Channel Estimation in MIMO OFDM System in 5G using Deep Learning

Project report submitted in partial fulfillment
of the requirements for the degree of

Bachelor of Technology
in
Electronics and Communication Engineering

by

Aditya Raj - 17UEC143
Yash Maheshwari - 17UEC146
Aniket Tiwary - 17UEC020

Under Guidance of
Prof. Purnendu Karmakar

Department of Electronics and Communication Engineering
The LNM Institute of Information Technology, Jaipur

December 2020

The LNM Institute of Information Technology
Jaipur, India

CERTIFICATE

This is to certify that the project entitled "Channel Estimation in MIMO OFDM System in 5G using Deep Learning" , submitted by Aditya Raj (17UEC143), Yash Maheshwari (17UEC146) and Aniket Tiwary (17UEC020) in partial fulfillment of the requirement of degree in Bachelor of Technology (B. Tech), is a bonafide record of work carried out by them at the Department of Electronics and Communication Engineering, The LNM Institute of Information Technology, Jaipur, (Rajasthan) India, during the academic session 2020-2021 under my supervision and guidance and the same has not been submitted elsewhere for award of any other degree. In my/our opinion, this thesis is of standard required for the award of the degree of Bachelor of Technology (B. Tech).

Date

Adviser: Prof. Purnendu Karmakar

Dedicated to our Family and Friends

Acknowledgments

The authors would like to acknowledge their Mentor **Professor Purnendu Karmakar** for providing the opportunities and resources that contributed heavily in completion of this research paper. Without the platform and exposure provided by the **Professor and Department of Electronics and Communication Engineering, The LNM institute of Information and Technology, Jaipur** this research paper would not have achieved its full potential. Lastly, Congratulations to the three authors for their efforts and cooperation.

Abstract

In this paper, using the wireless communication configuration of MIMO systems we try to estimate the channel for 5G technology. As the 5th Generation is used, its characteristic nature implores testing the channel estimation in Orthogonal Frequency Division Multiplexing (OFDM) system. Further, We investigate and visualise the step by step outputs of each block in the long chain of Modulation and demodulation for channel estimation. By employing the Least Squares Channel Estimator(LSE) and Minimum Mean-Square Error Estimator(MMSE) ,we try to provide meaningful results based on their comparisons. Further we generated a randomly designed input data that is taken and sent into the deep learning model. The model comprises of multiple layers for the purpose of accurate classification. This model is further fine tuned using training options and its validation accuracy is monitored. This leads to comparable results between the three techniques employed .Wireless technology is the modern method of quick and reliable telecommunication which when recruited under 5G has advanced the field to new horizons.

Contents

Chapter *1*

Introduction

This paper seeks to evaluate the effectiveness and accuracy of 5G channel estimation in a Multiple Input and Multiple Outputs (MIMO) system incorporating the Orthogonal Frequency-Division Multiplexing(OFDM) technique for minimisation of error and how to transmit and receive the OFDM signal through wireless communication in a very effective manner. In recent years, optimization of the wireless communication system has become critical with the rapid growth of mobile communication services and emerging broadband mobile Internet access services. So we need to come up with a high bandwidth and better data rate technology to overcome our need. We are using MIMO system as multiple antennas for both transmitters and receivers vastly improve communication performance, because of the increase in the number of antennas as compared to SISO(Single Input Single Output) it provides us with multiple degrees of freedom which is also referred to as spatial degree of freedom. This spatial degree of freedom could be used can either be used for diversity or multiplexing or both depending on one's need.

OFDM is one such technology which helps us achieve that because of plenty of reasons as : (i)As the symbol rate is inversely proportional to symbol duration, as a result every sub-carrier has comparatively long symbols. Robustness of long symbols against multi-path fading, as noticed in wireless channels, (ii) Due to frequency selective nature of the channel when the channel is in deep fade (i.e. the received energy on this carrier is very low), rather than the whole bit stream only the exact bit value is lost which is also a plus, (iii) As the system has more than one carrier therefore it allows more than one user benefits at the same time by assigning multiple sub-carrier . A figure that demonstrates the fundamental block of an OFDM system is shown below:

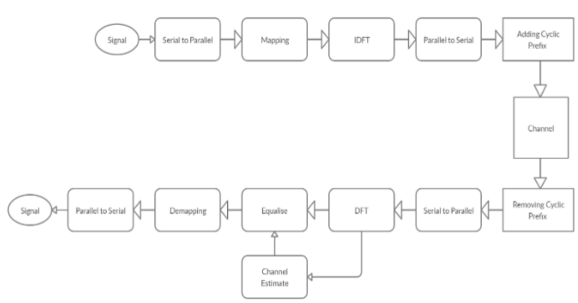

Figure 1.1 Block Diagram depiction of an OFDM model

Implementation of the OFDM system is achieved using MATLAB.The channel estimation techniques implemented at the re-ceiver end are MMSE, LSE and using Deep Learning method.

The Wireless Communication System has grown it's patterns, concepts and methods since we started using DL(Deep Learning Technology). Experimental and theoretical improvements can be seen when we use DL in various processes like channel estimation, signal detection, management of resources and compression of feedback. Out of all the benefits Deep Learning provides in wireless communication systems the process studied and refined most is channel estimation. Deep Learning methods can be applied to study various characteristics of wireless communication systems and tackle the nonlinear interference and distortion for OFDM(Orthogonal Frequency Division Multiplexing) systems. DL could be used in MIMO systems to estimate channel estimation and DoA(Direction of Arrival) estimation. We could establish a novel end-to-end DNNs(Deep Neural Networks) architecture which replaces all the modules at the transmitter and at the receiver in place of boosting only certain modules.

