel experiences various impairments such as ISI, Doppler shift and fading effects. All these effects tend to limit the throughput of the wireless communication systems. In order to achieve higher data rates, it is mandatory to mitigate the effects of channel induced impairments. This requires designing of an adaptive filter for equalization. The purpose of this filter is to nullify the effects of wireless channel and recover the originally transmitted data, which results in an overall better performance of the communication system. In recent years the use of ML techniques especially the ANN based methods have gained interest due to their remarkable success in the fields of Computer Vision (CV), Speech Recognition and Natural Language Processing (NLP). The technique although invented back in the mid of 20^{th} century was not very popular due to the lack of required computational power. Due to the availability of high speed computational resources and the success of ML in the various other fields have provoked its application in the development of robust communication systems. Many researchers have proposed the use of ML for designing communication systems and have demonstrated improved results in terms of BER. However, its application for this requires the solution to following problems.

- What maximum performance gain in terms of BER that can be achieved by the use of NN and its variants such as MLP, RBF, FLANN, SVM and LSTM.
- Is it possible to train the NN to estimate a wireless channel in real-time as required by the modern day channel equalizers to mitigate the channel in real-time. Typically, an equalizer is required to train its taps in less than few microseconds. What possible methods can be used to achieve this task?

The research aims to address the mentioned problems.

1.3 Motivation for Work

The problems highlighted in the previous sub-section provide an opportunity for advanced level research which is technologically challenging and may extend the knowledge boundaries in many dimensions.

The research focuses on the fundamentals of many engineering and computer science fields such as communication systems, AI, and ML etc. It provides a mechanism to implement new and existing ML methods to solve complex problem of channel equalization in wireless communication systems that is still an active area of research being pursued by the research community.

The merger of mentioned fields of research will open many new research dimensions for the future academic research which will use this work as a foundation.

1.4 Thesis Layout

The thesis is structured as follows:

Chapter 1 describes the research objectives, problem statement and motivation for work.

Chapter 2 provides the elementary concepts of the basic communication system and briefly explains the channel estimation techniques in single-carrier and multi-carrier systems. This chapter briefly introduces the basics of ML techniques and also gives the comprehensive literature review.

Chapter 3 describes the simulations of the NN based channel equalizers. It also discusses the simulation setup which we used to apply NN techniques. The simulation parameters of each technique are discussed. The performance is evaluated using the BER. The detailed simulation result analysis is also done.

Chapter 4 focuses on the channel estimation of OFDM systems. The comparison of single-carrier and multi-carrier communication is also done in this chapter. This chapter also explains the importance of OFDM in simplification of equalization process. This chapter also explains the usefulness of pilot structures in channel estimation. A comprehensive literature review of the ML based channel estimation techniques is also presented. This chapter also verifies the effectiveness of LSTM based channel estimation through simulations.

Chapter 5 summarizes the whole work presented in this thesis and presents some suggestions for future work.

CHAPTER 2

CHANNEL EQUALIZATION USING NEURAL NETWORKS

Digital wireless communication systems are designed to transmit information from source to destination through wireless transmission medium which is termed as channel. The wireless channel causes different types of impairments to the signal passing through it, such as inter symbol interference, which is caused by the multiple signals that combine at the receiver after traversing different paths between transmitter and receiver, and the noise which gets added to the signal also causes severe degradation [1]. The combination of several multiple path signals at the receiver results in either constructive or destructive interference along with the added noise, affect the phase and frequency of the signal, resulting in degradation of the transmitted signal which can be extremely difficult to demodulate at the receiver. In order to mitigate the channel effects the receiver attempts to learn the channel and eliminates the impairments caused by the channel using a method known as channel equalization. Channel equalization is traditionally achieved using adaptive filters such as LMS, NLMS, RLS etc. ANN and ML based techniques provide an alternate more robust and diverse method for channel equalization. The chapter provides a comprehensive overview of the channel equalization techniques used in single-carrier and multi-carrier systems. ANN based equalization techniques are also discussed.

This chapter is structured as follows. We first give an overview of a digital wireless communication system. This is followed by some discussion related performance issues of a wireless communication system. The later sections describe the channel equalization techniques, system model and literature review of NN based equalization techniques. A section on channel equalization for multi-carrier systems is also provided.

2.1 Overview of a Wireless Communication System

A high level block diagram of the communication system is depicted in Figure 2.1. In general, a basic communication system includes three basic blocks; transmitter, receiver, and the wireless channel. The transmitter modulates the input message signal and transmits it using communication channel. During the course of transmission, it experiences various types of impairments such as path loss which results in attenuation of the signal, AWGN and multipath effects caused by the reflections of the electromagnetic waves from various objects present between the receiver and transmitter.

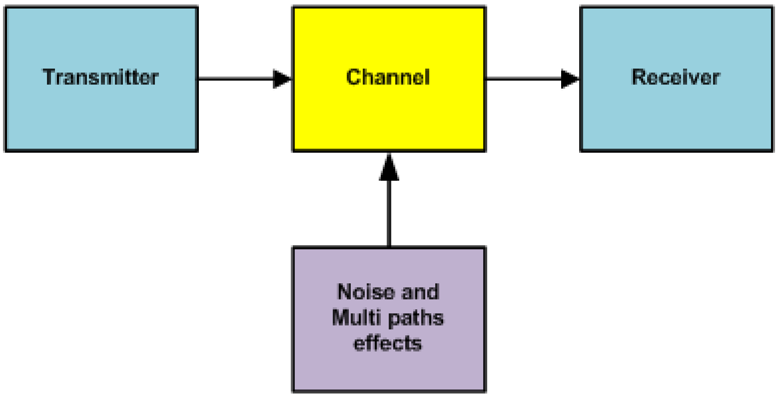

Figure 2.1 Overview of digital wireless communication system [1]

The transmitter and receiver block is subdivided into three blocks as shown in Figure 2.2. The input digital data is fed into the *source encoder* which effectively transforms the bit stream into the compressed form using Huffman encoding. The input can be an audio source, text, binary or any other sensor input, which may require A/D conversion prior to feeding to the source encoder block. The digital data at this stage can also be secured using encryption algorithms. The resulting data sequence at the output of source encoder is passed to the *channel encoder* which adds redundancy in a controlled manner, to help the receiver to detect and correct the channel induced errors. This step should make the data robust against harsh channel conditions. In the next step, output of a channel encoder is given to a *modulator* that applies digital modulation methods such as BPSK, QPSK or some variant of FSK. The output of the modulator is fed to the frequency up-converter which translates the baseband signals to passband frequency and finally the signal is amplified to appropriate levels and then transmitted through antenna.

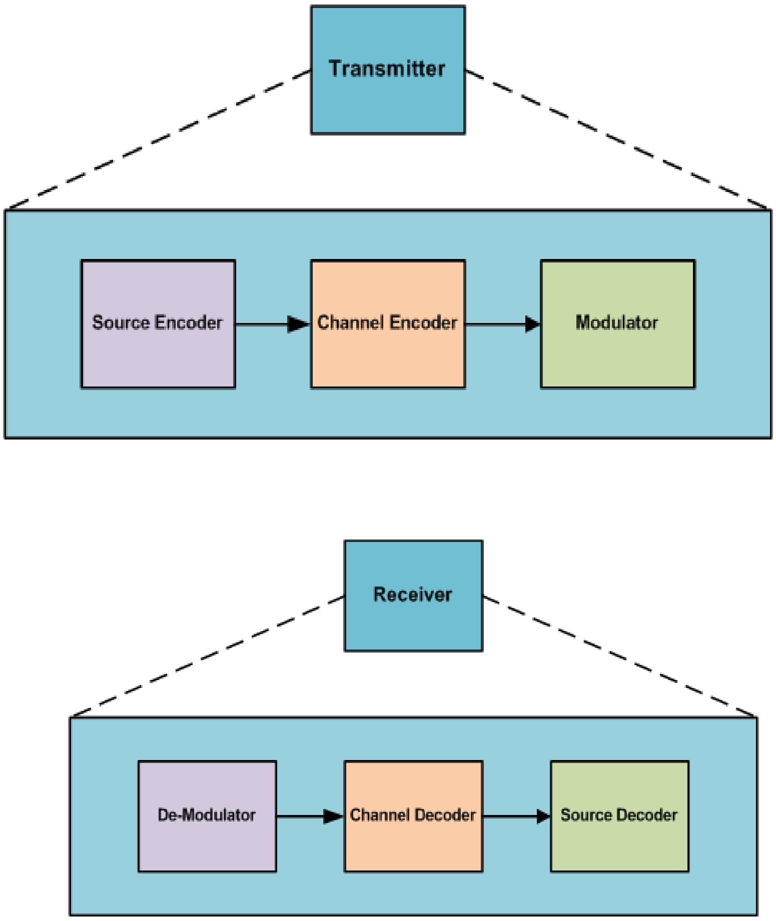

Figure 2.2 Transmitter and receiver block [1]

2.1.1 Mathematical Formulation

This section presents the mathematical formulation [1] of the communication system. Consider $s(t)$ be the transmitted signal. It is represented mathematically as

$$s(t) = Re[x(t)e^{j\omega_c t}] \tag{2.1}$$

Where $x(t)$ is the baseband signal and the $\omega_c = 2\pi f_c$ is the center frequency of the passband signal. The received signal is given as

$$r(t) = \sum_{m=0}^{N-1} \gamma_m(t)s(t-\tau_m) + w(t) \tag{2.2}$$

Where $\gamma_m(t)$ represents the complex amplitude of the channel, τ_m is the delay of the mth multipath and N represents the total number of multi paths. $w(t)$ represents the AWGN. The resulting received signal can be written as

$$r(t) = \sum_{m=0}^{N-1} \gamma_m(t)e^{j\omega_c \tau_m}x(t-\tau_m) + w(t) \tag{2.3}$$

$$r(t) = \int_{-\infty}^{\infty} h(\tau,t)\, x(t-\tau)d\tau + w(t) \tag{2.4}$$

Where $h(\tau,t) = \sum_{m=0}^{N-1} \gamma_m(t)e^{j\omega_c \tau_m}\delta(t-\tau)$ is the impulse response of the time-varying channel. It is the goal of wireless communication systems to estimate the $h(\tau,t)$ which is the channel impulse response for the desired level of performance.

2.1.2 Performance Issues in Wireless Communications

One of the primary goals while designing a communication system is to achieve performance as closer to the Shannon's capacity definition [2].

$$C = B\log_2(1+\gamma) \tag{2.5}$$

Where C is the capacity of wireless channel, B represents the bandwidth and γ represents the SNR. This theorem gives the fundamental bound on the achievable capacity of the wireless channel. All communication systems tend to achieved Shannon capacity. As of today this goal has not been fully achieved due to many reasons. Amongst the most notable reasons are:

1. The Wireless channel
2. Signal to Noise Ratio (SNR)

3. Link budget

Assuming the availability of required Bandwidth these three objectives must be served to achieve desired performance in wireless communication. SNR and link budget can be improved using high gain antennas, more transmit power and by using better antennas, however the effects caused by channel require more sophisticated handling. Its effects must be mitigated using channel equalization methods.

2.1.3 Wireless Channel

An insight into the wireless channel is mandatory for the design of effective wireless communication system. A wireless channel alters the transmitted signal in many ways.

1. Fading
2. Doppler shifts
3. ISI

Fading refers to the time variation of the received signal power.[1]. The variation of the received signal power is caused when either of the transmitter or receiver is moving, or when the channel is changing. Fading can be of two types.

1. The large scale fading (LSF)
2. The small scale fading (SSF)

The LSF results when the effect of fading is experienced over a larger geographical area where the effects of SSF are experienced over a smaller area, comparable to the wavelength of the transmitted signal. This is also termed as Rayleigh fading. The channel may consist of multiple paths.

Another impairment caused by the channel is Doppler shift in the frequency experienced by the moving transmitters and receivers, especially at the microwave frequencies. Typical Doppler shift is in several kilohertz. This results in the severe degradation of the SNR and results in increased BER. In most of the cases the channel related impairments can be reversed using channel equalization techniques, which nearly eliminate the multipath effects. The receiver estimates the effects caused by estimating the channel and then applying the advanced signal processing methods or the ML methods to eliminate them.

2.1.4 Channel Estimation Techniques

Channel estimation techniques are broadly categorized into three main types:

- The Pilot Aided Channel Estimation techniques(PACE)
- Blind and Semi Blind Channel Estimation techniques(BSB)
- Decision Directed Channel Estimation techniques (DDCE)

In the PACE based estimation techniques the transmitter sends a known sequence of data symbols to the receiver called pilot symbols. The receiver estimates the Channel with the help of received pilots using mathematical techniques. Let $X(z)$ represent the transmitted symbols known to both receiver and transmitter. $H_c(z)$ is the channel impulse response and $R(z)$ is the received signal. The received signal in frequency domain can be represented as

$$R(z) = H_c(z)X(z) \tag{2.6}$$

The estimate of the channel $\widehat{H}(z)$ is obtained as

$$\widehat{H}(z) = \frac{R(z)}{X(z)} \tag{2.7}$$

If the equalizer is employed, then the above equation can be written as

$$R(z) = H_c(z)H_e(z)(z)X(z) \tag{2.8}$$

Where $H_e(z) = \frac{1}{\widehat{H}(z)}$ represents the equalizer. The goal is to achieve $H_c(z)H_e(z) = 1$. If this is achieved then $R(z) = X(z)$. This means that the signal is correctly recovered.

The second technique is the BSB channel estimation technique. In this technique the receiver has no information about the input signal of the channel. This technique uses the data symbols for channel estimation by employing the pre-coding of the symbols at the transmitter. The receiver knows the parameters of the pre-coding used at the transmitter and then uses correlation based methods to estimate the channel information [3-5]. The semi-blind technique uses both the pilot symbols and the transmitted data symbols. This slightly increases the bandwidth but the technique is effective.

The DDCE technique uses the pilot symbols and the demodulated symbols for the channel estimation. In the absence of bit errors, the symbols can be used for estimation of the channel impairments and start acting as the pilot symbols. This technique proves to be more efficient as compared to the pilot symbol based channel estimation techniques because it reduces the bandwidth by saving the number of pilots required in pilot based channel estimation techniques.

2.1.5 Channel Equalization

Channel equalization and channel estimation both are interdependent. Inverse of the channel estimate can be used for channel equalization. The performance of equalizer is proportional to the accuracy of the channel estimation.

The equalization mechanism can be divided into two modes- a training mode and a decision-directed mode. In the first mode, the equalizer is trained by sending a training sequence. Training sequence is known apriori to the receiver. Equalizer weights are learned using the training sequence. In the second mode, the equalizer is operated on the channel to estimate the channel. Various types of equalizers are used in the digital communication receiver. Figure 2.3 depicts the classification of the equalizers [6]. Equalization is generally divided into two categories. The linear equalizers and the nonlinear equalizers, the linear equalizers employ only feed forward path and do not use the output of the equalizer in the equalization process. On the other hand, the nonlinear equalizers use the output of the equalizer in the determination of the future samples. Both the linear and nonlinear equalizers employ adaptive algorithms such as LMS, NLMS, RLS and Kalman filtering etc. for the adaptation of the equalizer weights. Amongst the nonlinear equalizers is the Maximum Likelihood Sequence Estimator (MLSE). This type of equalizer does not use filter for equalizing the channel but instead uses Viterbi algorithm to decode the sequence and chooses the sequence with maximum probability as the output.

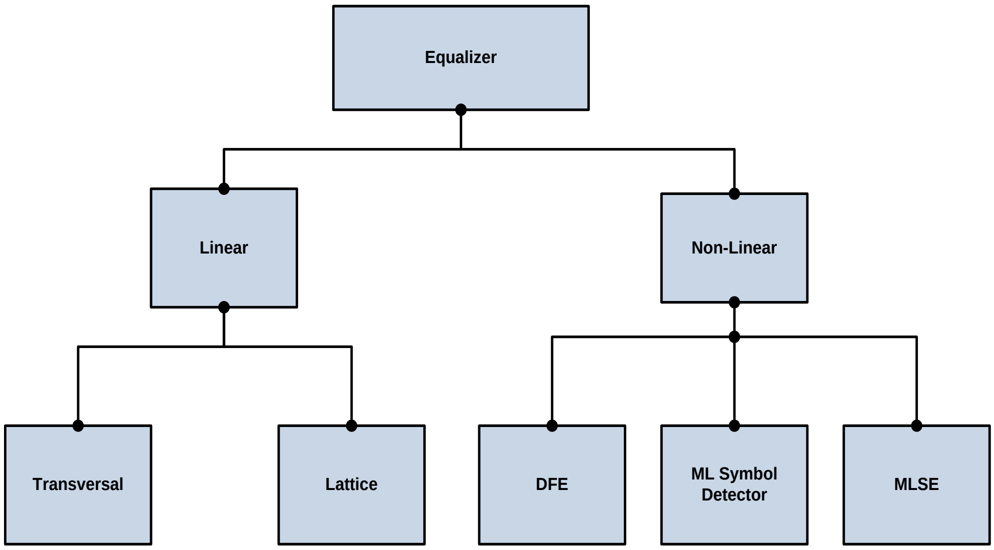

Figure 2.3 The classification of equalization techniques [6]

2.2 Introduction to ML

2.2.1 A Perceptron

ML is a subfield of computer science which focuses on the development of algorithms to learn and solve the complex problems. Unlike traditional approach it does not use predefined models or set of equations to solve the given problem, instead it learns to solve the problem. It consists of human brain like neurons termed as perceptrons. A perceptron is a simple mathematical model (function) that maps the set of inputs to the set of outputs and performs three basic operations: multiplication, summation and activation. Each input value is multiplied with its corresponding weight. The previously weighted inputs are then summed up and passed through the activation function. The activation function determines the output of neuron with respect to its input. The commonly used activation functions are threshold, linear, sigmoid and ReLU.