We should also note that despite of wide success of DL in wireless communication system, Wireless communication systems which are DNN embedded are regarded as a black box when it comes to communication of signal i.e, signal transmission/reception. Deep Learning methods does not provide analytical interpretation which confirms the disadvantages and advantages of Deep learning method we have only experimental and numerical evaluations.

Chapter *2*

Literature Survey

2.1 MIMO

Multiple-input multiple-output (MIMO) is a dynamic technique that has become the cornerstone of advanced generations of communication systems. MIMO is specially relevant in the field of practical wireless communication setups. The channel utilization peaks as well as the robustness of the channel is enhanced on MIMO system utilisation.An enlarged link capacity with improved stability to transmit data combined with improved spectral efficiency are prime characteristics of the system configuration. Owning to this it holds an advantage in controlled situations over other system configurations like Single input single output(SISO), single input multiple output (SIMO) and multiple input single outputs (MISO).

Further, Battikh, et al.[3] have showcased that intuitively MIMO outperforms MISO in terms of throughput and is able to accommodate a higher number of users which although leads to more divided cell capacity, provides a higher throughput in high load in the outer rings. They are enabled for wireless communication by incorporating a transmitting and a receiving antennae with a wireless channel in between for seamless transmission. The multi-path technology is used widely in WiFi, Wi-Max,LTE,etc.

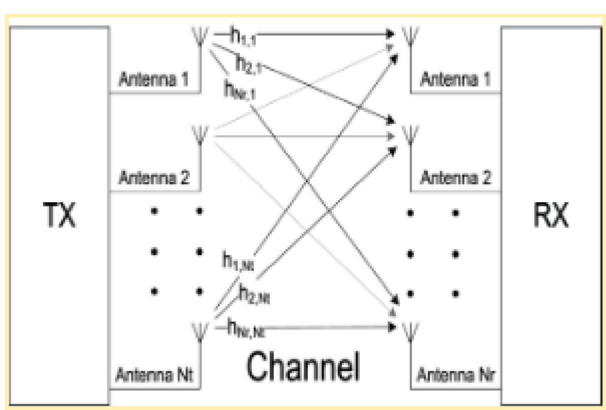

Figure 2.1 MIMO system [4]

The space time diversity is the most widely availed in case of MIMO systems. The Alamouti space time block code [11] is the most simplified version from the whole family of orthogonal space time codes. The capacity of MIMO channels by the Shannon formulae [5] is given by -

$$C = \max_{f(x)} K(X;Y) = \max_{f(x)} \int_{S_x, S_y} f(x,y) \log\left(\frac{f(x,y)}{f(x)f(y)}\right) \tag{2.1}$$

Here the variable values are as follows.The support for the random variable S_x is *X* and that for *Y* is S_y. *K(X;Y)*-mutual information factor of the channel.The probability density function(pdf) of *Y* is $f(y)$ and probability density function(pdf) of *X* is $f(x)$.Shannon showed that the channel capacity is equal to the mutual information of the channel maximized over all possible input distributions.

2.2 Massive MIMO channel

A single cell BS, when the input data is transmitted using a BS transmitter it goes through a number of stages of digital processing used in a typical digital communication system namely channel coding, data interleaving, modulation of signal, precoding of data signal in MIMO and filtering if needed. When the transmit MIMO signal leaves the transmitter and reaches the receiver at the receiver antenna then the processing done at the transmitter is reversed (like precoding, decoding, demodulation, space-time decoding in case of index modulation at the transmitter). This all processes are done for a MIMO channel but if the MIMO channel becomes massive as in our case when we talk about Massive MIMO, the two parameters we consider while designing Massive MIMO are : (a.) Multiplexing Gain, (b.) Diversity Gain.

Massive multiplexing gain is defined as the slope of the outage capacity of a channel fixed Frame Error Rate(FER) and given Signal to Power Ratio(SNR). If the MIMO transmitter - receiver are efficiently designed then massive multiplexing gain doubles for every 3dB increase in SNR for fixed FER. When we measure the negative asymptotic slope of FER vs SNR when both are given in log scales that gives us maximum diversity gain. When a efficient transmitter - receiver is designed over a MIMO channel the maximum diversity gain for a fixed FER when we increase the SNR 3dB the FER reduced by the factor of 2^{-NM} for a (N,M) MIMO transceiver.

To represent MIMO channel we use matrix generally defined as matrix H with N rows and M colums (NxM) where each and every element of the matrix h_{nm} is the defined SISO(single input - single output) channel between mth transmitter antenna and nth reciever antenna. Here a row say nth row symbolises the symbols which are recieved from the N-antenna arrays which are placed at the reciever, similarly each column or the mth column represents the mth antenna at the transmitter out of the M-antenna arrays. Characteristic of the individual SISO channel helps us to make the MIMO channel a zero mean and Gaussian random variable which is circularly symmetric in nature which has distributed power gains which are exponential and amplitudes which are Rayleigh distributed random variable. Individual channels correlation turns out to be a fuction of antenna arrays configuration done at the Base Station and scattering in the medium.