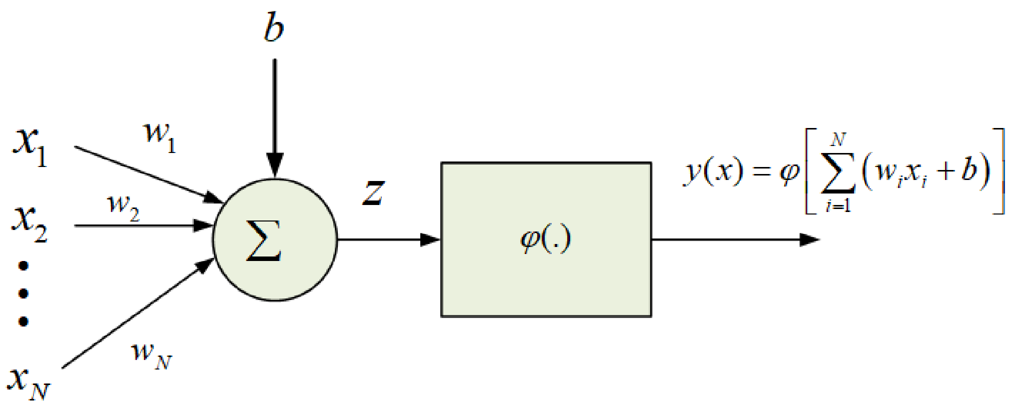

Figure 2.4 A single perceptron [7]

Mathematically, a perceptron can be defined as

$$y(x) = \varphi(\boldsymbol{w}^T x + b) \tag{2.9}$$

Here, $\boldsymbol{w}$ is a weight vector, $\boldsymbol{b}$ is a bias and $\boldsymbol{w^T x}$ is a dot product of w and x.

$$z = \sum_{i=1}^{N} w_i x_i \tag{2.10}$$

So the equation (2.9) becomes

$$y(x) = \varphi\left(\sum_{i=1}^{N} w_i x_i + b\right) \tag{2.11}$$

$\boldsymbol{\varphi}(.)$ is the activation function. A sigmoid and ReLU functions are defined in equation (2.12) and (2.13).

$$\varphi(x) = \frac{1}{1 + e^{-x}} \tag{2.12}$$

$$\varphi(x) = \max(0, x) \tag{2.13}$$

2.2.2 Neural Network (NN)

A single perceptron cannot perform complex nonlinear mappings. So, many perceptrons are linked together to make larger structures. For this, a layered network is designed where each layer consists of multiple perceptrons. This arrangement of perceptrons is termed as Neural Networks (NNs). It has at least three layers as shown in Figure 2.5.

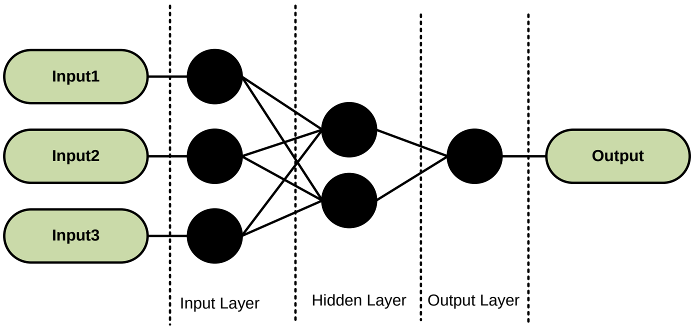

Figure 2.5 A basic structure of ANN [7]

These layers learn their respective weights through a feedback process called back-propagation [7]. The function of this algorithm is to modify the weights of the network in an orderly fashion to attain a desired design objective. This is done by comparing the output of network with the desired result, and use the difference between the outputs to adjust the weights of the connections in the network. The technique was known since early 1950's however it was not popular due to that lack of computational resources. In the last decade with the advancement in the CPU and GPU technologies, the use of NN has emerged and has revolutionized many fields especially the Computer Vision and AI, and has become a useful tool for many complex applications including, nonlinear system identification, pattern recognition, adaptive channel equalization and optimization [8]. It is also gaining attention in the area of wireless communications. Various NN based algorithms are employed that are influencing the legacy design methods used in communication systems development.

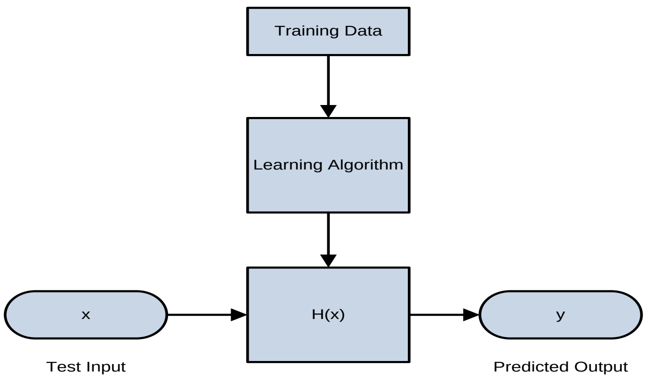

Figure 2.6 A basic flow diagram of supervised learning [7]

2.2.3 Types of Learning

The learning algorithms can be categorized as.

- **Supervised learning:** In this type, some tests are carried out by a human expert and his findings are noted down. These findings are arranged to form a training dataset that maps the inputs to the outputs. Model learns from the training data during a training process and then predicts the unseen data. This concept is depicted in Figure 2.6. Classification and regression are the examples of supervised learning.
- **Unsupervised learning:** In this type, training examples are not used. It is used for data analysis to find the grouping of data. It is more complex than supervised learning. Clustering is an example of this type of learning.

2.2.4 Building Blocks of a Learning Algorithm

Following are the building blocks of the learning algorithm.

- **Loss function:** Error is a difference between the actual output and the predicted output. loss function is used to computes this error. Mean Square Error (MSE) is the most widely used loss function.
- **Optimization algorithm:** The goal of the learning algorithm is to minimize the loss function by updating the weights. The well-known optimization algorithms are stochastic gradient descent (SGD) and Adam.

2.2.5 Model Evaluation

When a training algorithm is applied to the training dataset a model is created. Evaluation of this model is essential for the performance measurement. The commonly used performance metrics for classification problems are

- **Confusion matrix**: It is a two dimensional table that summarizes the usefulness of the model in predicting the class labels. One dimension indicates the actual labels and the other dimension indicates the predicted labels.
- **Accuracy**: It is defined as the ratio of correctly classified examples to the total classified examples.

2.3 Machine Learning Based Channel Equalization Techniques

2.3.1 Related Work

NNs are capable of processing non-linear data and can produce complex decision regions. Therefore, NNs can be employed for equalization purpose to overcome the difficulties associated with channel nonlinearities [8-10]. The performance of NN based equalizer has been reported to be superior to other conventional adaptive equalizers. In recent past the use of NNs has gained popularity in the design of Software Defined Radios where DNN, CNN and RNN have been applied for many of the classical radio operations [11-14].

In [15] the Deep NNs have been used for the channel estimation of doubly selective channels which experience variations both in time and frequency. The deep learning based algorithm is trained in three steps. These are the *Pre-Training step, Training* and *Testing* stages. During the first two steps the model is developed offline using training data. During the testing stage the channel is estimated and equalized. The results show improved BER performance as compared to Linear MMSE.

In [16] the ML and NN have been used in the Frequency Division Duplexing (FDD) systems which is a double selective channel and the results show improvements in term of MMSE in prediction of the channel.

In [17] the NN and DL methods have been used to predict the behavior of Rayleigh channel and it has been reported through simulations that the MSE performance compared with the traditional algorithms has improved.

In [12, 18] DL has been thoroughly investigated and provides a review of the various ML based techniques for the wireless communication. It has been shown that traditional theories do not meet the higher data rate requirements of communication and limits the efficiency due to complex undefined channel requirements, fast processing and limited block structure. On the other hand, AI based communication systems faces some challenges that needs to be addressed. These challenges includes the availability of large amount of data and how easily it can be integrated in classical infrastructure [19]. Similarly, ML has been applied to the physical layer for modulation recognition and classification [18, 20-23] .

In the later sections, we will briefly explain the fundamental concepts of different types of NNs along with the related work in literature.

2.3.2 Multi Layer Perceptron (MLP)

An MLP is feedforward NN which consists of an input layer, a hidden layer and an output layer. It has nonlinear decision making capabilities. The training of MLP is done through the back-propagation algorithm [24]. The MLP was the first neural network used for channel equalization [9, 10, 25-28]. Its structure is depicted Figure 2.7. Gibson et al. [10] introduced a MLP based non-linear equalizer structure and demonstrated its superior performance over the linear equalizer (LMS). The major drawback of the MLP network is its slow convergence [29]. This is due to the back-propagation algorithm which operates on the basis of first-order information. Genetic algorithm [30] can be used to solve this problem. The convergence can be improved by using the second-order data like the Hessian matrix, which is defined as the second-order partial derivatives of the error performance.

Zerguine [31] has proposed a MLP-based DF equalizer with lattice filter to overcome the convergence problem to improves the performance of MLP. However, this improvement increased the complexity of MLP structure.

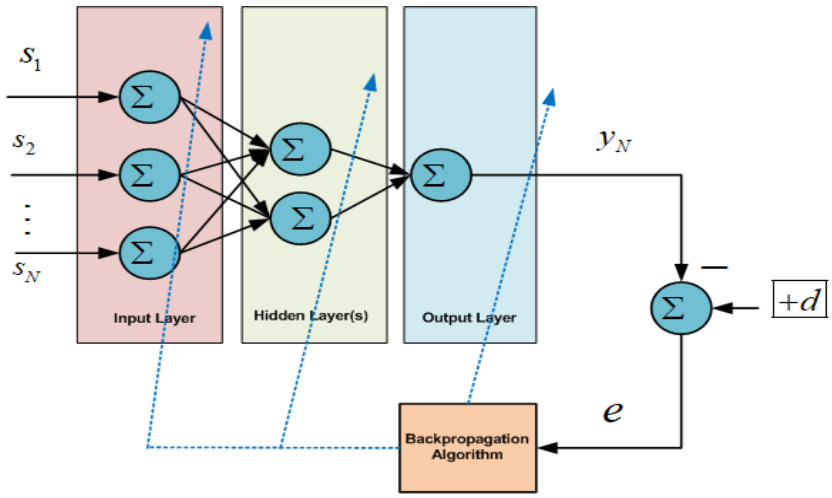

Figure 2.7 The MLP based adaptive equalizer [32]

2.3.3 Radial Basis Function Neural Network (RBFNN)

The RBFNN is a three-layer network which comprises an input layer, a nonlinear hidden layer and a linear output layer. The input layer contains the source symbols. In the hidden layer, the input space is transformed into a high dimensional space by using non-linear basis functions. The output layer linearly combines the output of the previous layers. The structure of RBFNN is depicted in Figure 2.8.

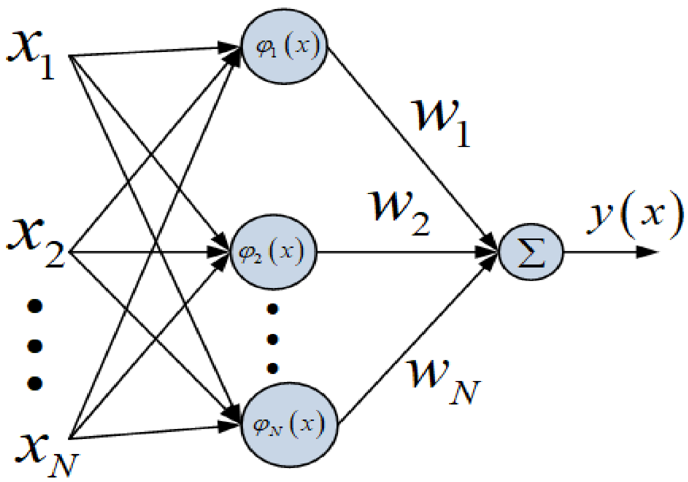

Figure 2.8 The structure of RBFNN [33]

Mathematically, the output of RBFNN can be derived using Gaussian radial function in hidden layer. A typical Gaussian radial function is expressed as follows.

$$\varphi(x) = \exp\left[-\frac{(x-c)^2}{r^2}\right] \tag{2.14}$$

A linear model is expressed as follows.

$$y(x) = \sum_{i=1}^{N} w_i\, \varphi_i\,(x) \tag{2.15}$$

By substituting equation (2.14) in (2.15), we get

$$y(x) = \sum_{i=1}^{N} w_i \exp\left(-\frac{(x-\ c_i)^2}{r^2}\right) \tag{2.16}$$

Where $\boldsymbol{x}$ is an input vector at input layer. Hidden layer maps this input vector to an output scalar value by using equation (2.16). N indicates the number of neurons in hidden layer. $\boldsymbol{c_i}$ is a center vector at i^{th} neuron and w_i is the weight of the i^{th} neuron of the output layer.

RBFNN provides an appealing alternative to MLP for channel equalization. Many techniques have been developed to solve the equalization problem using RBF [34-36]. In 1991, Chen et.al [9] used RBFNN for equalization. Similarly, a RBF based equalizer has been reported which shows the satisfactory performance [37, 38]. Another work has demonstrated the use of RBFNN for equalization and found the improvement in BER [39]. The performance of RBFNN is compared with maximum likelihood sequence estimator (MLSE) over the Rayleigh fading channel [37, 40, 41]. Simulations have confirmed that RBFNN is a reasonable choice with low computational complexity. Hen et al. [42] and Cha and Kassam [43] have proposed a complex RBF (CRBF) network and improved performance has been observed.

The drawback of RBFNN is that it is not suitable for hardware implementation. The network needs a large number of hidden nodes to achieve a desired performance.

2.3.4 Functional Link Artificial Neural Network (FLANN)

In the last few years, FLANN has become very famous [44]. It is a single-layer NN that can form complex decision boundaries. FLANN provides less computational complexity and greater convergence speed than other traditional NNs. From the perspective of hardware implementation, FLANN has simple design, less computational complexity, and higher computation performance [45, 46].

The input dimension is expanded by using nonlinear functions which may lead to better nonlinear approximation. The expansion is done using three commonly used functions i.e. trigonometric, Chebyshev expansion and Legendre expansion. A traditional FLANN uses trigonometric functions, whereas other two expansions are based on Legendre [33, 47] and Chebyshev polynomials [48]. Their plots are shown in Figure 2.9 and Figure 2.10 respectively. The mathematical formulation of the FLANN can be given as follows.

$$DF(X) = \omega^T X^* \tag{2.17}$$

Where $DF(X)$ is the decision function or the output of the FLANN, ω^T is the weight vector and X^* is the N dimensional vector of the functional each represented as $f_i(X)$.

$$X^* = \begin{bmatrix} x_1 \cos(\pi x_1) \sin(\pi x_1) \ldots \cos(2\pi x_1) \sin(2\pi x_1) \\ x_2 \cos(\pi x_2) \sin(\pi x_2) \ldots \cos(2\pi x_2) \sin(2\pi x_2) \\ x_1 x_2 \end{bmatrix} \tag{2.18}$$

The output $y(n)$ is given as

$$y(n) = \phi(\omega^T X^*) \tag{2.19}$$

Where $\phi(.)$ represents the activation function. In this work $\phi(z) = \tanh(z)$ is used. The error is computed as

$$X' = e(n) = d(n) - y(n) \tag{2.20}$$

The weights are adopted using LMS algorithm. As given by equation (2.17) to (2.20) , the functional expansion is performed using orthogonal trigonometric functions. After the initial training the weights calculation is stopped and the equalization is carried out in decision directed mode.

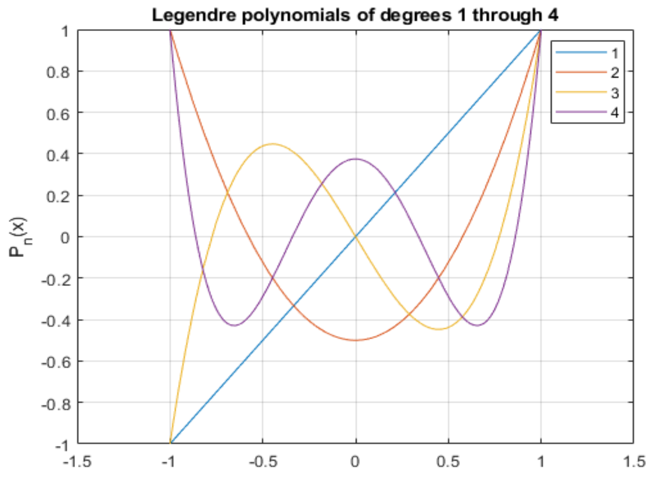

Figure 2.9 Legendre polynomials

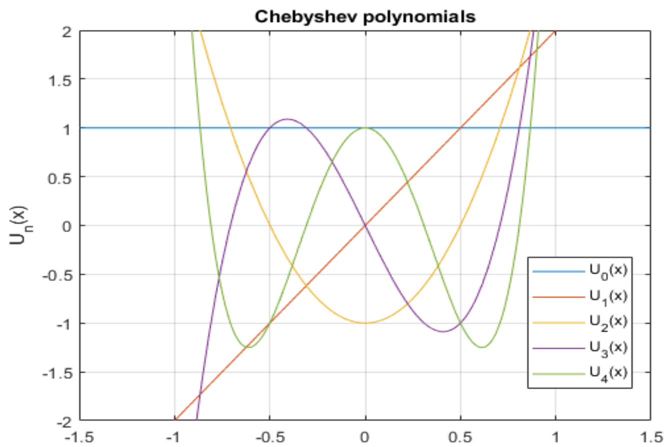

Figure 2.10 Chebyshev polynomials

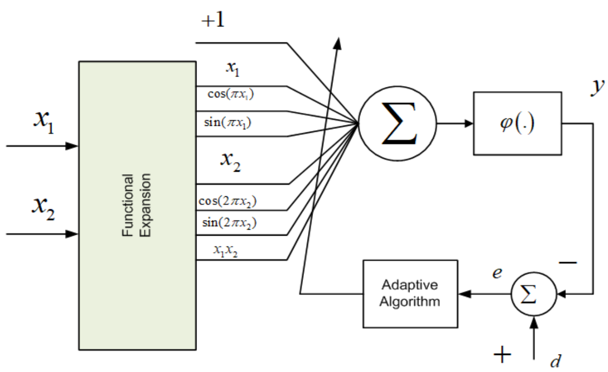

Figure 2.11 FLANN structure [32]

2.3.5 Chebyshev FLANN (Ch-FLANN)

Ch-FLANN is another computational efficient network. It is depicted in Figure 2.12. It has many applications in functional approximation [49], nonlinear dynamic system identification [50, 51] and nonlinear channel equalization [52]. In these networks the expansion is performed using Chebyshev polynomials. These polynomials are generated using the equation (2.21).

$$T_{n+1}(x) = 2xT_n(x) + T_{n-1}(x) \tag{2.21}$$

The first few polynomials are defined in equation (2.22).

$$\begin{aligned} T_0(x) &= 1 \\ T_1(x) &= x \\ T_2(x) &= 2x^2 - 1 \\ T_3(x) &= 4x^3 - 3x \\ T_4(x) &= 8x^4 - 8x^2 + 1 \\ T_5(x) &= 16x^5 - 20x^3 + 5x \end{aligned} \tag{2.22}$$

Using the above polynomials, the input pattern can be enhanced as:

$$X' = [1, T_1(x), T_2(x), T_3(x), T_4(x), T_5(x), \dots \dots] \tag{2.23}$$

This enhanced expression is then applied to single layer perceptron.