2.3 OFDM : Orthogonal Frequency Division Multiplexing

Orthogonal frequency division multiplexing is a technology developed on the backbone of Multiple antenna systems. The most primitive systems of single input single output were relatively simple to implement but had the problem of transmission like no other system before it. The single channel developed in these was prone to 'deep fade'. This was the phenomena of the channel degradation constant tending to a very high value. Dynamic environment conditions coupled with hurdles and obstacles led to this high unnatural behaviour which was quite common in the real world. Added to this Additive white Gaussian noise (AWGN) Noise, the single channel of communication between transmitter and receiver was rendered useless and no information transfer could take place between the two. This led to complete deterioration of the communication system.

To overcome this deep fade problem, MAS were designed. Diversity techniques were used to hedge the risk of the system falling prey to deep fade. Diverse Systems had Multiple transmitter and receiver antennae which led to creation of multiple Line of sight or Non line of sight channels between one node and another. Even if one these channels went into deep fade the other channels ensured that the information transfer took place without substantial loss. A new problem arose with the usage of MAS.

There was the unwanted occurrence of Inter channel Interference between the multiple channels created due to this diversity techniques.The information stream embedded in one carrier wave would constructively or destructively interact with the information in another channel carrier wave leading to corruption of data. So to tackle this issue of Inter Channel Interference, Index modulation techniques were utilised. In these, Different physical or virtual indices of the antennae were used as the source of information coupling with the actual information.

One such method was spatial modulation which used the physical location of the antennae in the 3D space as the source of information. But in these systems the symbol time was often less than the delay spread. This led to Inter Symbol Interference which caused heavy distortions in the data encapsulated in the symbol. As in these single carrier systems, the Symbol time was inversely proportional to the bandwidth of the carrier, A pragmatic solution for avoiding ISI to divide the bandwidth into smaller sub bands to increase the symbol time was adopted. This led to the symbol time exceeding the lag in the first and the last multipath component. As multiple carriers in this setup would need an extensive network of crystals oscillating at a different frequency. This would be cost heavy as well as have high implementation complexity.

To Overcome this a completely new approach was taken up. Rather than transmitting symbols in frequency domain An Inverse fast Fourier transform (IFFT) of input signal is done at the transmitter end to send information in time domain. A Fast Fourier transform (FFT) of received signal is done at the receiver end. This is done to get back the symbol in the frequency domain and do the proper analysis.

Further to avoid the problem of Interblock Interference, Cyclic prefix is added to the start of each sub block to act as a buffer. This leads to loss of spectral efficiency as redundant data is transmitted. This IFFT-FFT of information at the transmitter and receiver end is the basis of OFDM systems. The technology of OFDm had to wait a long time before it got practically feasible to implement. The further technologies developed could be seen as a variation of this basic OFDM.

Our next goal in communication and data transmission is to meet high data rates and achieve efficient use of limited spectrum. The idea is to allow the exploitation of locally available frequency bands, on a temporary basis but under a non interfering constraint.Most modern digital communication standards use Orthogonal Frequency Division Multiplexing (OFDM) as the air interface, because of its flexibility and robustness in frequency-selective channels. OFDM allows high- speed data transmission across a dispersive channel, so is used in many new and emerging high speed wired and wireless communications.

Orthogonal frequency division multiplexing (OFDM) transmission scheme is another type of a multi-channel system, which is similar to the FMT transmission scheme in the sense that it employs multiple sub-carriers [2].It does not use individual band-limited filters and oscillators for each sub-channel and furthermore, the spectra of sub-carriers are overlapped for bandwidth efficiency, unlike the FMT scheme where the wide-band is fully divided into N orthogonal narrow-band sub-channels. The multiple orthogonal sub-carrier signals, which are overlapped in spectrum, can be produced by generalizing the single-carrier Nyquist criterion in Equation into the multi-carrier criterion.

OFDM is symbol based, and can be viewed as a large number of low bit-rate carriers transmitting in parallel. Using synchronized time and frequency, these carriers transmit simultaneously, thus consequently forming a single block of spectrum and further ensures that the orthogonal nature of the structure is maintained.

The orthogonality property of OFDM signals can be better visualized by looking at its spectrum. In the frequency domain each OFDM sub-carrier has a sinc, $sin(x)/x$, frequency response as shown in the Fig.2.2

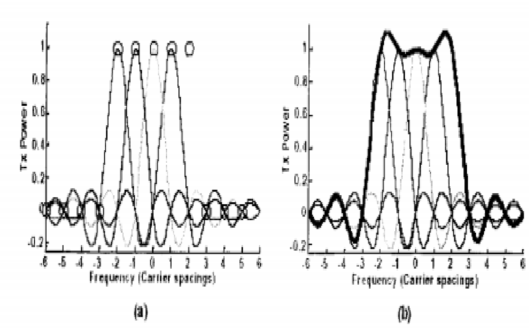

Figure 2.2 Frequency Response of the sub-carrier in a 5 Tone OFDM Signal [2]
(a)Shows the spectrum of each carrier and the discrete frequency samples seen by an OFDM receiver
(b)Shows the overall combined response of the 5 sub-carriers (thick black line).

These multiple carriers can be referred as sub-carriers as they form a single OFDM transmission. Also it is worth noting that the orthogonality between sub-carriers in OFDM systems allows inserting and extracting pilots without any sort of interference. Each sub-carrier is modulated with a typical digital modulation scheme (such as 64QAM, QPSK etc.) at low symbol rate. However, when we combine sub-carriers, it enables data rates similar to a typical single-carrier modulation schemes and that too within equivalent bandwidths.