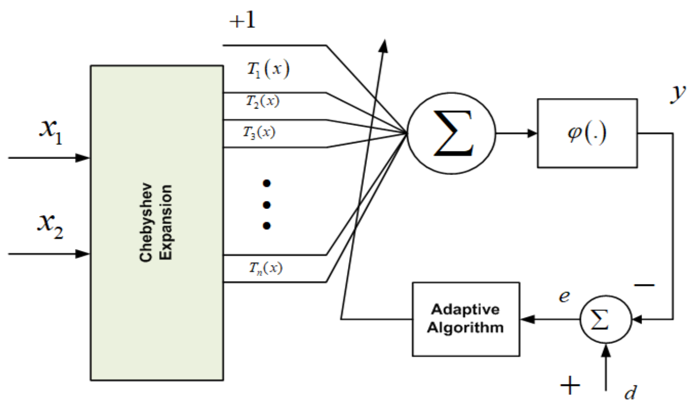

Figure 2.12 Chebyshev FLANN [33]

Ch-FLANN offers comparable efficiency and, in some cases, better than MLP but with much lower computational load.

2.3.6 Legendre FLANN (Le-FLANN)

Le-FLANN is similar to Ch-FLANN. It is depicted in Figure 2.13. It is computationally efficient and gives better performance. The input pattern is enhanced into a nonlinear high dimensional using Legendre polynomials. These polynomials are generated using the equation (2.24).

$$P_{n+1}(x) = \frac{1}{n+1}[(2n+1)xP_n(x) - nP_{n-1}(x)] \tag{2.24}$$

The first few Legendre polynomials are as follows:

$$\begin{aligned} P_0(x) &= 1 \\ P_1(x) &= x \\ P_2(x) &= \frac{1}{2}(3x^2 - 1) \\ P_3(x) &= \frac{1}{2}(5x^3 - 3x) \\ P_4(x) &= \frac{1}{8}(35x^4 - 30x^2 + 3) \end{aligned} \tag{2.25}$$

The input data is reshaped into the high dimensional using these polynomials. The enhanced pattern is then given to single layer RBFNN.

$$X' = [1, P_1(x), P_2(x), P_3(x), P_4(x), P_5(x), \ldots \ldots] \tag{2.26}$$

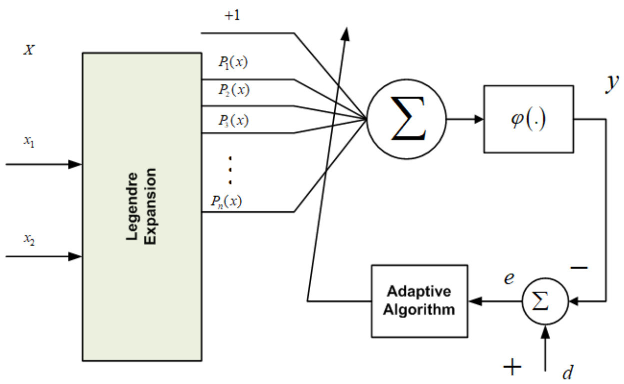

Figure 2.13 Legendre-FLANN [33]

Since there is no hidden layer in both the Ch-FLANN and Le-FLANN networks as opposed to RBF, the computational complexity for these networks is comparatively lower. As opposed to trigonometric FLANN, Ch-FLANN is more efficient in terms of computational complexity. The downside of this network is its increased complexity in an attempt to reduce the BER by increasing its dimensions.

2.3.7 Recurrent Neural Network (RNN)

RNN is a popular DL technique which was first introduced for processing sequential data [24] and gained a lot of attention in recent past. They have been proven better than traditional signal processing methods in modeling and predicting nonlinear and time series [53] in a wide variety of applications ranging from speech processing and adaptive channel equalization [54-58].

Unlike ANN, which does not have memory and cannot deal with temporal data, RNN has feedback loops which make them attractive for the equalization of nonlinear channels. This means data can be fed back to the same layers. Figure 2.14 illustrates this concept. It has been demonstrated through simulations that a reasonable size of RNN can model the inverse of the channel. RNNs are known to outperform FLANN, MLP and RBF [59, 60]. In [61] it has been shown that equalizers based on CNN and RNN reduce the channel's fading effects but also increase the overall coding gain by more than 1.5 dB.

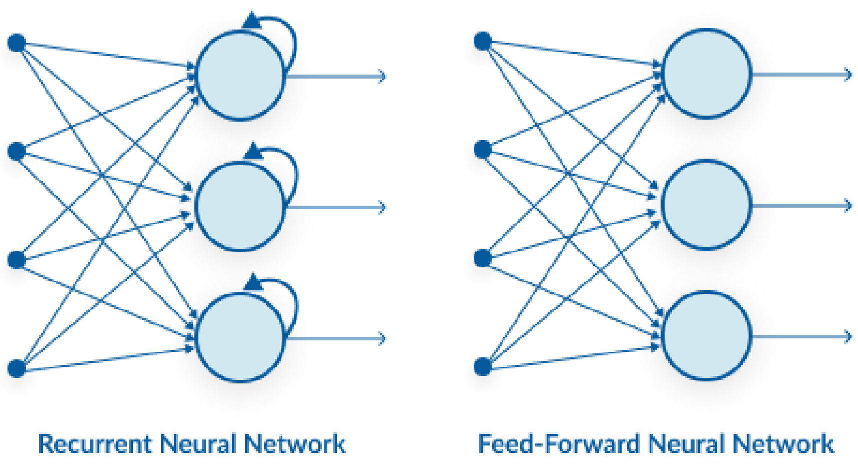

Figure 2.14 RNN vs ANN [24]

RNN has one problem of *exploding and vanishing gradient* [62]. This problem arises when there is a long dependency in a sequence. To solve this problem, LSTM has been proposed [63]. LSTM is slightly different than RNN. It has some special units in addition to standard units. These special units are called memory cells. These units can retain the information for a long period. This means that LSTM can now detect the patterns even in a long sequence. The sequence problems can be efficiently solved by LSTM and can also solve the channel equalization problem. In this case, future samples can be predicted by taking previous symbols into account. This means that variations in a channel can be easily tracked. We can specify the number of samples that LSTM can hold for prediction of future sequences. If it is selected according to the delay spread of a channel, more accurate results may be observed.

2.3.8 Support Vector Machines (SVM)

SVM lies in the category of supervised learning. Originally, it was developed for binary classification. Then it has been extended to perform regression and multi-class classification problems [64-66]. It has the potential to generalize well in classification problems by maximizing the margin. The trained classifier contains support vectors on the margin boundary and summarizes the information required to separate the data. It uses the parametric learning algorithm, in which a model has fixed learnable parameters which are adapted during the training process. Once the model is trained, these parameters are then used exclusively for testing while discarding all the training examples.

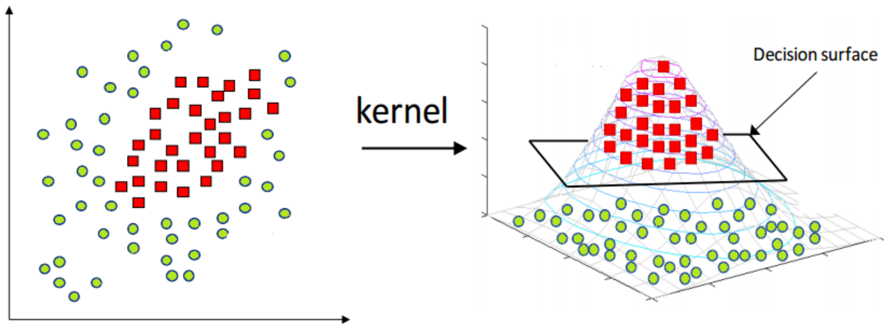

Figure 2.15 Kernel methods in SVM [64]

This makes the SVM more computationally efficient. On the other hand, NNs are non-parametric as the number of parameters increases with the number of layers. NN introduces nonlinearity by using nonlinear activation function whereas SVM uses kernel methods that implicitly transform the input space into higher dimensions. The concept of kernel methods is illustrated in Figure 2.15. RBF kernel is the most commonly used kernel method. The SVM has been suggested to address the number of digital communications issues due to its nonlinear processing capability. A DFE, based on SVM has been proposed and it has been observed that the performance of this equalizer is superior to MMSE DFE [67]. Similar work has been done in [68].

2.3.9 Autoencoder Based Communication System

Autoencoder was first introduced in [69]. It is a feedforward deep NN, based on encoder-decoder architecture as shown in Figure 2.16. The model is divided into two parts. First part is encoder which maps the input data to a latent form or code. The decoder then tries to reproduce the original input data from the latent representation. The code layer is used for dimensionality reduction and de-noising. It has many applications in communication systems as the aim of every

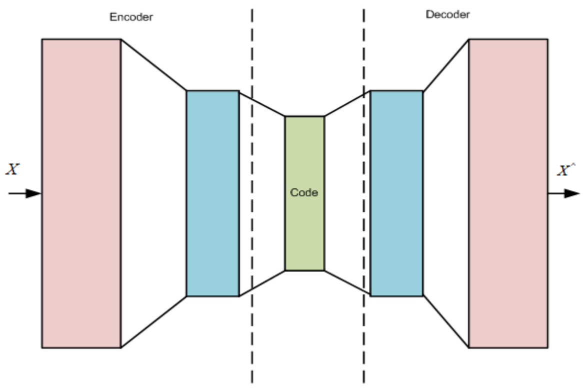

Figure 2.16 Autoencoder [69]

communication system is to receive the same transmitted symbols. Keeping this in view, a communication system has been proposed, based on Autoencoder in [11, 70] which does not use any predefined conventional block based structure. Its design is depicted in Figure 2.17. This approach makes the design simple. It represents the transmitter, channel, and receiver as a single deep neural network.

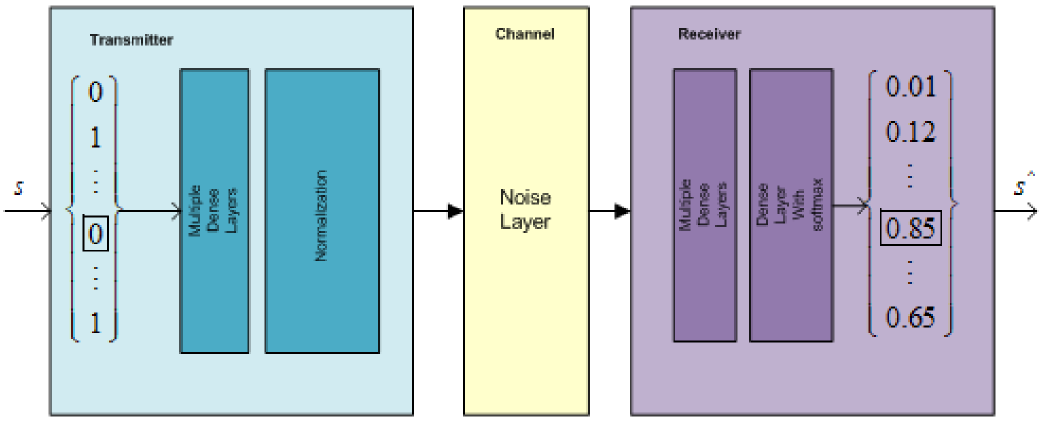

Figure 2.17 Autoencoder based communication system [11]

In this section NN based channel equalization techniques for the single-carrier systems have been critically reviewed. In the next section the multi-carrier OFDM based communication systems have been discussed in detail along with the NN base channel equalization.

2.4 ANN based Channel Equalization in Multi-carrier OFDM Systems

Multi-carrier communication systems as the name shows use multiple sub carriers for the transmission of information [71, 72]. These systems offer robustness against ISI and are more bandwidth efficient as compared to single-carrier system. Channel equalization for multi-carrier systems is based on the similar principles as described in the previous sub-sections. The techniques are based on legacy communication methods already discussed. The use of NN based channel equalization in OFDM is currently an active area of research. The next sub-section provides an overview of the OFDM communication system followed by the application of NN for channel estimation.

2.4.1 Introduction to OFDM

OFDM was introduced in late 60 and with time progressed so rapidly that it has become the potential candidate for future wireless communications [73-80]. It is basically an extension of a traditional technique, Frequency Division Multiplexing (FDM) where sub-carriers were far apart from each other. In OFDM, subcarriers are overlapped and closely spaced. This overlapping is allowed due to orthogonal subcarriers. This overlapping serves the following benefits listed as:

- Requires less bandwidth to carry the same amount of data as depicted in Figure 2.18.
- Makes the spectrum of each subcarrier nearly zero at other subcarrier frequencies.

This means sub-carriers will not cause interference to the neighboring subcarriers. This phenomenon is called Inter Channel Interference (ICI) and it can further be reduced by inserting a Cyclic Prefix (CP), which is a copy of symbols tail placed at the front of the symbol.

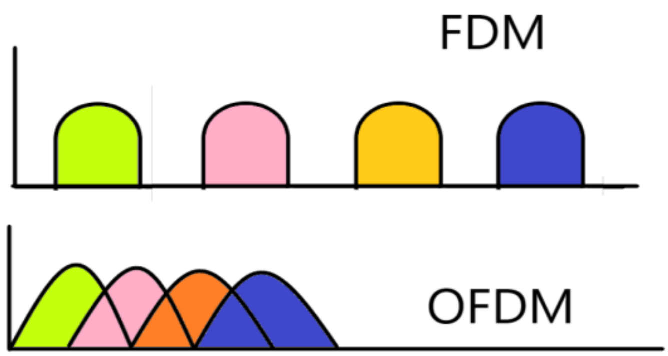

Figure 2.18 FDM vs OFDM

2.4.2 System Model

The overall diagram of OFDM communication system is depicted in the Figure 2.20. The transmitter performs these operations listed as:

- Modulates the input bit sequence using PSK/QAM modulation

- Converts it to parallel bit streams
- Inserts pilots on each stream
- Performs IFFT
- Inserts CP
- Converts back to single serial stream

At receiver, the reverse operation is employed as follows.

- Converts the time domain single bit stream to parallel stream
- Removes CP from each stream
- Performs FFT
- Applies channel estimation techniques on each stream using pilots
- Coverts to serial stream and perform demodulation
- The original transmitted data is then recovered after demodulation

Mathematically the OFDM system can be formulated [1] as follows

$$d_k = \sum_{n=0}^{N-1} D_n e^{\frac{j2\pi kn}{N}} \tag{2.27}$$

Where $D_n = [D_0 D_1 D_2, \dots, D_{N-1}]$ are the complex QAM/QPSK modulated symbols and d_k represents the kth OFDM symbol. In the next step the CP is added to the OFDM symbol d_k this is typically done by copying the last few symbols of d_k to the start of the symbol. This makes the OFDM symbol robust against ISI and also makes the OFDM periodic and help in channel estimation and demodulation. The symbols are then serialized and transmitted which passes through the time-varying channel $h(t, \tau)$.

$$r(t) = \int_0^{\infty} x(t-\tau)h(t,\tau)d\tau + n(t) \tag{2.28}$$

Where $r(t)$ is the received signal. It is sampled at $t = \frac{k}{f_s}$. and$k$ is the sampling instances. The output of the FFT block is represented as

$$\widetilde{D}_m = \frac{1}{N}\sum_{k=0}^{N-1} r_k e^{-j\frac{2\pi mk}{2N}} \tag{2.29}$$

Where r_k is given as

$$r_k = \sum_{k=0}^{N-1} H_n D_n e^{j\frac{2\pi nk}{2N}} + n(k) \tag{2.30}$$

$H_n(z)$ represents the discrete channel impulse response which is to be equalized.

2.4.3 Channel Estimation in OFDM

The wireless channels are frequency selective and time-varying in mobile communication systems [78, 81, 82]. Therefore equalization is mandatory at the receiver. OFDM channel estimation techniques can be classified into two categories.

- **Blind:** These techniques estimate the channel state without the prior knowledge.
- **Pilot-aided:** The known data is appended to the transmitted signals to estimate the channel. This technique is mostly used which employ various interpolation techniques to estimate the channel response of the subcarriers between pilot symbols [83].

2.4.4 Pilot Structures

There are three different types of pilot structures according to the arrangements of pilots: block type, comb type, and lattice type. These are depicted in Figure 2.19. The OFDM symbols with pilots are regularly transmitted at all subcarriers for channel estimation. A time-domain interpolation is performed using these pilots to estimate the channel along the time-axis. To keep track of the time-varying characteristics of the channel, the pilot symbols must be positioned as often as the coherence time is. It was built under slow fading channel assumption. In this type, the channel estimation may be based on either LS or MMSE. The MMSE estimation showed a 10-15 dB gain in SNR over LS estimation [84].