Fig. 2.3 illustrates the main concepts of an OFDM signal and the inter-relationship between the time and frequency domains. In the frequency domain, multiple adjacent tones or sub-carriers are each independently modulated with complex data. In order to produce the OFDM symbol in the time-domain, an Inverse FFT (IFFT) transform is performed on the frequency-domain. Then in the time domain, guard intervals are inserted between each of the symbols to prevent inter-symbol interference (ISI) and inter-channel interference (ICI) at the receiver which are primarily caused by multi-path delay spread in the radio channel. The final OFDM burst signal can be created by concatenating multiple symbols. FFT is performed on the OFDM symbols to recover the original data bits at the receiver end.

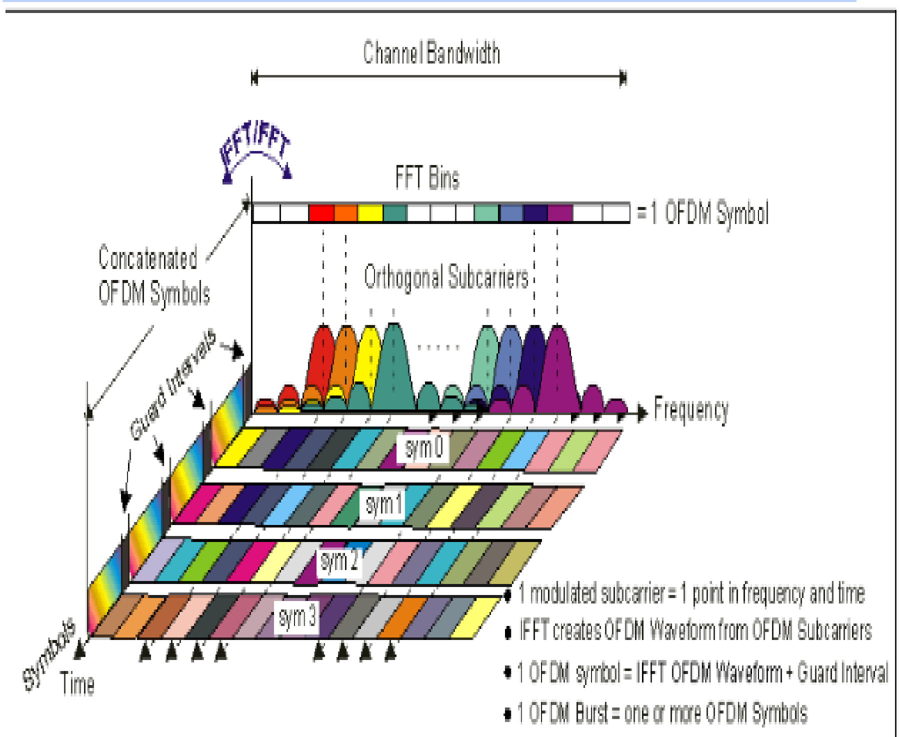

Figure 2.3 Frequency-Time Representative of an OFDM signal [10]

The relationship between all the carriers must be carefully controlled to maintain the orthogonality of the carriers and in order to generate OFDM successfully. For this very reason, based on the Input data and the modulation scheme used, OFDM is generated by firstly choosing the spectrum required, where each carrier to be produced is assigned some data to transmit. The input serial data stream is formatted into the word size required for transmission and then shifted into a parallel format. The data is then transmitted in parallel by assigning each data word to one carrier during transmission..

On each carrier, the data to be transmitted is mapped into a phase shift keying (PSK) format. Based on the modulation method, the data on each symbol is then mapped to a phase angle. The symbol rate for an OFDM signal is significantly lower than a single carrier transmission scheme for a given system bandwidth.

The system bandwidth for OFDM is broken up into N sub-carriers, resulting in a symbol rate that is N times lower than the single carrier transmission. This low symbol rate makes OFDM naturally resistant to effects of Inter-Symbol Interference (ISI) caused by multi-path propagation. As multi-path propagation is caused by the radio transmission signal reflecting off objects in the propagation environment, such as walls, buildings, mountains, etc. so these multiple signals arrive at the receiver at different times due to the transmission distances being different. This spreads the symbol boundaries causing energy leakage between them. On an OFDM signal, Inter-Symbol Interference effects can be further improved by the addition of a guard period to the start of each symbol which is basically a cyclic copy that extends the length of the symbol waveform. Also the guard period provides protection against time-offset errors in the receiver.

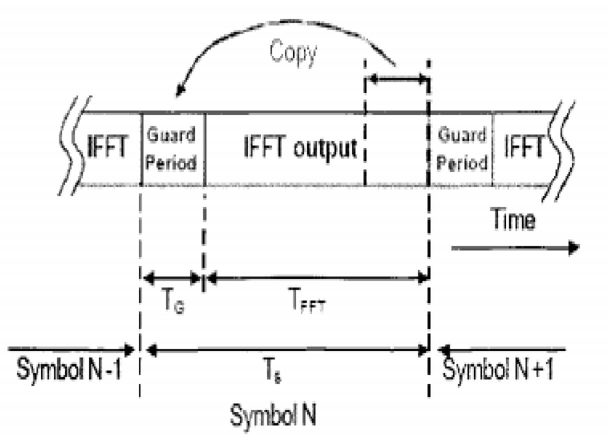

Figure 2.4 Addition of a guard period of an OFDM signal [10]

The total length of the symbol is Ts=Tg + Tfft, where Ts is the total length of the symbol in samples, Tg is the length of the guard period in samples, and Tfft is the size of the IFFT used to generate the OFDM signal. The receiver basically does the reverse operation to the transmitter to retrieve the information transmitted. The guard period/interval is removed and the FFT of each symbol is then taken to find the original transmitted spectrum. We evaluate the phase angle of each transmission carrier and which is then converted back to the data word by demodulating the received phase. The data words are then assembled back to the same word size as the original data.