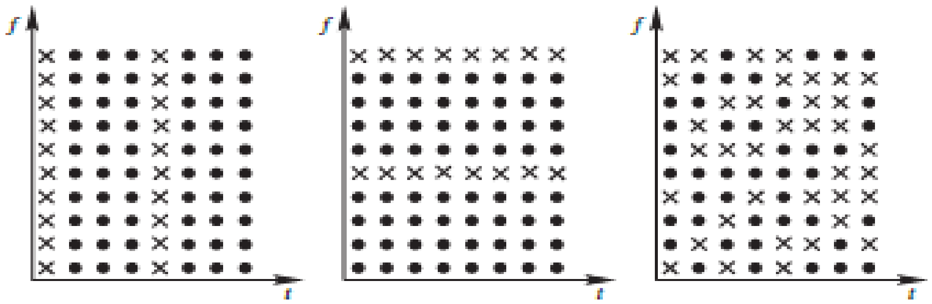

Figure 2.19 Pilot structures of OFDM [84]

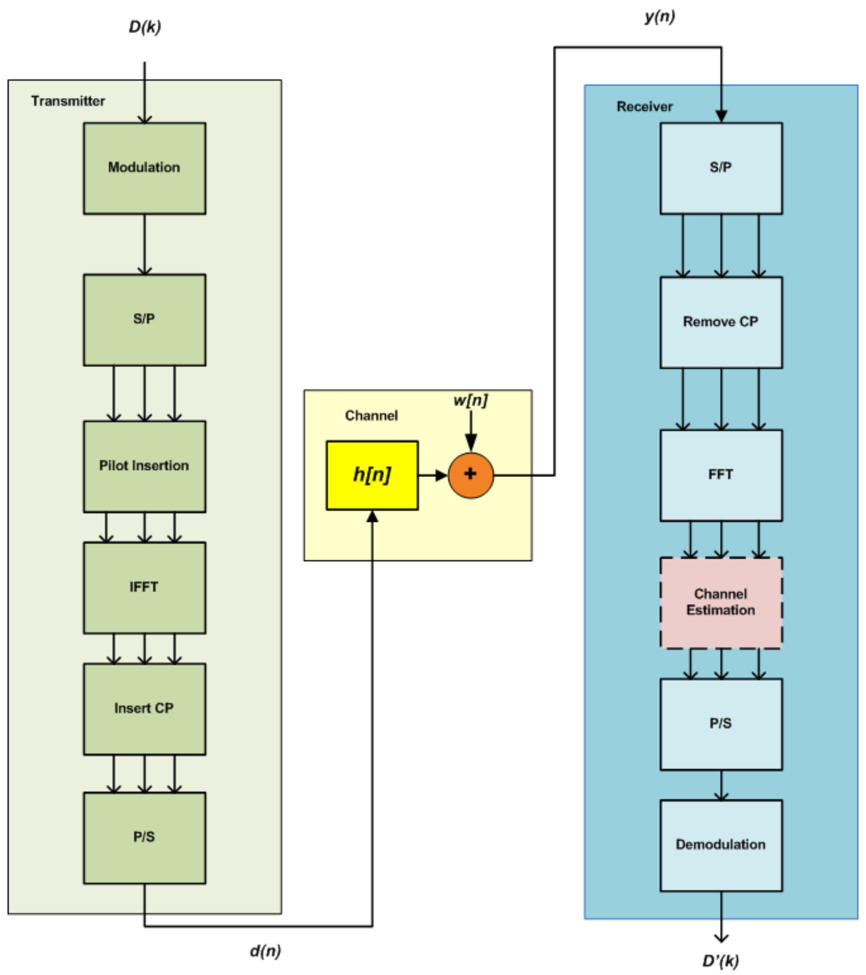

Figure 2.20 The system model of OFDM [85]

In comb type block structure, pilot tones are placed periodically along frequency axis. In contrast to the block type structure, comb type is suitable for fast fading channels. In Lattice type, pilot tones are positioned along both the time and frequency axes at certain given interval. This allows interpolations in both axes to estimate the channel.

2.4.5 Related Work

The literature includes many channel estimation techniques for OFDM. For example, the work in [86] considers the LS channel estimation for MIMO OFDM. This scheme uses the LS algorithm on pilot tones to compute the MSE and then derived the expression for the optimal pilot structure and sequence. This work is further improved by using RLS algorithm in which the parameter forgetting factor is derived. The simulations have shown that RLS is better and can achieve a significant gain in terms of SNR, particularly when a channel is time-varying.

In this work [87], two channel estimation techniques, Maximum likelihood Estimation (MLSE) and Bayesian MMSEE have been compared. The plus point of MLSE is that it does not require the prior knowledge of the channel statistics, thus makes it simple to implement. It is concluded that if pilot tones are greater than the channel's impulse response (CIR's) than MMSE and MLSE performs equally good, however, MMSE is better with high computational complexity.

In [88], pilot signals with comb type structure, are estimated using LS and MMSE algorithms together with interpolation schemes. A similar work is done in [89] where channel estimation based on comb type pilot structure is studied. LS and LMS algorithms are applied on pilot sub carriers. The channel is interpolated using many ways such as linear interpolation, low pass interpolation, spline cubic etc. This work also includes the implementation of the DFE. The performance of all these techniques is also compared in terms of BER.

In [90] channel is estimated and equalized using the proposed back propagation neural network (BPNN) algorithm for OFDM systems. Channel models considered in this work was Rayleigh, Rician and TU6. It was observed through simulations that BER is improved but it has a huge computational complexity because of the training algorithm.

In [85] NN has been applied for the channel estimation and equalization. The offline model is used which was trained separately. Deep learning based channel estimation shows better BER as compared to MMSE estimator.

A new design of the receiver using the feed forward neural network is proposed [91]. This work has the advantage that only one neural network was used for estimation, equalization and demodulation. The downside of this work is that it is not computationally efficient due to neural networks. In [92] SVM robust version is proposed for demodulation of OFDM symbols in presence of impulse noise. This work applies the regression approach instead of classification approach of SVM. Simulations showed superior performance.

2.5 Performance Comparison of NN based Channel Equalization Schemes

Up to this point in this work two types of communication systems i.e. the single-carrier systems and the multi-carrier OFDM systems have been described. Channel equalization methods of the respective systems have also been highlighted. A critical review of the methods has been provided. All the methods have been found to perform well in Rayleigh communication channels. However there is a need to compare the schemes and highlight the best possible NN scheme for the channel equalization. To the best of the knowledge of the author this work has not been carried out in the literature. In this work the selected NN's are used for channel equalization and their performance has been compared. Simulation frameworks as described in the later chapters are developed for the single-carrier and for the multi-carrier OFDM systems. The performance criteria used was selected to be the BER as this is the primary criteria in communication systems for the performance measurement of the equalizers. The results show that the performance of the NN based channel equalizers perform better than LMS based channel estimators. Further results are discussed in detail in the following chapters.

2.6 Summary

This chapter provides a comprehensive overview of the channel estimation and equalization techniques. Different neural network structures are discussed in the context of channel equalization. The MLP network implementation is simple, but training takes a lot of time. The main disadvantage of the FLANN structure is its computational and time complexity which gradually increases as the number of input nodes increases. RBF-based neural network equalizer is an interesting alternative and has been successfully used for blind equalization. LSTM equalizers are superior to NN feed forwards, including MLPs, RBFs, and FLANNs.

CHAPTER 3

WIRELESS CHANNEL EQUALIZATION USING NEURAL NETWORKS

Wireless communication is one of the most exciting areas in the communication field today. It has grown rapidly over the last two decades due to the rapid rise in demand for wireless connectivity. Wireless communication, however, is not as effective as wired media communication, because of fading and other channel induced propagation effects [93]. Thus the techniques for enhancing its efficiency and reliability become the basic objectives of current research. To enhance the performance of wireless communication, numerous NN based channel estimation and equalization techniques have been proposed in the literature. These techniques include MLP, RBF, FLANN, SVM, and LSTM.

This chapter deals with the implementation and simulation of these techniques and organized as follows. We first describe the QPSK modulation scheme and then describe the channel used in our work. This is followed by the description of the simulation parameters of NNs used for channel estimation and equalization. At the end of this chapter, a detailed result analysis has been made.

3.1 Implementation of NN based equalizers

ML techniques are setting a path to replace the conventional communication techniques and the combination of these two fields has led to a lot of successful work. NNs are capable of processing nonlinear data and can produce complex decision regions. Therefore, NNs can be employed for equalization to overcome the difficulties associated with channel nonlinearities [8-10]. The simulation setup is depicted in Figure 3.1.

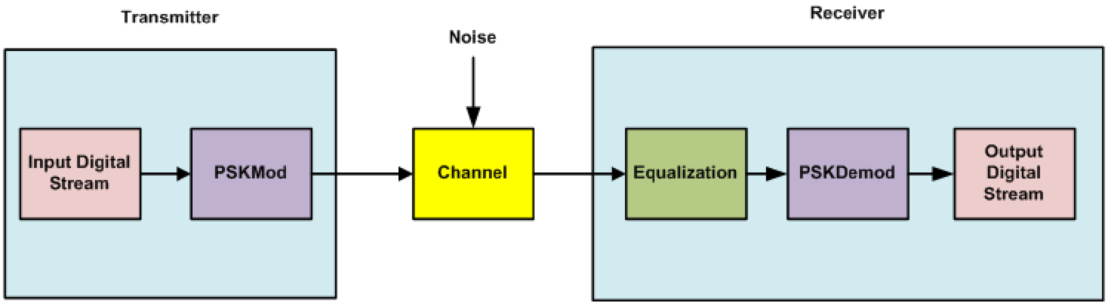

Figure 3.1 The simulation setup

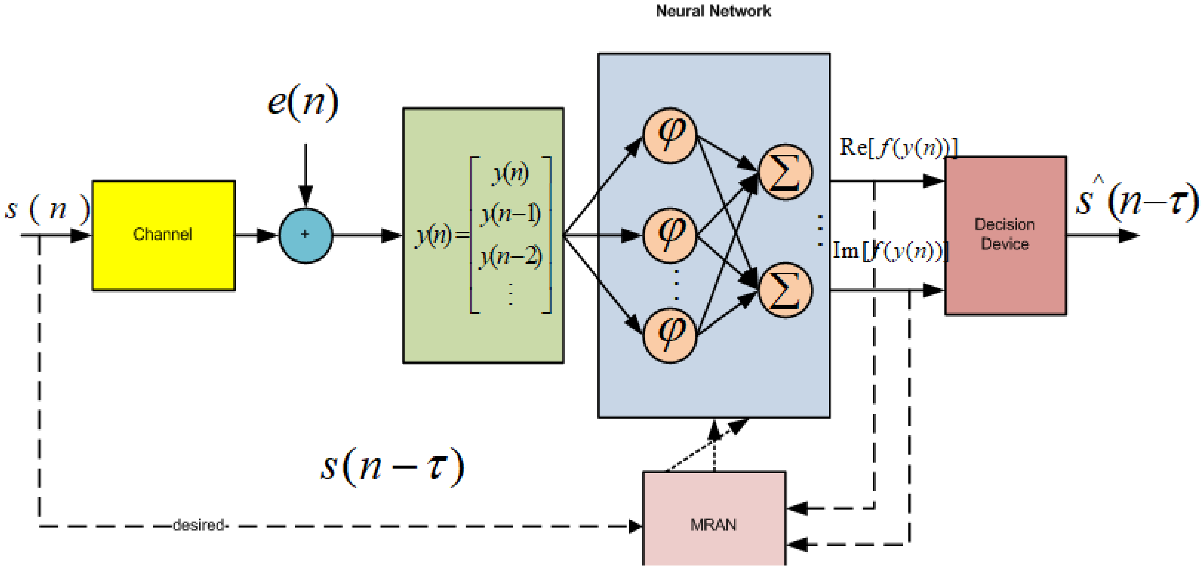

Figure 3.2 NN based equalizer [94]

A typical NN based channel equalizer is depicted in Figure 3.2. The transmitter first transmits the training symbols which are known to both the receiver and transmitter and then transmits the actual data. The equalizer uses the received training symbols to learn the equalizer weights. The optimization criterion is to minimize the MSE.

3.1.1 Data Generation and QPSK Modulation

Data is randomly generated using MATLAB rand function. It generates uniformly distributed data between 0 and 1. The data is QPSK modulated and then passed through the channel filter. QPSK uses two signals I and Q, where I is a in-phase signal and Q is a quadrature signal. Both of these signals are at 90^0 phase difference. This modulation is popular due to its simpler design and efficient hardware realization. The block diagram of QPSK modulator is shown in Figure 3.3. The following steps are performed to produce a QPSK modulated signal.

- The incoming digital data is converted into two streams. One stream contains the odd bits and the other takes the even bits from the original stream.
- The streams are then pulse shaped using root raised cosine pulses. The duration of the pulse determines the data rate of the transmitter. In this phase the incoming data is first up-sampled by a factor N which corresponds to the symbol duration and then convolved with the RRC pulse. The resulting signal is termed as baseband signal.
- The resulting I and Q stream are then multiplied with I/Q carrier signals. In other words, these streams are amplitude-modulating using I/Q signals.
- Finally, the two modulated signals are summed up to form a QPSK-modulated signal. In QPSK, two bits are used in one symbol.

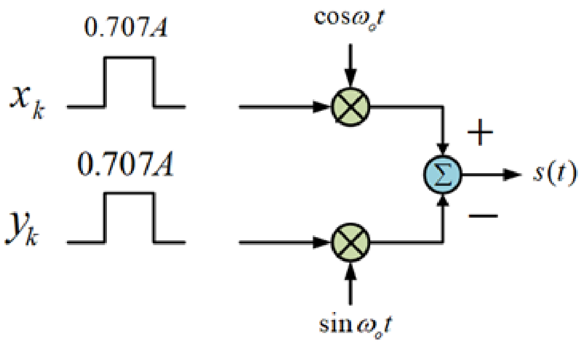

Figure 3.3 Block diagram of QPSK modulator [1]

Mathematically, QPSK modulation can be derived as follows:

Let m_k represent the message signal. Where $m_k = x_i + jy_i$ is the complex representation of the i[th] message signal. This complex representation represents the group of bits together. One is represented as the real and the second one represents the imaginary bit. The message signal is QPSK modulated as

$$\begin{aligned} s_{QSPK}(t) &= Re\,\{m_k e^{j2\pi f_o t}\} \\ s_{QSPK}(t) &= Re\,\{(x_i + jy_i)\big(\cos(2\pi f_o t) + jsin(2\pi f_o t)\big)\} \\ s_{QSPK}(t) &= x_i \cos(2\pi f_o t) - y_i \sin\,(2\pi f_o t) \end{aligned} \tag{3.1}$$

Where $x_i = 0.7071A$ and $y_i = 0.7071A$ are the amplitudes of the pulses.

$$s_{QSPK}(t) = 0.7071A \cos(2\pi f_o t) - 0.7071A\sin\,(2\pi f_o t) \tag{3.2}$$

Using trigonometric relations the equation can be simplified as

$$s_{QSPK}(t) = A \cos\left(2\pi f_o t + \frac{\pi}{4}\right) \tag{3.3}$$

From equation (3.3) the four reference constellation points of QPSK modulation are given as

$$m_i = \begin{cases} \;\;0.7071 + j0.7071 \\ -0.7071 + j0.7071 \\ -0.7071 - j0.7071 \\ \;\;0.7071 - j0.7071 \end{cases} \tag{3.4}$$

And the constellation plot is shown in Figure 3.4.

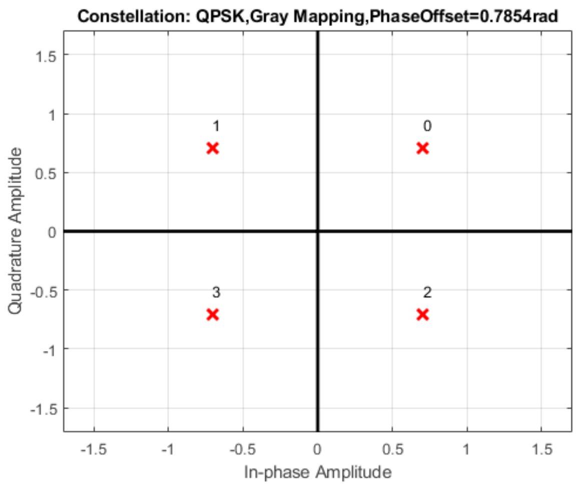

Figure 3.4 Constellation plot of QPSK

A typical QPSK receiver is depicted in Figure 3.5. At the received signal is demodulated as follows. The received QPSK signal is multiplied with the local oscillators which are at 90 degrees phase difference and are called I and Q. The resulting signals are low pass filtered using the RRC filters. This result in the recovery of the baseband pulses which are further down sampled by N and are signal is received. A detailed discussion of the receivers and transmitter can be found in [1]

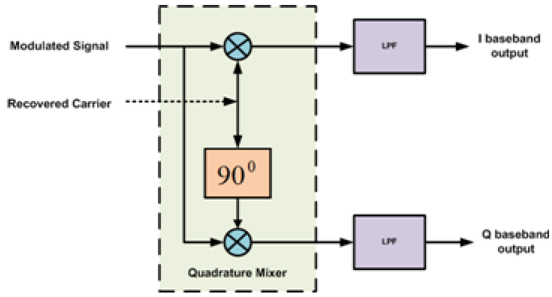

Figure 3.5 QPSK demodulator [1]

The received signal can be expressed mathematically as

$$r(t) = s(t) * h(t) + \eta(t) \tag{3.5}$$

Equation (3.5) shows that the received signal r(t) is the sum of convolution of h(t) with transmitted signal s(t) and with noise n(t) added.

3.1.2 Wireless Channel Model

The wireless channel model describes the underlying communication medium. The performance of communication system is dependent on the condition of the channel. Rayleigh and Rician fading channel models are widely used to simulate the channel in that realistic wireless environment. Rayleigh fading channel [95-97] is the conceptual model assuming the fact that there are several objects in the atmosphere. Due to these objects, the transmitted signal may be dispersed and replicated. It is also presumed that the there is no direct path between the transmitter and the receiver. On the other hand, Rician channel [95, 96, 98] assumes that there is a direct path between the transmitter and the receiver.

The received signal contains both the dispersed and scattered (or reflected) paths. In this case, the scattered (or reflected) paths appear to be weaker than the direct path.

We have considered a complex valued multipath channel mentioned in [43]. The scatter plot of this channel is depicted in Figure 3.6. The coefficients of this channel are defined as:

$$c = [1 - 0.3434j \quad 0.5 + 0.2912j] \tag{3.6}$$

$$H(z) = c_1 X(z) + c_2 z^{-1} X(z) \tag{3.7}$$

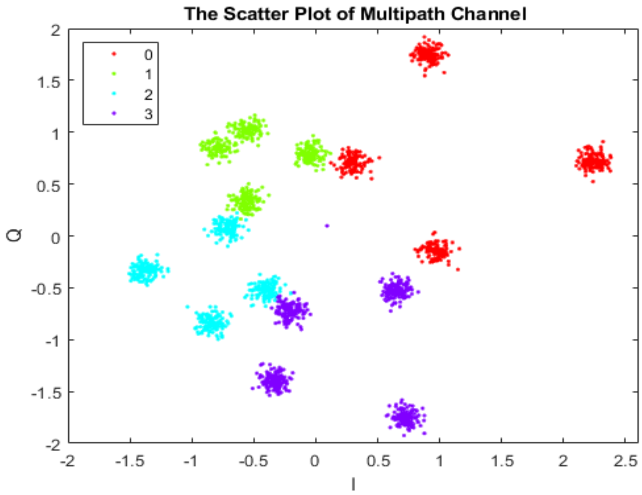

Figure 3.6 The scatter plot of channel

3.2 Simulation Parameters of NN

Different NN equalizers were plugged into the configuration of Figure 3.1 and results were obtained. These configurations and the respective results are discussed in the sequel. The primary performance criteria used was BER. Loss function analysis and the computational complexity are also calculated. The detailed results are compared and discussed in the later sections.

3.2.1 The Flowchart of NN based Equalizer

The flow chart of the NN based equalizer is depicted in Figure 3.7.

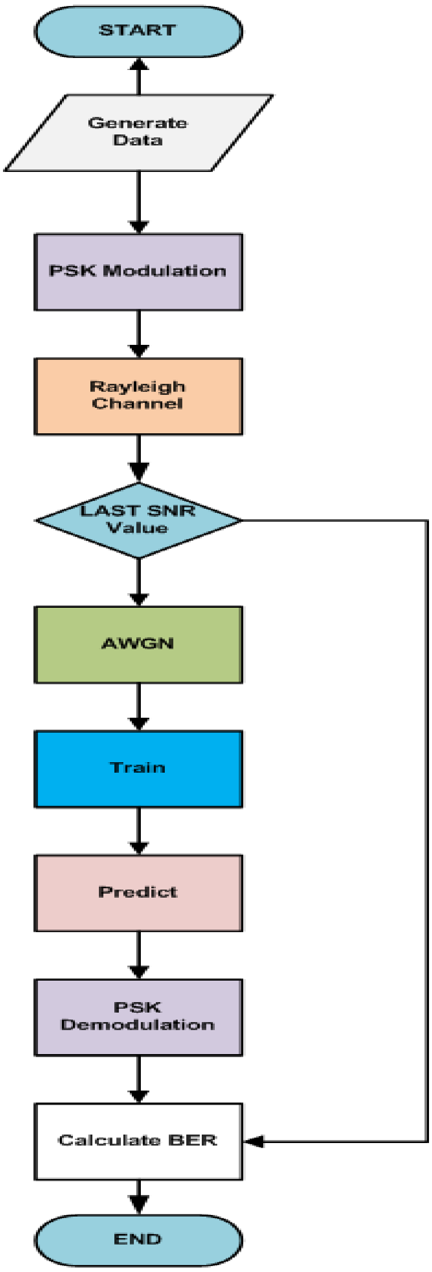

Figure 3.7 The flowchart of NN based equalizer

3.2.2 MLP based Equalizer

MLP is a simple three-layer network that maps the input to the output. The detailed description of MLP is presented in section 2.3.2. MLP is designed using ***nntraintool*** of MATLAB. It comprises of an input layer, a hidden layer and an output layer. The structure of MLP is shown in Figure 3.8. The input layer contains two vectors. One vector is the real part of the input signal (X) and another is the complex part of the signal. The output layer generates four vectors Y_0 to Y_3. The MLP is trained with these parameters as shown in Table 3.1. The MATLAB output during the training phase is depicted in Figure 3.9. The convergence of the loss function is depicted in Figure 3.10. Total epochs run were thirty three. The error was reduced significantly.

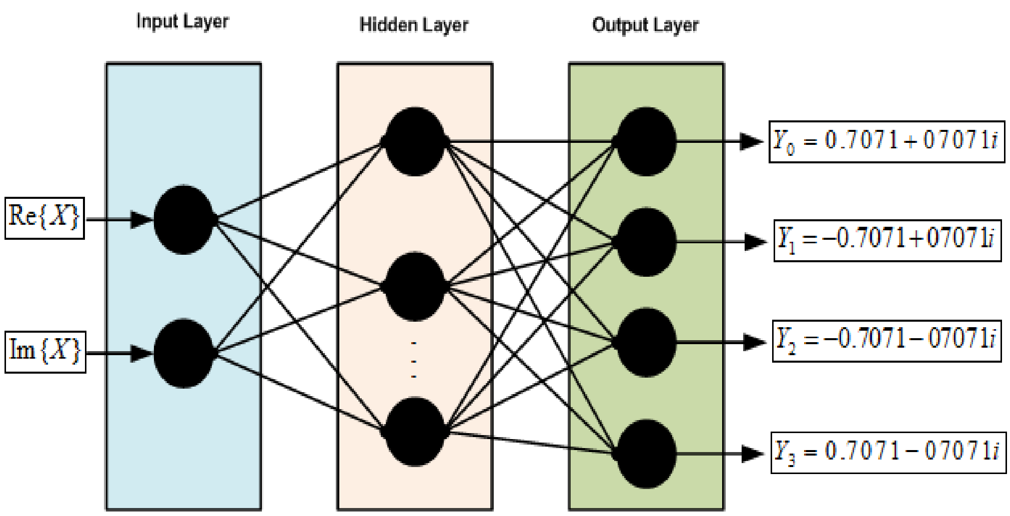

Figure 3.8 MLP based equalizer

Table 3.1 The simulation parameters of MLP

Parameter	**Value**
Hidden Nodes	30
Input Size (X)	1,000,000
Training Algorithm	Scaled Conjugate Gradient (SCG) [99]

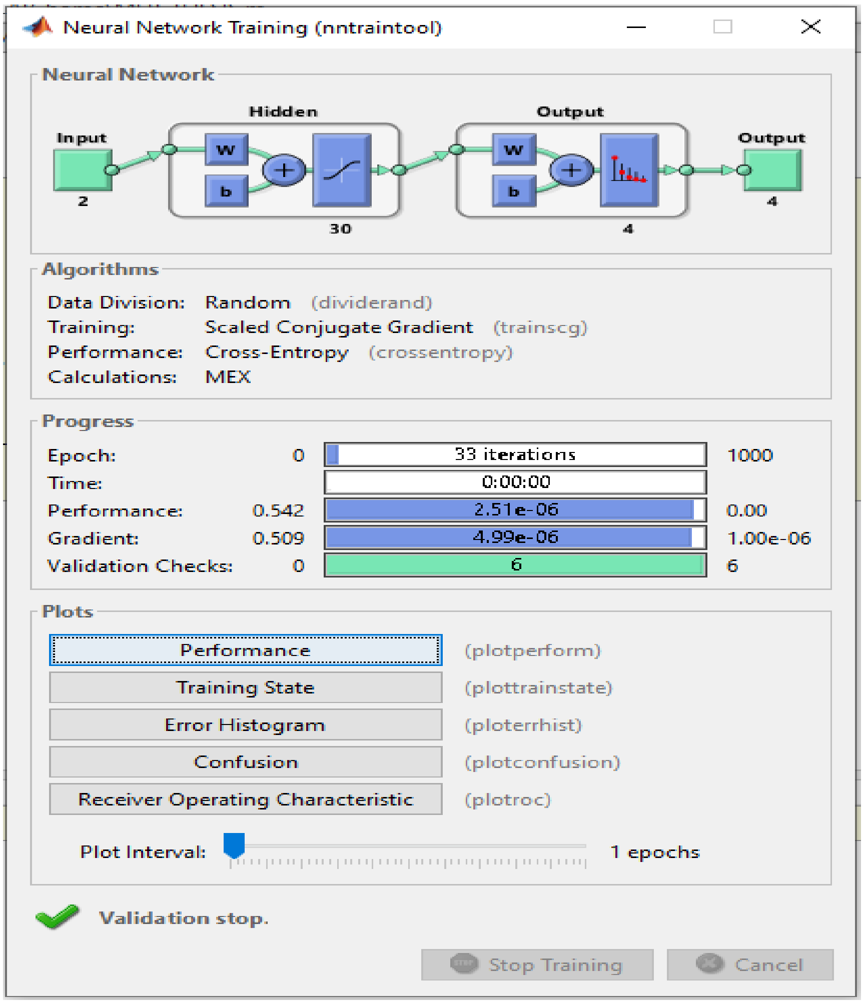

Figure 3.9 The training progress of MLP

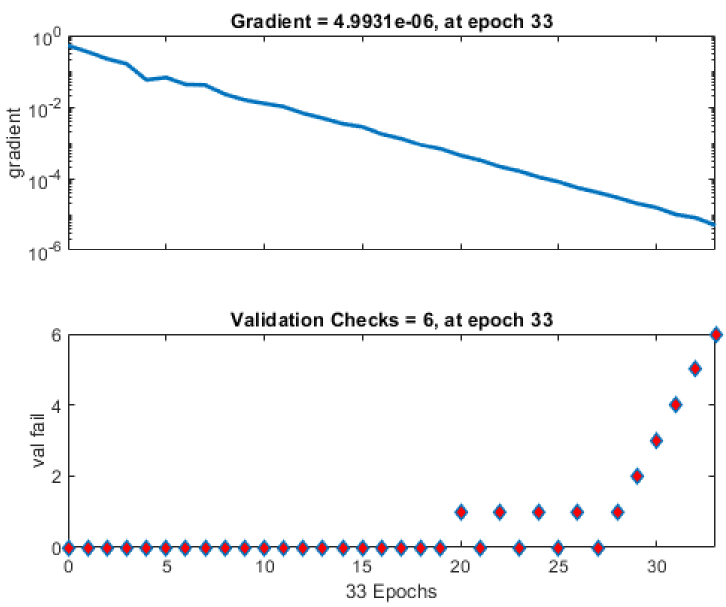

Figure 3.10 The gradient loss during training

3.2.3 RBFNN

It is a three-layer network which comprises an input layer, a nonlinear hidden layer and a linear output layer. Radial functions are used as an activation function. Radial functions are special functions. The output of these functions increases or decreases monotonically with distance from a center. K-MEAN algorithm is used to find the centers. So first, centers of clusters are determined in an unsupervised manner and then classification is performed to recover the signal. We have implemented this work[43] and observed the improved BER.

3.2.3.1 Simulation Parameters

Table 3.2 depicts the simulation parameters for the RBF NN.

Table 3.2 The simulation parameters of the RBFNN

Parameter	Value
Data set size	2000
Noise variance	0.01
Centers	16

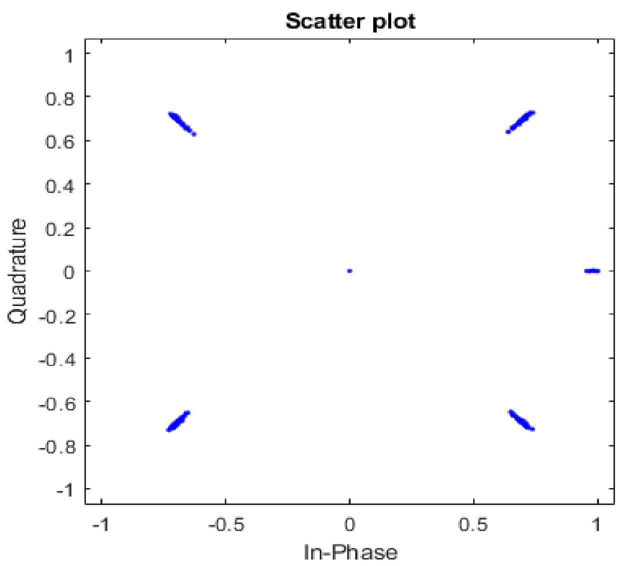

Figure 3.11 The scatter plot of equalized signal using RBF

3.2.3.2 Simulation Results

The scatter plot depicts the equalized received signal. It is clear from the Figure 3.11 that the RBF based equalizer has removed all the channel related impairments.

3.2.4 FLANN

It is a single layer neural network. The main concept of FLANN is to convert the input data to a higher dimension by using different functional expansions. Due to absence of hidden layers, these networks have following advantages listed below.

- Low computational complexity with very few adjustable parameters.
- Faster training time.
- Simple design so can be implemented on hardware.

Using the work in [33, 46], we have implemented the FLANN based equalizer. The block diagram of equalizer is shown in Figure 3.12.

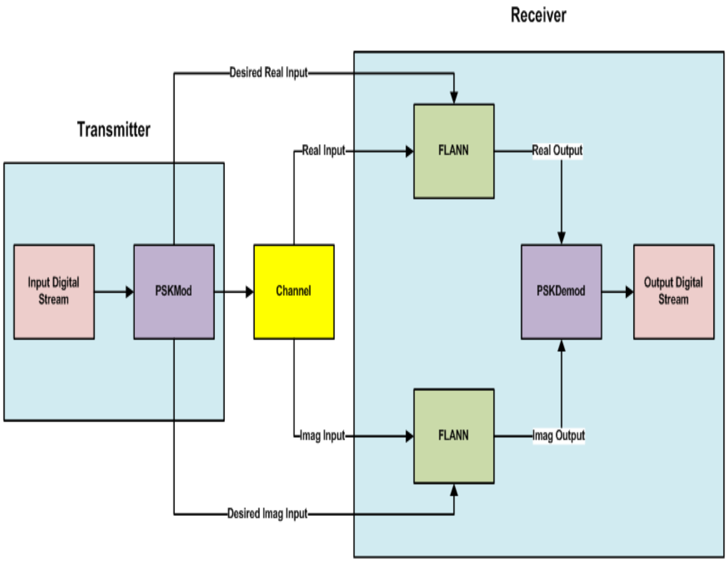

Figure 3.12 The block diagram of the implemented FLANN based equalizer.

3.2.4.1 FLANN based Equalizer Simulation Parameters

Table 3.3 depicts simulation parameters of the FLANN based NN Equalizer.

Table 3.3 The simulation parameters of FLANN, Le-FLANN and Ch-FLANN

Parameters	Value
Length of input	2000
FLANN order	30
Input size	4
μ	0.01
No of iterations	10
Channel Noise Variance	0.01

3.2.4.2 Simulation Results

Simulation results of trigonometric FLANN are depicted in Figure 3.13.

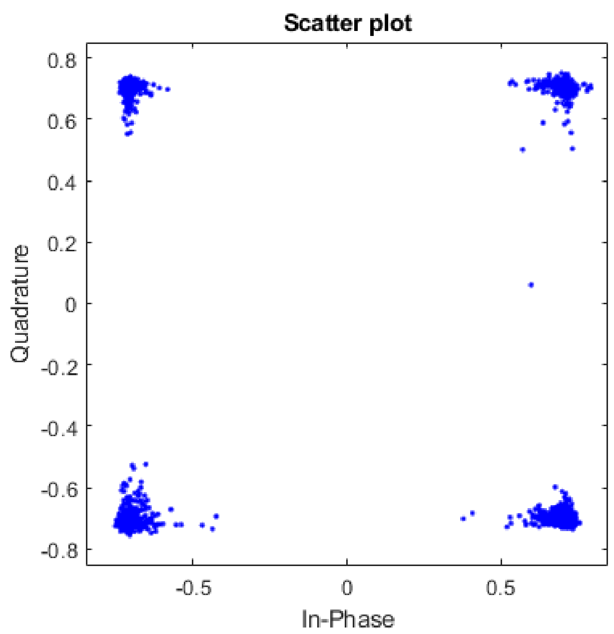

Figure 3.13 Scatter plot for FLANN

3.2.5 SVM Based Channel Equalizer

SVM is a supervised algorithm used for classification problems. Channel estimation is a classification problem so it can be used to deal with the nonlinear channel effects. In our work, we have implemented a basic SVM model equalization. Simulation parameters of the SVM are shown in Table 3.4. The generalization error computed during simulations is 0.00001 which indicates the best performance.

Table 3.4 The simulation parameters of SVM

Parameters	Value
Input Size	1,000,000
Kernel Function	KNN

3.2.6 LSTM Channel Equalizers

LSTM is a popular RNN based DL technique. It is different from feedforward NN which does not have memory and cannot deal with temporal data. A detailed description of LSTM is provided in section 2.3.7. The LSTM network used in this work is depicted in Figure 3.14.

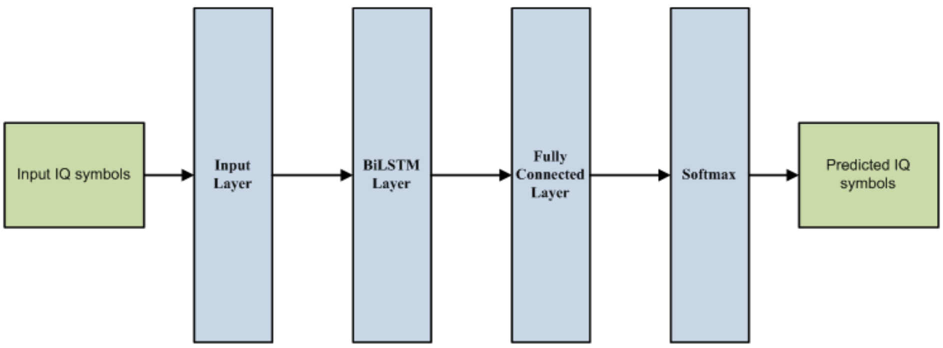

Figure 3.14 The LSTM model

3.2.6.1 LSTM Simulation Parameters

The Simulation parameters of LSTM are shown in the Table 3.5. The simulation of LSTM was executed in MATLAB. The convergence of the training phase is depicted in Figure 3.15. The minimum value of loss function achieved was 0.0001.