We can use a simple analog based implementation to show the basic principles of generating an OFDM system. In this analog based OFDM system there are N sinusoidal input signals and each sub-carrier transmits one bit of information (N bits total) as indicated by its presence or absence in the output spectrum. The frequency of each sub-carrier is selected to form an orthogonal signal set which are also available at the receiver for signal recovery. We note that the output is updated at a periodic interval T (symbol period). To maintain orthogonality, T must be the reciprocal of the sub-carrier spacing. This is a result of the symbol time corresponding to the inverse of the carrier spacing. As far as the receiver is concerned each OFDM symbol transmitted for a fixed time (Tfft). This symbol time corresponds to the inverse of the sub-carrier spacing of 1/Tfft (Hz).

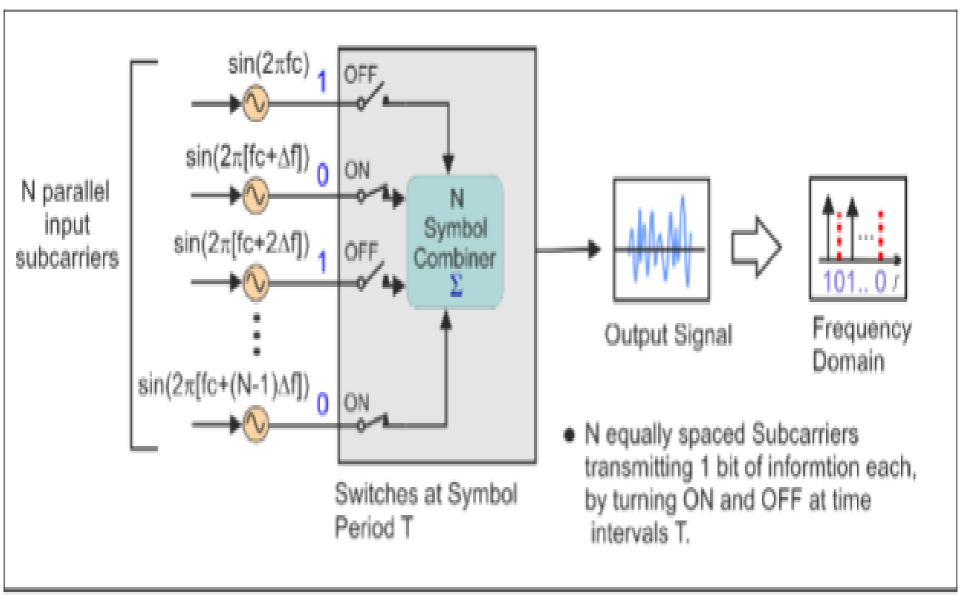

Figure 2.5 Simple OFDM Generation [10]

The concepts used in the simple analog OFDM implementation can be further extended to the digital domain by using a combination of Fast Fourier Transform (FFT) and Inverse Fast Fourier Transform (IFFT) digital signal processing. These transforms are important from the OFDM perspective because they can be viewed as mapping digitally modulated input data (data symbols) onto orthogonal sub-carriers. In principle, the IFFT takes frequency-domain input data (modulated sub-carriers represented by complex numbers) and converts it to the time-domain output data (analog OFDM symbol waveform).

In a digitally implemented OFDM system, the input bits are grouped and mapped to source data symbols that are a complex number representing the modulation constellation point (e.g., the BPSK or QAM symbols that would be present in a single sub-carrier system). These complex source symbols are treated by the transmitter as though they are in the frequency-domain and are the inputs to an IFFT block that transforms the data into the time-domain. The IFFT takes in N source symbols at a time (where N is the number of sub-carriers in the system). Each of these N input symbols has a period of T seconds (symbol period). We recall that the output of the IFFT is N orthogonal sinusoids with each having a different frequency, while the lowest frequency is DC.

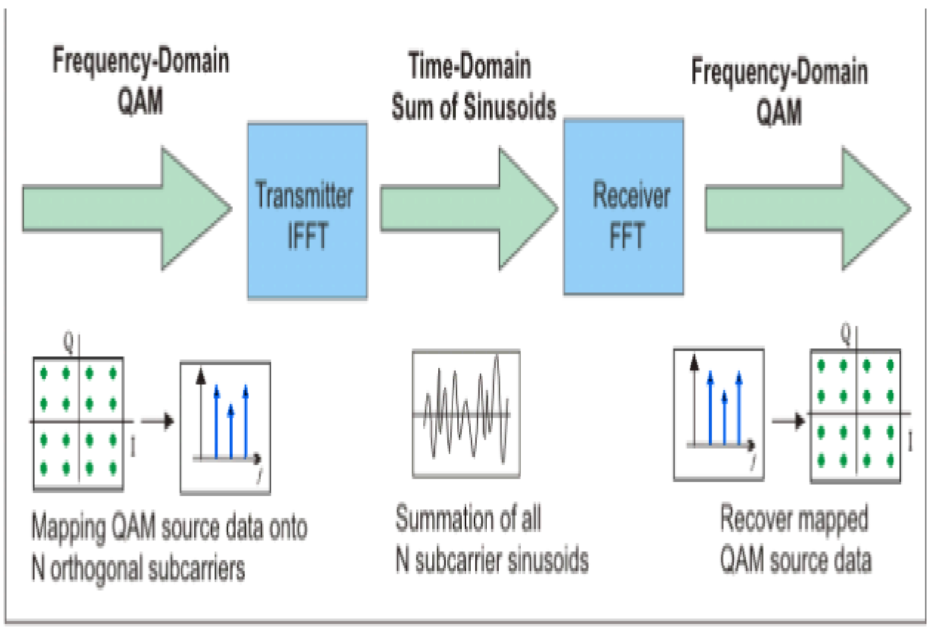

Figure 2.6 Simplified OFDM Signal Block Diagram [10]