Table 3.5 The simulation prameters of LSTM model

Parameters	Value
Training SNR	12dB
Channel Noise Variance	0.01
LSTM Nodes	16
Learning rate	0.01

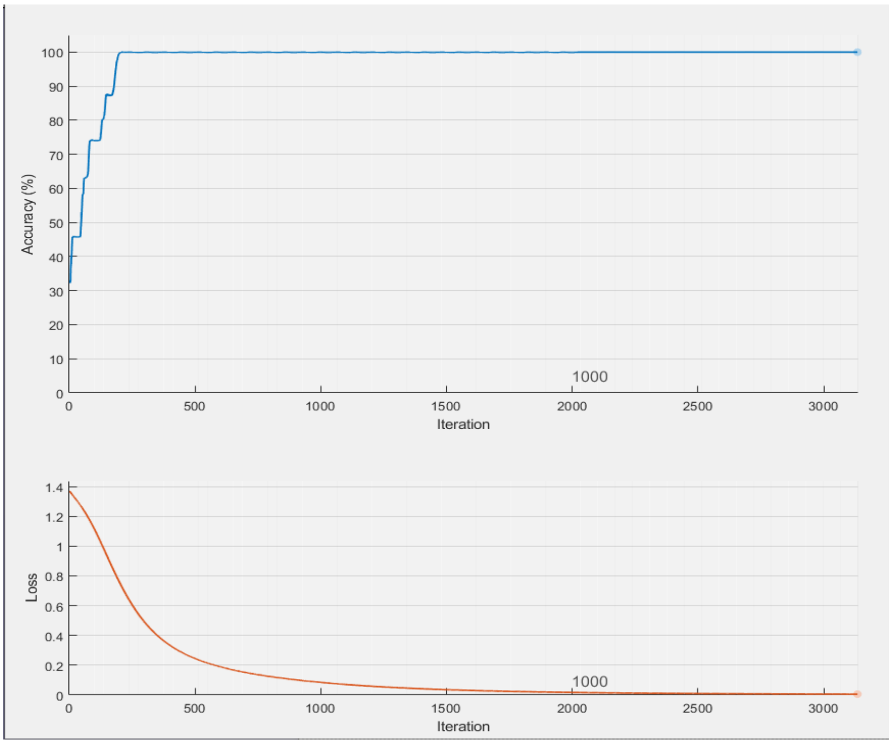

Figure 3.15 Training progress of LSTM model

3.3 Simulation Result Analysis

All the simulations executed in the Section 3.2 are compared in this section. Figure 3.16 depicts the BER comparison of all the simulated NN's. Generally, the trend verifies already establish theory. As the SNR increases the BER performance is getting better and better. The performance of FLANN is slightly bad as compared to the rest of the schemes due to its single layer architecture. The performance of the traditional LMS algorithm is worst. In this work [43] the similar results are observed. All the other ML based schemes are having same BER performance. In Figure 3.17 the zoomed version of the BER graph is depicted. The LSTM is slightly bearing higher BER then SVM and RBF based ML methods. The performance of FLANN when compared with rest is almost 4dB poorer than the rest. The performance of LSTM is about 0.7 dB poorer then the RBF and SVM and MLP. This may be reduced by further tuning or by increasing the size of Neural Network. However, this will be at the cost of time and computational resource which can be very expensive in the communication systems.

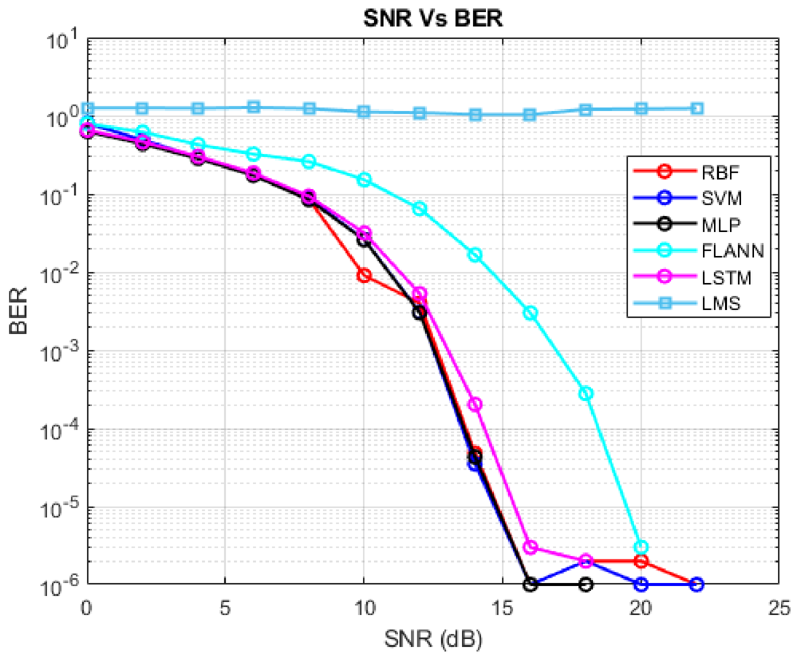

Figure 3.16 The comparison of NN techniques

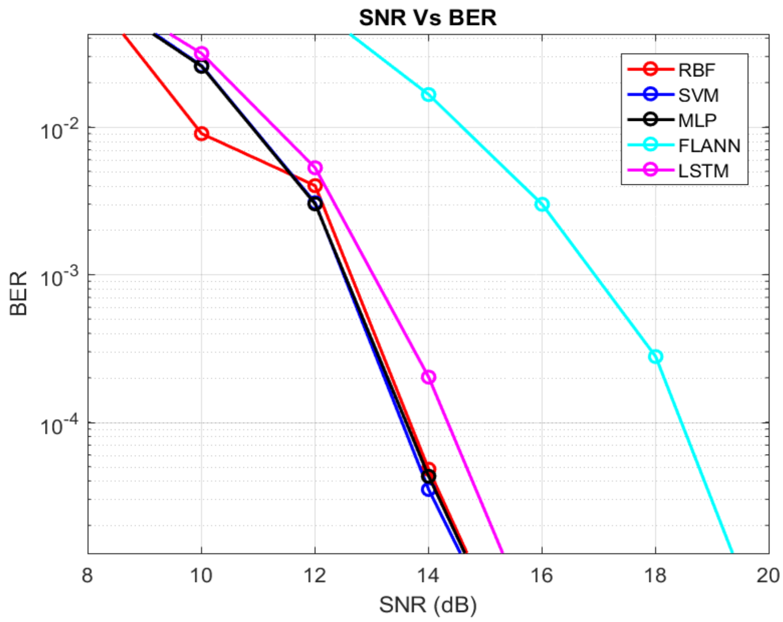

Figure 3.17 The zoomed comparison plot of BER

Loss function is an important parameter of the optimization and is therefore discussed. The lesser the value of the loss function the better the performance is considered. In Table 3.6 the values of the loss function for all the algorithms used in this text are depicted. It shows that all the NN's are well trained. The minimum value of the loss function achieved is in the case of SVN where the value is 0.00001. The BER results depicted in the Figure 3.16, and Figure 3.17 are very much in line with these results. The loss function values of RBF, FLANN and LSTM can be further reduced by using more training data and by using better optimization algorithms.

Table 3.6 Loss function values of ANN

ANN	Loss function value
Radial Basis Function	0.001
SVN (KNN based)	0.00001
MLP	0.003
FLANN	0.005
LSTM	0.003

3.3.1 Computational Complexity

Computational complexity analysis of the algorithms is presented. This presents the number of computations resources required to perform the respective ANN. Table 3.7 presents the computational complexity of various algorithms. The number of additions, multiplications and other computational resources such as exponentiation, powers and trigonometric functions are enlisted [100]. This analysis is useful for the HW implementations and for estimating the computational requirements for embedded systems.

Table 3.7 Computational complexity of various ANN

ANN	**Computational Complexity**				
	No of Mul	**No of Additions**	**Division**	**Tanh**	**Exp ()**
RBF	$n_0n_1 + 2n_1 + n_0 + 1$	$2n_0n_1 + n_0 + n_1 + 1$	$n_0 + n_1$	-	n_o
FLANN	$3n_1(n_0 + 1) + n_0$	$2n_1(n_0 + 1)$	-	n_1	-
MLP	$4\sum_{i=0}^{L-1} n_in_{i+1} + 3\sum_{i=0}^{L-1} n_i - n_0n_1 + 2n_L$	$3\sum_{i=0}^{L-1} n_in_{i+1} + 3n_L - n_0n_1$	-	-	-
SVM	$4T_d$	$3T_d$			T_d
LSTM	2N+2N	N_0N_1			

n_0 is the number of input nodes. n_i is the ith node in network. T_d is the total number of Euclidean distances in the KNN based SVM. N is the number of nodes in the convolution layer of LSTM network. N_0N_1 are the number of Nerons in the BiLSTM layers [100].

The computational complexity analysis of the mentioned algorithms was verified by timing the MATLAB® implementations. The time of all the algorithms used in this work was measured using MATLAB® built-in function called 'timeit'. Number of iterations performed for each algorithm was 10^6. Machine used for computation was DELL® 7920 running MATLAB® 2019b. The CPU was Intel® Xeon(R) Silver 4116 CPU running at 2.1GHz. The time is enlisted in the Table 3.8. The computational time computed endorses the computational complexity as given in the Table 3.7. The minimum computational time achieved is for the SVM. SVM is running KNN algorithm which is computationally efficient. Its BER results are also amongst the best. RBF and MLP bear good performance but their computational time is more.

Table 3.8 Time of various ANN algorithms

ANN	**Total Time in seconds**	**Iterations**	**Time Single Iteration (seconds)**
SVM	5.2315 sec	1000000	5.2315×10^{-6}
LSTM	102.9576 sec	1000000	1.029576×10^{-4}
RBF	12.2729 sec	1000000	1.2273×10^{-5}
FLANN	36.3450 sec	1000000	3.6345×10^{-5}
MLP	13.0859 sec	1000000	1.3086×10^{-5}

3.4 Summary

This chapter describes the implementation and simulations of NN based equalizers. The performance of these equalizers is also compared. Rayleigh fading channel with varying noise and multi-paths have been implemented to simulate the actual propagation environment. QPSK modulation scheme is applied on data. Simulations have confirmed that these techniques are better than the conventional methods. BER is used as the primary criterion to measure performance. Other performance parameters which are compared using simulation results are the loss function and the computational complexity. The computational complexity calculated with the help of simulations is endorsing the mathematical formulations of the computational complexity as listed in this chapter.

CHAPTER 4

CHANNEL EQUALIZATION TECHNIQUES IN OFDM

The wireless communication has been evolved from the single voice transmission (2G) to massive real-time HD video streaming (4G) and is now progressing towards 5G. Thus, the requirements for high-speed transmission with more capacity are increasing on daily basis. Multi carrier techniques can fulfill these requirements. OFDM is a multi-carrier technique which has been utilized for many high rate transmission applications such as WiFi and 4G LTE etc. Traditionally, in OFDM multiple modulators and demodulators were used at transmitter and receiver for the generation of multiple carriers. OFDM has gained popularity due to the incorporation of a simple FFT operation, which eliminates the need of these modulators. The basic principle behind this transmission technique is to split the total available bandwidth into several narrowband sub channels that are equally and closely spaced and the spectrum of these sub channels are overlapped to save the bandwidth. Other than that, the primary motivation for using this technique is its robustness against multi path channel effects which makes the equalization process much simpler than in a single-carrier communication. Thus, a simple one tap equalizer can be employed that makes the design of a receiver much simpler. In OFDM, frequency domain equalization is used instead of time domain equalization as in single-carrier communication.

This chapter is structured as follows. The first section gives an overview of the channel estimation techniques in OFDM. Section 4.2 presents the LSTM model for channel estimation. The later sections explain the simulation parameters and simulation results in comparison with LMS and MMSE.

4.1 An Overview of Channel Estimation Techniques in OFDM

OFDM channels suffer from multipath effects, Doppler spreads, timing and frequency offsets. Therefore, channel estimation is necessary on the receiver as it directly affects the BER performance of the OFDM systems. So, channel estimation techniques need to be robust enough to cater all these channel induced effects. Extensive work has been done to resolve these issues. LS and MMSE are traditional channel estimation methods [101]. LS based estimation is simple and does not require any prior knowledge of channel. However, its detection performance is not good in comparison with MMSE based estimation, which is based on the second order statistics of the channel. The MMSE performs better, but at the cost of high computational complexity. OFDM uses multiple carriers to transmit data and FFT defines the total number of sub carriers. Some sub carriers are used for data symbols. Some sub carriers are reserved only for pilots. The pilot sub-carriers are utilized to estimate the channel. LS and MMSE estimation techniques are applied on the pilot symbols. The estimated response on these pilot symbols are then used on the data symbols. There are also some unused sub carriers called null carriers used to avoid interference.

DL [102] has been successfully applied in many fields and showed up to the mark performance over traditional techniques. It is possible because of its nonlinear flexibility, which makes the complex problems simple to solve. Due to its promising results, it has also been applied in the communication field as well, particularly in the domain of channel estimation and showed remarkable results in terms of BER. DL based techniques can extract the underlying nonlinear characteristics of the channels. This chapter deals with the application of DL techniques in channel estimation of OFDM systems.

4.2 Channel Estimation using LSTM based Neural Network

LSTM is a type of RNN which deals with the sequence data and use it to predict the future sequence data. It has been successfully applied in video classification, speech recognition and NLP. LSTM has also been considered in the communication domain and showed better results. For example, this work [85] utilizes LSTM for joint channel estimation and detection of symbols. Using this work as a reference, we have implemented the LSTM based channel estimator with slight modifications. We have found similar and sometimes better results than MMSE.

4.2.1 Simulation Model

A baseband OFDM system with LSTM model is depicted in Figure 4.1. The OFDM system is same as the traditional one and has been discussed in the previous chapter. The design involves two stages:

- In offline stage, a model is trained by using the simulation data.
- In online stage, the pre-trained model is deployed online into recover the data.

It is obvious from this diagram that the use of LSTM or other NN based OFDM receiver structures are simpler when compared with the traditional OFDM receivers where the channel estimation is performed in a very complex manner by using many of the pilot symbol based techniques already discussed.

The time domain representation of the transmitted signal $x[n]$ after passing through the multipath channel $h[n]$ is:

$$y[n] = x[n] * h[n] + w[n] \tag{4.1}$$

Where $w[n]$ represents the Gaussian noise (AWGN).The receiver performs DFT on $y[n]$ to recover the data. The frequency domain representation of $y[n]$ is:

$$Y[K] = X[K]H[K] + W[K] \tag{4.2}$$

Where $Y[K]$, $X[K]$, $H[K]$ and $W[K]$ are the DFT of their corresponding time domain signals.

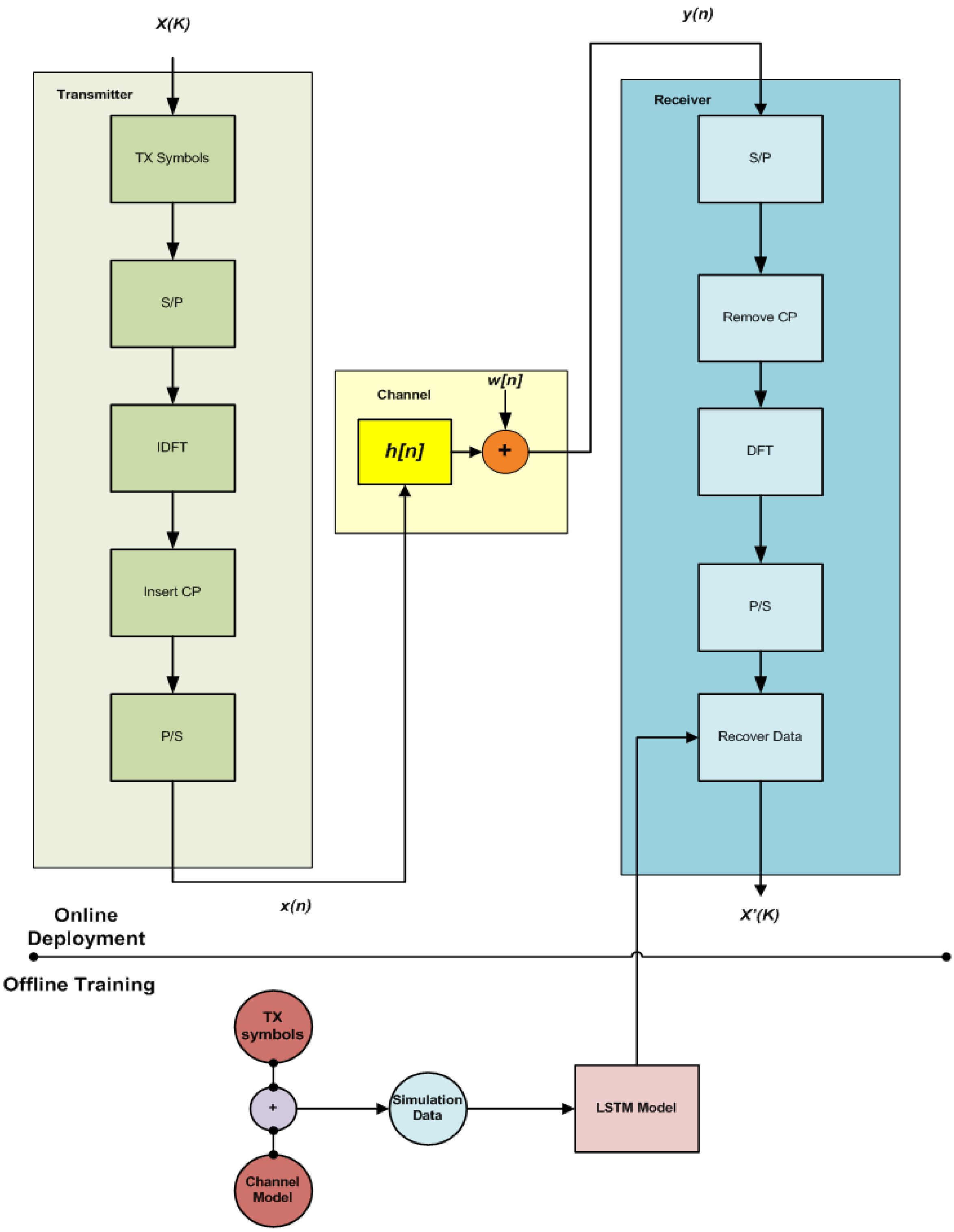

Figure 4.1 OFDM simulation setup [85]

4.2.2 OFDM Simulation Parameters

OFDM system parameters used in this simulation are listed in Table 4.1

Table 4.1 The simulation parameters of OFDM

Parameter	Value
Total sub-carriers	64
Pilots	64
Sub carrier Modulation	QPSK
Cyclic Prefix	16

4.2.3 The LSTM Model

The block diagram of the LSTM model is shown in the Figure 4.2. It consists of input layer, LSTM layer, dense layer, softmax layer and an output layer. The number of hidden units in LSTM layer are 16 with learning rate of 0.01.

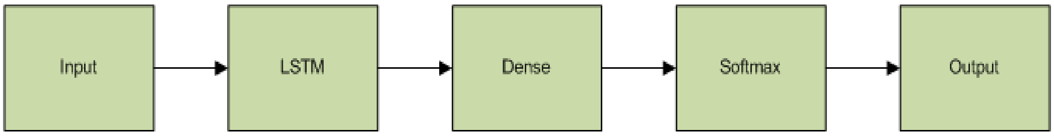

Figure 4.2 The LSTM model

4.2.4 Dataset Generation

The training data is generated through simulations. The simulation framework is shown in Figure 4.1. First, OFDM frames are generated which consists of one data block and one pilot block. A pilot block contains a fixed known sequence, followed by a data block. A data block contains randomly generated data symbols. All subcarriers contain the same pilot sequences, thus having a block type pilot structure. After passing through the multipath channel, these frames are given as an input to model. The received signal and the originally transmitted signals are collected as a training data set.

4.2.5 Model Training

A model is trained on the training dataset. A model compares the originally transmitted data with the predicted data of the LSTM and tries to minimize the error by using loss function like Mean Square Error (MSE). MSE can be expressed as:

$$MSE = \frac{\sum_{i=1}^{N}(y_i - y_i')^2}{N} \tag{4.3}$$

4.2.6 Simulation Results

This section presents the simulation results with various multipath channels with different path delays. Figure shows the training progress of the LSTM model and shows that the training accuracy 98% is achieved just after 10-15 epochs. However, LSTM converges nearly after 100 epochs in case of second multipath channel. The BER plots are shown in the Figure 4.4, Figure 4.5 and Figure 4.6 respectively. These plots indicate that the LSTM performs better than the LS and is equally good as MMSE. Simulation was run for three different Rayleigh channels. In the first simulation the channel is based on Rayleigh Narrowband fading channel. Channel consists of twenty multi paths and is defined as

$$H(z) = \sum_{n=0}^{19} a_n z^{-n} \tag{4.4}$$

Where the values of a_n are complex and selected randomly. The results are depicted in Figure 4.4. The results show three plots. It is clear from the graph that the MMSE and LSTM are equally performing on low SNR's. LS is slightly having higher BER at low SNR's. At higher SNR the performance of LSTM is again comparable to MMSE. The LS has slightly lower BER performance when compared to LSTM and MMSE.

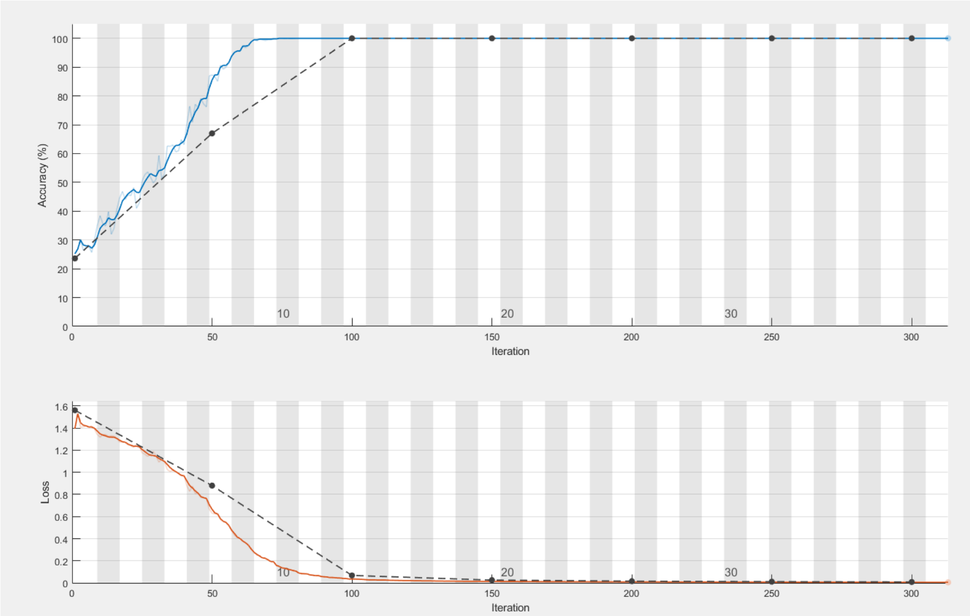

Figure 4.3 The training progress of LSTM model

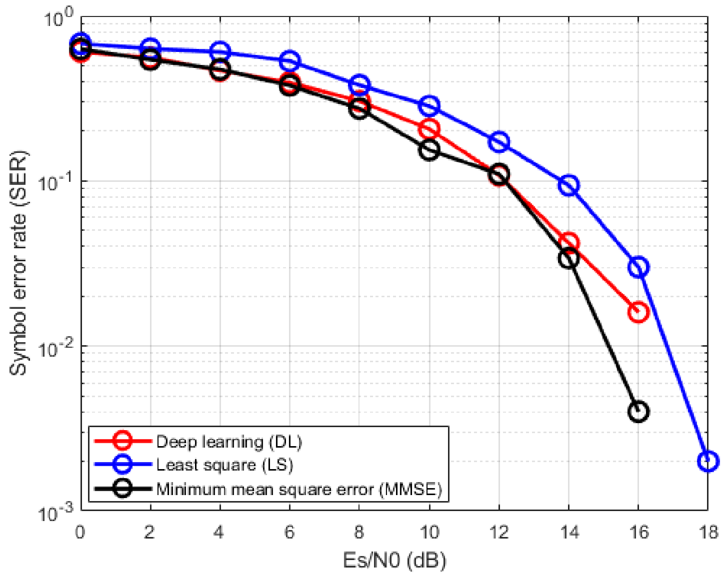

Figure 4.4 The BER plot of first channel

The channel used for the second simulation is taken from [43]. This channel can be given as

$$c = [1 - 0.3434j \quad 0.5 + 0.2912j] \tag{4.5}$$

$$H(z) = c_1 X(z) + c_2 z^{-1} X(z) \tag{4.6}$$

The results in Figure 4.5 are in line with the results depicted in the Figure 4.4. LS is slightly lower in the performance when compared with the LSTM and MMSE channel equalizers. The results show that LSTM and the LS are very close in the performance terms.

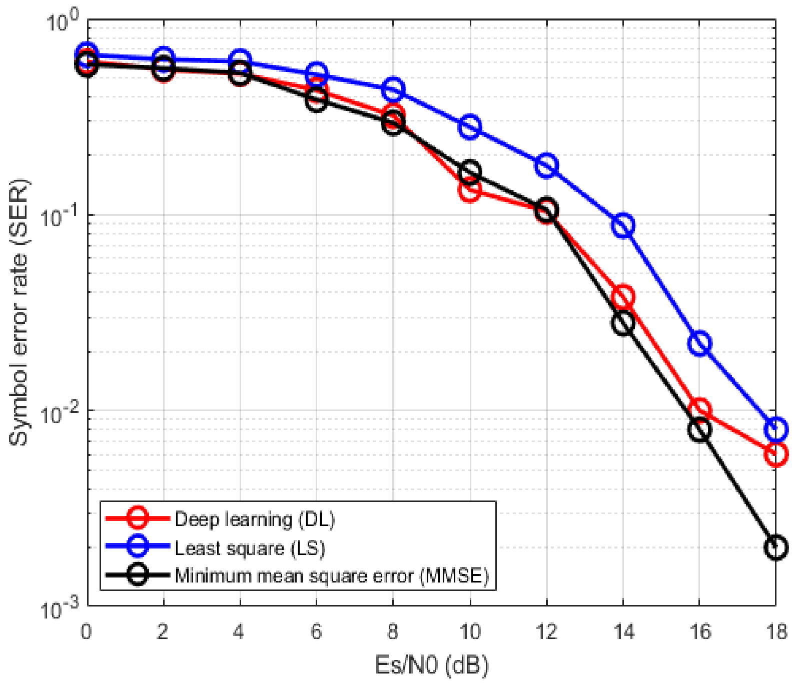

Figure 4.5 The BER plot of second channel

In Figure 4.6 the results are obtained by using another channel from [43] and is a non-minimum phase channel. A non-minimum phase channel may rotate the symbols more than 90 degrees thus severely degrading the symbols. The channel has three coefficients and can be given as

$$c = [0.34 - 0.27j \;\; 0.87 + 0.43j \quad 0.3 - 0.21j] \tag{4.7}$$

$$H(z) = c_1 X(z) + c_2 z^{-1} X(z) + c_3 z^{-2} X(z) \tag{4.8}$$

Again the results are very much in line with the results shown in the case of previous two simulations depicted in Figure 4.4 and Figure 4.5.

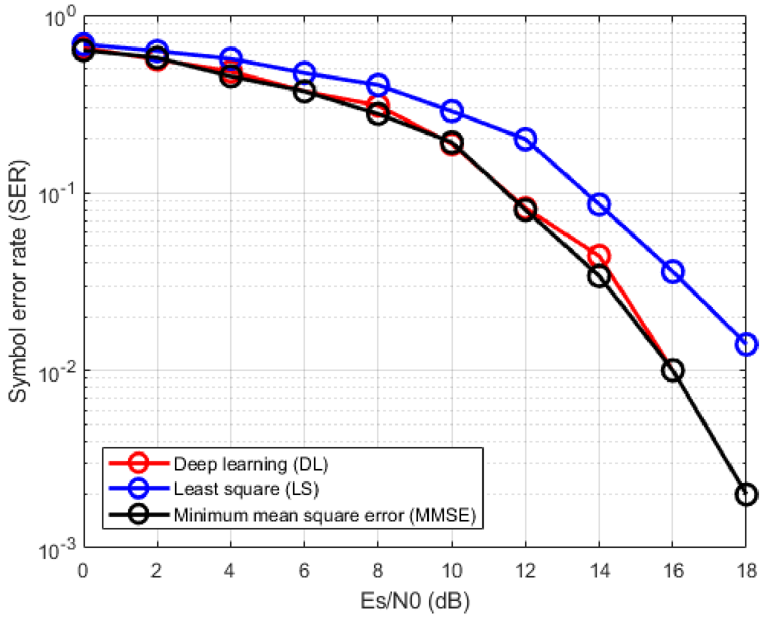

Figure 4.6 The BER plot of third channel

The LS has slightly poor BER performance when compared with the LSTM and MMSE. However, the LSTM and MMSE bear almost similar performance. Thus the ML methods have potential for usage as a channel estimation method in communication systems.

4.3 Summary

In this chapter, we have performed the pilot-aided channel estimation in OFDM systems using the LSTM model. A Model is first trained offline using simulation data, and then deployed online for recovering the actual symbols. It has shown promising results that are comparable to the conventional techniques like LS and MMSE. The simulations have confirmed that DL techniques like LSTM can perform equally good.

CHAPTER 5

CONCLUSIONS AND FUTURE WORK

This chapter summarizes the whole work reported in this thesis. It also specifies the research activities and concludes the work. This chapter also highlights future directions that can extend this work.

5.1 Thesis Summary

Communication system is an ever evolving, well established field of research and has shown major advances in signal estimation, equalization and other fields such as channel coding etc. Channel equalization being very critical for achieving high data rates and improved spectral efficiency, has been achieved using traditional theory of least squares estimation and minimum mean squares estimation techniques such as LMS, NLMS, RLS and Kalman filtering etc. Use of NN and SVM based channel equalization methods are currently under the research and are proving to be performing better than the conventional methods mentioned above.

In this thesis we have addressed the application of information theory based methods for the channel equalization comprising of Neural Networks and SVM techniques. It revealed that the methods used in traditional communication systems are difficult to understand and implement as compared to the ANN based methods. The channel equalization when treated as a classification problem using ANN techniques resulted in the simpler receiver structures especially in the case of OFDM. The results achieved are also found to be improved in terms of BER. Another advantage with the use of ANN based methods is that this has resulted in the relatively simpler way to understand the communication systems and many of the computer scientists who are not well versed with the communication systems theories can also attempt developing better communication systems by using their computer science and software development skills. Our research activities are summarized as follows.

1. We have presented a detailed literature review of various NNs (MLP, RBF, FLANN, SVM and LSTM) and discussed their applicability in wireless communications, especially in the domain of channel estimation and equalization. The literature review also covers the background communication systems theories related to single-carrier communication systems and the multi carrier OFDM based communication systems. The respective channel equalization techniques used for both the single-carrier and multi carrier OFDM have also been reviewed in detail.
2. We have developed a comprehensive simulation framework for carrying out the simulations where the basic blocks of communication system like transmitter, receiver and multipath Rayleigh channel have been developed. These blocks can always be replaced to add new functionalities such as changing the modulation scheme can be done by adding a new block of code.
3. The implementation of NN based equalizers has been done. Many of the NN methods such as RBF, FLANN, LSTM, SVM, etc have been implemented in MATLAB®. Their performances have also been compared. BER and constellation plots have validated the performance in simulation.
4. Training of NNs requires powerful hardware resources to accelerate this process. For this purpose, we have used a high-end GPU (RTX 2080) which uses CUDA toolkit and provides a very fast NN training capability.

5. Two major learning algorithms used in this work were LMS and RLS and it was observed that the convergence performance of the RLS is much better than the LMS algorithm.
6. Channel equalization in multi-carrier systems i.e. OFDM has been carried out and the performance of LSTM has been compared with the LS and MMSE methods. It has been found that the performance of the LSTM and MMSE is better than the performance of LS base channel equalizers.

5.2 Future Work

This thesis has laid the foundations for the further research work in this field. The fundamental methods used have been critically reviewed. Performance of the key algorithm has been compared. This work can be extended in many ways. Following is the list of possible emerging research areas.

1. Computational complexity analysis and computing platform optimizations of the algorithms is mandatory for efficient implementation on Hardware platforms such as ARM processors and FPGA and GPU’S. In this work the preliminary computational complexity analysis has been worked out. However this can be further extended when the implementation of these algorithms will be carried out on the FPGA’s or when optimized for the implementation on the microcontrollers and DSP processors.
2. Two dimensional treatment of the received signal which is similar to time frequency analysis where a number of frames are gathered and then processed as a block. This will enable the use of advanced neural network methods such as CNN, DNN and RNN methods. Existing frameworks such as AlexNet etc. may also be used.
3. Use of channel coding, MMSE and Viterbi for performance evaluation of the ANN is also recommended as future work.

4. Offline training of channel equalizer weights is another interesting research area where this work can be extended. This will be especially useful when a computationally less powerful platform is to be utilized.
5. NN based Channel estimators for the stochastic models of the various channels as defined in the literature such as COST-231 [6], Watterson model [103] for HF communication etc.
6. Currently, the performance evaluation is performed using QPSK modulation. The performance evaluation using higher order constellations such as 16QAM, 64QAM, 8PSK etc may also be carried out in the future.
7. Validation by developing hardware may be carried out.

REFERENCES

[1] A. Goldsmith, *Wireless communications*: Cambridge university press, 2005.

[2] C. E. Shannon, "A mathematical theory of communication," *The Bell system technical journal,* vol. 27, pp. 379-423, 1948.

[3] A. Petropulu, R. Zhang, and R. Lin, "Blind OFDM channel estimation through simple linear precoding," *IEEE Transactions on Wireless Communications,* vol. 3, pp. 647-655, 2004.

[4] H. H. Zeng and L. Tong, "Blind channel estimation using the second-order statistics: algorithms," *IEEE Transactions on Signal Processing,* vol. 45, pp. 1919-1930, 1997.

[5] E. D. Carvalho and D. T. M. Slock, "Cramer-Rao bounds for semi-blind, blind and training sequence based channel estimation," in *First IEEE Signal Processing Workshop on Signal Processing Advances in Wireless Communications*, 1997, pp. 129-132.

[6] T. S. Rappaport, *Wireless communications: principles and practice* vol. 2: prentice hall PTR New Jersey, 1996.

[7] S. Haykin, *Neural networks: a comprehensive foundation*: Prentice-Hall, Inc., 2007.

[8] G. J. Gibson, S. Siu, and C. F. Cowan, "The application of nonlinear structures to the reconstruction of binary signals," *IEEE Transactions on Signal processing,* vol. 39, pp. 1877-1884, 1991.

[9] S. Chen, G. Gibson, C. Cowan, and P. Grant, "Reconstruction of binary signals using an adaptive radial-basis-function equalizer," *Signal Processing,* vol. 22, pp. 77-93, 1991.