Input symbols for any sub-carrier specify both the amplitude as well as the phase of the sinusoid as it represents the constellation points which are mapped and are complex values. We get the summation of all N sinusoids as IFFT output and the IFFT block provides a convenient and easy way to modulate data onto N orthogonal sub-carriers and finally a single OFDM symbol is formed from the IFFT block of N output samples.

So, we can infer that a OFDM carrier signal is the sum of one or more OFDM symbols. Inverse FFT is computed when the IFFT block is completely loaded, thus giving a set of complex time-domain samples representing the combined OFDM sub-carrier waveform. To create a 3.2 microseconds (20Msps/64) duration OFDM waveform, the samples are clocked out at 20 Msps (mega samples per second). To complete the OFDM symbol, a 0.8 microseconds duration Guard Interval is then added to the beginning of the OFDM waveform. A guard interval is added to each symbol to minimize Inter Symbol Interference (ISI) and the channel delay spread. This produces a ”single” OFDM symbol with a time duration of (3.2 + 0.8) microseconds = 4 microseconds in length. Now, for the remaining input data bits, this process is replicated to create additional OFDM symbols. Finally the single OFDM symbols are concatenated together and then appended to a 4 microseconds SIGNAL symbol and a 16 microseconds Preamble to complete the OFDM frame structure, where SIGNAL symbol provides Rate and Length information while Preamble is used for synchronization. OFDM frame is finally completed and is now ready to be transmitted as an OFDM Burst.

One of the biggest challenges is estimation of channel properties during data transmission.This step needs to be carried to know the Initial channel conditions before any Symbol demodulation can take place correctly. Prior to demodulating the symbols transmitted in OFDM systems, Channel estimation needs to be performed. As the channel is under the effect of frequency-selective fading and time-varying factors, This step becomes essential. Its done in two ways. The first one sees a comb type usage in which pilot symbols are transmitted on some of the sub carriers of each OFDM symbol. A different kind of interpolation techniques like time domain,Linear interpolation need to be used in this methodology. In the second type of method, Often called the pilot based block type channel estimation, pilot symbols are inserted in some of the sub carriers of all OFDM symbols or are placed in all sub carriers of a symbol of OFDM system[1]. In this channels considered are slow faded as well as a detailed study of this system while using Least Square (LS) Estimator,Minimum Mean Square Error (MMSE) Estimator and a Deep Learning Neural Network(DNN) is showcased in this paper [1].

2.4 Pilot Addition

The channel for transmission is often with faults and complications. A pilot signal is a helpful tool in checking the channel robustness. It's usually a mono-frequency single which is deployed over the communication system for multiple reasons.Pilot assists in reversal of signal distortions when signal travels through a channel,coordination and synchronisation of system and control purpose. Channel estimations can be easily done if a pilot symbol known at both the transmitter and receiver end is utilised for interpolation for estimation of channel responses.However choice of channel estimation technique is largely dependent on time variation, computation power, complexity and threshold of performance metrics. Pilot structure classification based on the arrangement of pilot signals is as follows.

1. **Block Type** - In this OFDM symbols with pilot symbols are transmitted periodically for channel estimation. Using these pilots, a time-domain interpolation is performed to estimate the channel along the time axis.This type is most suited for frequency-selective channels. Coherence time is an important parameter as pilot symbols frequency should match the coherence time which in turn is denoted as the inverse of Doppler frequency($f_{doppler}$). Hence S_t, the period of pilot symbol in time must satisfy -

$$S_t \leq \frac{1}{f_{\text{Doppler}}} \tag{2.2}$$

The arrangement is shown below -

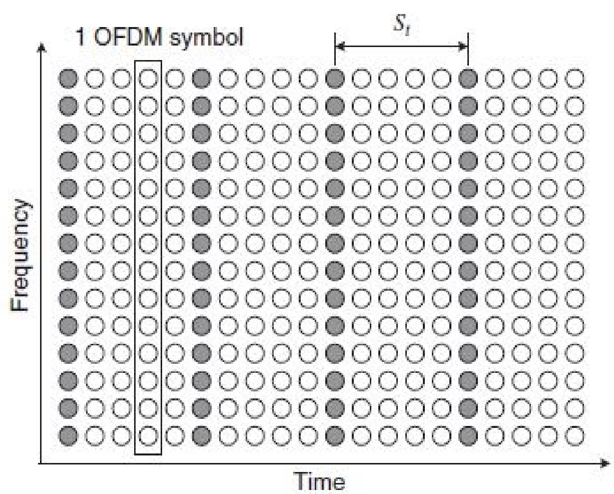

Figure 2.7 Block Type - Pilot [2]

2. **Comb type** -In this every OFDM symbol has pilot tones at the periodically-located sub-carriers, which are used for a frequency-domain interpolation to estimate the channel along the frequency axis. It's most suitable for fast-fading channels, but not for frequency-selective channels.Coherence bandwidth is an important parameter as pilot symbols frequency should match the coherence bandwidth which in turn is inverse of the maximum delay spread $\sigma_{\max}$, the This is done to track

the frequency-selective channel characteristics.Hence S_f, the period of pilot tone in frequency must satisfy -

$$S_f \leq \frac{1}{\sigma_{\max}} \tag{2.3}$$

The arrangement is shown below -

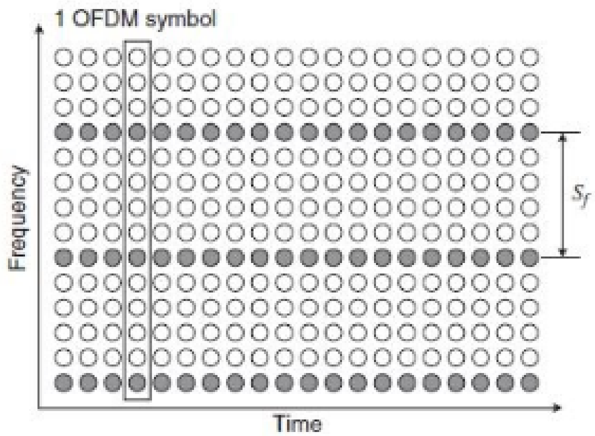

Figure 2.8 Comb Type - Pilot [2]

3. **Lattice type** - In this the pilot tones are inserted along both the time and frequency axes with given periods. The pilot tones scattered in both time and frequency axes facilitate time/frequency-domain interpolations for channel estimation.The pilot symbol arrangement must satisfy -

$$S_t \leq \frac{1}{f_{\text{Doppler}}} \quad \text{and} \quad S_f \leq \frac{1}{\sigma_{\max}} \tag{2.4}$$

where $f_{Doppler}$ and $\sigma_{\max}$ denote the Doppler spreading and maximum delay spread, respectively. The arrangement is shown below -

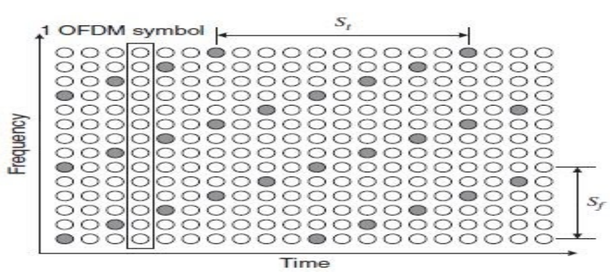

Figure 2.9 Lattice Type - Pilot [2]

In this paper, We have opted for frequency domain based pilot symbols as it's a fast-fading channel in accordance with the cyclic prefix. Reference used for this [2].

2.5 Cyclic Prefix in OFDM System

In order for OFDM systems to work efficiently and reliably one of the key element for OFDM data transmission is cyclic prefix. The main purpose of cyclic prefix is that it acts as a buffer region or in other words a guard interval that the protects the OFDM signals from inter-symbol interference(ISI) and inter-channel interference(ICI). One of the way to implement cyclic prefix is to use a cyclic prefix with almost same length as the length of the channel eventually eliminating ISI and ICI, but also using a cyclic prefix of a very short length can make the ISI and ICI to be spectrally concentrated which almost nullifies their impact on the performance of the system [16].

The cyclic prefix however two main functions in an OFDM data transmission : one was mentioned earlier that it basically provides with a guard interval to eliminate inter-symbol interference and second is that it repeats the signal so the convolution of frequency selective channel can be modeled as circular convolution which helps in converting to a frequency domain via Discrete Fourier Transform (DFT) or Inverse Fast Fourier Transform (IFFT).

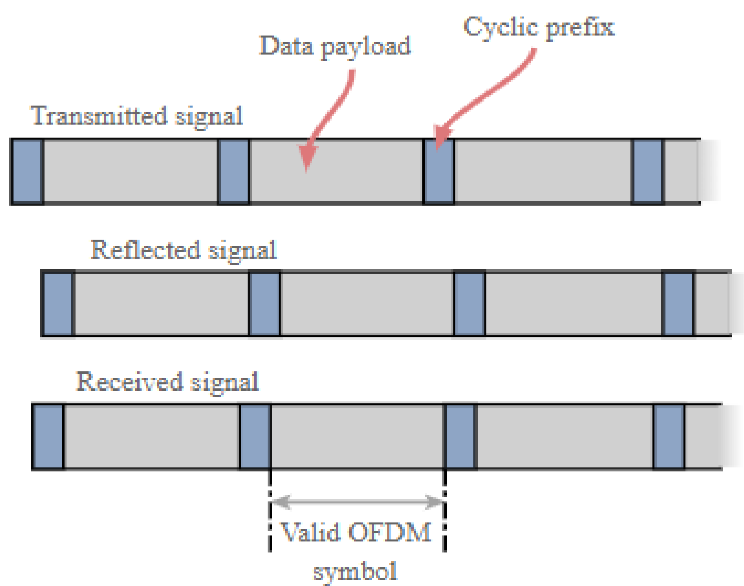

Figure 2.10 OFDM Cyclic Prefix [9]

Let us consider a model for frequency selective channel without some noise as :

$$y(n) = h(0)x(n) + h(1)x(n-1) + h(2)x(n-2) + \ldots + h(L-1)x(n-L+1) \tag{2.5}$$

As, we can see that this channel not only depends on $x(n)$ at time instant n but also **n-1** and **n-2**, so it has inter-symbol interference in time domain which represents a frequency selective channel. Here, $h(0), h(1), \ldots, h(L-1)$ are the co-efficients.