[10] G. J. Gibson, S. Siu, and C. F. Cowan, "Application of multilayer perceptrons as adaptive channel equalisers," in *Adaptive Systems in Control and Signal Processing 1989*, ed: Elsevier, 1990, pp. 573-578.

[11] T. O'Shea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Transactions on Cognitive Communications and Networking,* vol. 3, pp. 563-575, 2017.

[12] T. Wang, C.-K. Wen, H. Wang, F. Gao, T. Jiang, and S. Jin, "Deep learning for wireless physical layer: Opportunities and challenges," *China Communications,* vol. 14, pp. 92-111, 2017.

[13] C. Zhang, P. Patras, and H. Haddadi, "Deep learning in mobile and wireless networking: A survey," *IEEE Communications Surveys & Tutorials,* vol. 21, pp. 2224-2287, 2019.

[14] G. Gui, H. Huang, Y. Song, and H. Sari, "Deep learning for an effective nonorthogonal multiple access scheme," *IEEE Transactions on Vehicular Technology,* vol. 67, pp. 8440-8450, 2018.

[15] Y. Yang, F. Gao, X. Ma, and S. Zhang, "Deep Learning-Based Channel Estimation for Doubly Selective Fading Channels," *IEEE Access,* vol. 7, pp. 36579-36589, 2019.

[16] Y. Yang, F. Gao, G. Y. Li, and M. Jian, "Deep Learning-Based Downlink Channel Prediction for FDD Massive MIMO System," *IEEE Communications Letters,* vol. 23, pp. 1994-1998, 2019.

[17] Q. Bai, J. Wang, Y. Zhang, and J. Song, "Deep learning based channel estimation algorithm over time selective fading channels," *IEEE Transactions on Cognitive Communications and Networking,* 2019.

[18] M. Ibnkahla, "Applications of neural networks to digital communications–a survey," *Signal processing,* vol. 80, pp. 1185-1215, 2000.

[19] A. Zappone, M. Di Renzo, and M. Debbah, "Wireless networks design in the era of deep learning: Model-based, AI-based, or both?," *IEEE Transactions on Communications,* vol. 67, pp. 7331-7376, 2019.

[20] E. E. Azzouz and A. K. Nandi, "Modulation recognition using artificial neural networks," in *Automatic Modulation Recognition of Communication Signals*, ed: Springer, 1996, pp. 132-176.

[21] T. J. O'Shea, T. Erpek, and T. C. Clancy, "Deep learning based MIMO communications," *arXiv preprint arXiv:1707.07980,* 2017.

[22] T. J. O'Shea, T. Roy, and T. C. Clancy, "Over-the-air deep learning based radio signal classification," *IEEE Journal of Selected Topics in Signal Processing,* vol. 12, pp. 168-179, 2018.

[23] T. J. O'Shea, J. Corgan, and T. C. Clancy, "Convolutional radio modulation recognition networks," in *International conference on engineering applications of neural networks,* 2016, pp. 213-226.

[24] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *nature,* vol. 323, pp. 533-536, 1986.

[25] G. J. Gibson, S. Siu, and C. Cowen, "Multilayer perceptron structures applied to adaptive equalisers for data communications," in *International Conference on Acoustics, Speech, and Signal Processing*, 1989, pp. 1183-1186.

[26] S. Chen, G. Gibson, C. Cowan, and P. Grant, "Adaptive equalization of finite non-linear channels using multilayer perceptrons," *Signal processing,* vol. 20, pp. 107-119, 1990.

[27] D. P. Mandic, "A generalized normalized gradient descent algorithm," *IEEE signal processing letters,* vol. 11, pp. 115-118, 2004.

[28] S. Siu, G. Gibson, and C. Cowan, "Decision feedback equalisation using neural network structures and performance comparison with standard architecture," *IEE Proceedings I-Communications, Speech and Vision,* vol. 137, pp. 221-225, 1990.

[29] A. Sarangi, S. Priyadarshini, and S. K. Sarangi, "A MLP equalizer trained by variable step size firefly algorithm for channel equalization," in *2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES)*, 2016, pp. 1-5.

[30] X. Lyu, W. Feng, R. Shi, Y. Pei, and N. Ge, "Artificial neural network-based nonlinear channel equalization: A soft-output perspective," in *2015 22nd International Conference on Telecommunications (ICT)*, 2015, pp. 243-248.

[31] A. Zerguine, A. Shafi, and M. Bettayeb, "Multilayer perceptron-based DFE with lattice structure," *IEEE transactions on neural networks,* vol. 12, pp. 532-545, 2001.

[32] G. Kaur and G. Kaur, "Various Types of Functional Link Artificial Neural Network Based Nonlinear Equalizers used in CO-OFDM System for Nonlinearities Mitigation," *SSRG International Journal of Electronics and Communication Engineering (SSRG-IJECE),* vol. 5, 2018.

[33] J. C. Patra, W. C. Chin, P. K. Meher, and G. Chakraborty, "Legendre-FLANN-based nonlinear channel equalization in wireless communication system," in *2008 IEEE International Conference on Systems, Man and Cybernetics*, 2008, pp. 1826-1831.

[34] Y. Tan, J. Wang, and J. M. Zurada, "Nonlinear blind source separation using a radial basis function network," *IEEE transactions on neural networks,* vol. 12, pp. 124-134, 2001.

[35] A. Uncini and F. Piazza, "Blind signal processing by complex domain adaptive spline neural networks," *IEEE transactions on neural networks*, vol. 14, pp. 399-412, 2003.

[36] N. Xie and H. Leung, "Blind equalization using a predictive radial basis function neural network," *IEEE transactions on neural networks,* vol. 16, pp. 709-720, 2005.

[37] S. Chen, B. Mulgrew, and P. M. Grant, "A clustering technique for digital communications channel equalization using radial basis function networks," *IEEE Transactions on neural networks,* vol. 4, pp. 570-590, 1993.

[38] P. C. Kumar, P. Saratchandran, and N. Sundararajan, "Minimal radial basis function neural networks for nonlinear channel equalisation," *IEE Proceedings-Vision, Image and Signal Processing,* vol. 147, pp. 428-435, 2000.

[39] J. Satapathy, K. Subhashini, and G. L. Manohar, "A highly efficient channel equalizer for digital communication system in neural network paradigm," in *2009 Innovative Technologies in Intelligent Systems and Industrial Applications,* 2009, pp. 11-16.

[40] S. Chen, B. Mulgrew, and S. McLaughlin, "Adaptive Bayesian decision feedback equaliser based on a radial basis function network," in *SUPERCOMM/ICC'92 Discovering a New World of Communications,* 1992, pp. 1267-1271.

[41] S. Chen, B. Mulgrew, and S. McLaughlin, "Adaptive Bayesian equalizer with decision feedback," *IEEE Transactions on Signal Processing,* vol. 41, pp. 2918-2927, 1993.

[42] S. Chen, S. McLaughlin, and B. Mulgrew, "Complex-valued radial basis function network, Part II: Application to digital communications channel equalisation," *Signal processing,* vol. 36, pp. 175-188, 1994.

[43] I. Cha and S. A. Kassam, "Channel equalization using adaptive complex radial basis function networks," *IEEE Journal on Selected Areas in communications,* vol. 13, pp. 122-131, 1995.

[44] S. S. Ranhotra, A. Kumar, M. Magarini, and A. Mishra, "Performance comparison of blind and non-blind channel equalizers using artificial neural networks," in *2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN),* 2017, pp. 243-248.

[45] G. Panda and D. P. Das, "Functional link artificial neural network for active control of nonlinear noise processes," in *International Workshop on Acoustic Echo and Noise Control,* 2003, pp. 163-6.

[46] J. C. Patra and R. N. Pal, "A functional link artificial neural network for adaptive channel equalization," *Signal Processing,* vol. 43, pp. 181-195, 1995.

[47] D. M. Sahoo and S. Chakraverty, "Functional link neural network approach to solve structural system identification problems," *Neural Computing and Applications,* vol. 30, pp. 3327-3338, 2018.

[48] T. J. Rivlin, *Chebyshev polynomials*: Courier Dover Publications, 2020.

[49] T.-T. Lee and J.-T. Jeng, "The Chebyshev-polynomials-based unified model neural networks for function approximation," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics),* vol. 28, pp. 925-935, 1998.

[50] J. Liu, K. Mei, X. Zhang, D. Ma, and J. Wei, "Online extreme learning machine-based channel estimation and equalization for OFDM systems," *IEEE Communications Letters,* vol. 23, pp. 1276-1279, 2019.

[51] J. C. Patra and A. C. Kot, "Nonlinear dynamic system identification using Chebyshev functional link artificial neural networks," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics),* vol. 32, pp. 505-511, 2002.

[52] J. C. Patra, W. B. Poh, N. S. Chaudhari, and A. Das, "Nonlinear channel equalization with QAM signal using Chebyshev artificial neural network," in *Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005.*, 2005, pp. 3214-3219.

[53] G. I. Kechriotis and E. S. Manolakos, "Using neural networks for nonlinear and chaotic signal processing," in *1993 IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1993, pp. 465-468.

[54] A. Waibel, T. Hanazawa, G. Hinton, K. Shikano, and K. J. Lang, "Phoneme recognition using time-delay neural networks," *IEEE transactions on acoustics, speech, and signal processing,* vol. 37, pp. 328-339, 1989.

[55] G. Kechriotis, E. Zervas, and E. Manolakos, "Using recurrent neural networks for blind equalization of linear and nonlinear communications channels," in *MILCOM 92 Conference Record*, 1992, pp. 784-788.

[56] M. J. Bradley and P. Mars, "Application of recurrent neural networks to communication channel equalization," in *1995 International Conference on Acoustics, Speech, and Signal Processing*, 1995, pp. 3399-3402.

[57] J. D. Ortiz-Fuentes and M. L. Forcada, "A comparison between recurrent neural network architectures for digital equalization," in *1997 IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1997, pp. 3281-3284.

[58] S. Hu, Y. Pei, P. P. Liang, and Y.-C. Liang, "Robust modulation classification under uncertain noise condition using recurrent neural network," in *2018 IEEE Global Communications Conference (GLOBECOM)*, 2018, pp. 1-7.

[59] A. D. Back and A. C. Tsoi, "FIR and IIR synapses, a new neural network architecture for time series modeling," *Neural computation*, vol. 3, pp. 375-385, 1991.

[60] G. Kechriotis, E. Zervas, and E. S. Manolakos, "Using recurrent neural networks for adaptive communication channel equalization," *IEEE transactions on Neural Networks*, vol. 5, pp. 267-278, 1994.

[61] C.-F. Teng, H.-M. Ou, and A.-Y. Wu, "Neural network-based equalizer by utilizing coding gain in advance," *arXiv preprint arXiv:1907.04980*, 2019.

[62] S. Hochreiter, "The vanishing gradient problem during learning recurrent neural nets and problem solutions," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 6, pp. 107-116, 1998.

[63] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, pp. 1735-1780, 1997.

[64] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, pp. 273-297, 1995.

[65] V. Vapnik, S. E. Golowich, and A. J. Smola, "Support vector method for function approximation, regression estimation and signal processing," in *Advances in neural information processing systems*, 1997, pp. 281-287.

[66] S. R. Gunn, "Support vector machines for classification and regression," *ISIS technical report*, vol. 14, pp. 5-16, 1998.

[67] S. Chen, S. Gunn, and C. Harris, "Decision feedback equaliser design using support vector machines," *IEE Proceedings-Vision, Image and Signal Processing*, vol. 147, pp. 213-219, 2000.

[68] D. J. Sebald and J. A. Bucklew, "Support vector machine techniques for nonlinear equalization," *IEEE transactions on signal processing*, vol. 48, pp. 3217-3226, 2000.

[69] Y. Lecun, "PhD thesis: Modeles connexionnistes de l'apprentissage (connectionist learning models)," 1987.

[70] S. Dörner, S. Cammerer, J. Hoydis, and S. ten Brink, "Deep learning based communication over the air," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, pp. 132-143, 2017.

[71] R. v. Nee and R. Prasad, *OFDM for wireless multimedia communications*: Artech House, Inc., 2000.

[72] K. El Baamrani, A. A. Ouahman, and S. Allaki, "Rate adaptive resource allocation for OFDM downlink transmission," *AEU-International Journal of Electronics and Communications,* vol. 61, pp. 30-34, 2007.

[73] A. Jamalipour and S. Tekinay, "Next-generation broadband wireless networks and navigation services [Guest Editorial]," *IEEE Communications Magazine*, vol. 40, pp. 110-111, 2002.

[74] W. W. Lu, "Fourth-generation mobile initiatives and technologies [Guest Editorial]," *IEEE Communications Magazine,* vol. 40, pp. 104-105, 2002.

[75] R. Berezdivin, R. Breinig, and R. Topp, "Next-generation wireless communications concepts and technologies," *IEEE communications magazine,* vol. 40, pp. 108-116, 2002.

[76] H. Sampath, S. Talwar, J. Tellado, V. Erceg, and A. Paulraj, "A fourth-generation MIMO-OFDM broadband wireless system: design, performance, and field trial results," *IEEE Communications Magazine,* vol. 40, pp. 143-149, 2002.

[77] K. Fazel and S. Kaiser, *Multi-carrier and spread spectrum systems: from OFDM and MC-CDMA to LTE and WiMAX*: John Wiley & Sons, 2008.

[78] A. R. Bahai, B. R. Saltzberg, and M. Ergen, *Multi-carrier digital communications: theory and applications of OFDM*: Springer Science & Business Media, 2004.

[79] B. Farhang-Boroujeny, "OFDM versus filter bank multicarrier," *IEEE signal processing magazine,* vol. 28, pp. 92-112, 2011.

[80] M. Speth, S. A. Fechtel, G. Fock, and H. Meyr, "Optimum receiver design for wireless broad-band systems using OFDM. I," *IEEE Transactions on Communications,* vol. 47, pp. 1668-1677, 1999.

[81] H. Steendam and M. Moeneclaey, "Analysis and optimization of the performance of OFDM on frequency-selective time-selective fading channels," *IEEE Transactions on Communications,* vol. 47, pp. 1811-1819, 1999.

[82] F. Tufvesson and T. Maseng, "Pilot assisted channel estimation for OFDM in mobile cellular systems," in *1997 IEEE 47th Vehicular Technology Conference. Technology in Motion*, 1997, pp. 1639-1643.

[83] J. H. Manton, "Optimal training sequences and pilot tones for OFDM systems," *IEEE communications letters,* vol. 5, pp. 151-153, 2001.

[84] O. Edfors, M. Sandell, S. K. Wilson, and P. Borjesson, "On channel estimation in OFDM systems," in *IEEE VTC,* 1995, pp. 815-819.

[85] H. Ye, G. Y. Li, and B.-H. Juang, "Power of deep learning for channel estimation and signal detection in OFDM systems," *IEEE Wireless Communications Letters*, vol. 7, pp. 114-117, 2017.

[86] I. Barhumi, G. Leus, and M. Moonen, "Time-varying FIR equalization for doubly selective channels," *IEEE Transactions on Wireless communications*, vol. 4, pp. 202-214, 2005.

[87] M. Morelli and U. Mengali, "A comparison of pilot-aided channel estimation methods for OFDM systems," *IEEE Transactions on signal processing*, vol. 49, pp. 3065-3073, 2001.

[88] M.-H. Hsieh and C.-H. Wei, "Channel estimation for OFDM systems based on comb-type pilot arrangement in frequency selective fading channels," *IEEE Transactions on Consumer Electronics*, vol. 44, pp. 217-225, 1998.

[89] S. Colieri, M. Ergen, A. Puri, and A. Bahai, "A study of channel estimation in OFDM systems," in *Proceedings IEEE 56th vehicular technology conference*, 2002, pp. 894-898.

[90] M. M. Moustafa and S. H. El-Ramly, "Channel estimation and equalization using backpropagation neural networks in OFDM systems," in *2009 IFIP International Conference on Wireless and Optical Communications Networks*, 2009, pp. 1-4.

[91] M. Onder and A. Akan, "Pilot tone investigation for joint channel estimation, equalization, and demodulation based on neural networks," in *2015 9th International Conference on Electrical and Electronics Engineering (ELECO)*, 2015, pp. 749-752.

[92] M.-G. Garcia, J. L. Rojo-Álvarez, F. Alonso-Atienza, and M. Martínez-Ramón, "Support vector machines for robust channel estimation in OFDM," *IEEE Signal Processing Letters*, vol. 13, pp. 397-400, 2006.

[93] D. TseandP, "Viswanath, Fundamentals of Wireless Communication," ed: Cambridge, UK: Cambridge Univ. Press, 2005.

[94] D. Jianping, N. Sundararajan, and P. Saratchandran, "Communication channel equalization using complex-valued minimal radial basis function neural networks," *IEEE Transactions on neural networks*, vol. 13, pp. 687-696, 2002.

[95] S. S. Haykin and M. Moher, *Modern wireless communications*: Pearson Education India, 2011.

[96] A. Doukas and G. Kalivas, "Rician K factor estimation for wireless communication systems," in *2006 International*

Conference on Wireless and Mobile Communications (ICWMC'06), 2006, pp. 69-69.
[97] U. Madhow, *Fundamentals of digital communication*: Cambridge University Press, 2008.
[98] P. M. Shankar, *Introduction to wireless systems*: Wiley New York, 2002.
[99] M. F. Møller, "A scaled conjugate gradient algorithm for fast supervised learning," *Neural networks*, vol. 6, pp. 525-533, 1993.
[100] J. C. Patra, R. N. Pal, B. N. Chatterji, and G. Panda, "Identification of nonlinear dynamic systems using functional link artificial neural networks," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics),* vol. 29, pp. 254-262, 1999.
[101] Y. Li, L. J. Cimini, and N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels," *IEEE Transactions on communications,* vol. 46, pp. 902-915, 1998.
[102] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural networks,* vol. 61, pp. 85-117, 2015.
[103] C. Watterson, J. Juroshek, and W. Bensema, "Experimental confirmation of an HF channel model," *IEEE Transactions on communication technology*, vol. 18, pp. 792-803, 1970.