Now lets us transmit two consecutive symbols in OFDM and process them using IFFT :
First symbols are X_0 X_1 X_2...,$X_(n-1)$ and these symbols are loaded in the sub-carriers and are processed using the IFFT operation to get $x(0)$ $x(1)$ $x(2)$ ... $x(n-1)$ and,
Second symbols are X_1^1 X_2^1 X_3^1 ...X_{n-1}^1 and these symbols are loaded in the sub-carriers and are processed using the IFFT operation to get get $x^1(0)$ $x^1(1)$ $x^1(2)$...x_{n-1}^1.
So we successfully transmitted two consecutively OFDM symbol namely x and x^1.
Hence the output of the $x^1(0)$ signal would be :

$$y(0) = h(0) + h(1)x(n-1) + h(2)x(n-2) + \cdots + h(L-1)x(n-L+1) \tag{2.6}$$

And from the above equation makes it evident that we have inter-block interference(IBI) because frequency selective nature of the channel and now we can use a cyclic prefix to avoid these IBI between blocks of a channel.
Therefore instead of transmitting only the IFFT of the signals we will add a prefix which is :

$$x(n-\bar{L})\ldots\ldots x(n-2)x(n-1) \tag{2.7}$$

which results in transmitted signal into transforming like :

$$x(n-\bar{L})\ldots\ldots x(n-2)x(n-1)x(0)x(1)x(2)\ldots x(n-1) \tag{2.8}$$

Basically last $\bar{L}$ samples (or $\bar{L}$ samples from the tail) are taken and prefixed into the transmitted signal which results in transmission of (n+ $\bar{L}$) samples and thus we are cycling samples from the tail of OFDM block to the beginning of the OFDM block to add a prefix hence this is called **CYCLIC PREFIX**.
Because of equation 6, equation 8 turns out to be :

$$y(0) = h(0)x(0) + h(1)x(n-1) + h(2)x(n-2) + \cdots + h(L-1)x(n-L+1) \tag{2.9}$$

Thus the interference is now restricted to the samples from the same block and this is how we avoided the inter block interference using cyclic prefix.
Addition of cyclic prefix has resulted in circular convolution at the output of the OFDM signal.

Here,$Y(k)$ is the symbol on the k^{th} sub-carrier , $H(k)$ is the channel co-efficient of k^{th} and $X(k)$ is the symbol on the k^{th} sub-carrier . And we can notice that each sub-carrier is experiencing **FLAT FADING**. Therefore one of the main motivation we set for OFDM was to eliminate the frequency selective nature of the wireless channel and now we are witnessing with the help of cyclic prefix we are able to convert this wireless communication channel into a frequency flat fading channel across each sub-carrier.

This also proves the advantage of OFDM over single carrier transmission as we are transmitting n symbols in parallel (as compared to single signal in single carrier transmission) over n sub-carriers and the time of each OFDM symbol is n/B here B is net band of transmission and we are expanding the time of each symbol in OFDM (as in single carrier transmission the time was $1/B$) thereby removing the frequency selective nature of the OFDM.

Chapter *3*

Problems in the Literature

In Hao Ye et. al.[17] the approach of Deep learning is used in channel estimation in MIMO systems with OFDM configuration. In this paper rather than first estimating the channel state information and then demodulating the transmitted symbols at the receiver using the CSI, Here CSI are calculated implicitly and then recovers the symbols directly.In this paper, the channel distortions caused due to the propagation coefficient are removed by training the Deep Learning model offline with the dataset generated by simulation models. A key difference is that in this paper, fewer pilot symbols are added and the whole concept of cyclic prefixes is not used. Further the noise used here is nonlinear clipping rather than typical AWGN.

The paper clearly showcases how MMSE may outperform Deep learning models which is attributed to the fact that the second order statistics of CSI are known and integrated into the calculation. As expected, Lesser number of Pilot symbols would lead to better spectral efficiency but would also have the unwanted effect of having higher BER. Here the CP are omitted even though they mitigate the Inter Symbol interference.This is done to improve the time and cost of transmitting the information in the system. Further here the clipping is done to reduce the peak-to-average power ratio (PAPR). One major problem faced here is that data generated from real wireless channels can be used to better tune the model and get lower BER.

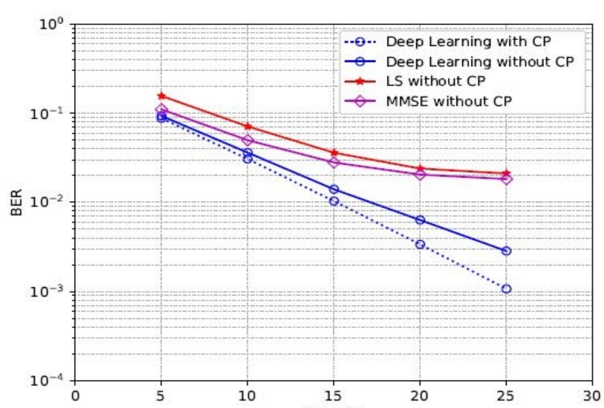

Figure 3.1 SNR vs BER for LSE,MMSE and DL [17]

In Chuan et. al. [7] To demodulate the transmitted symbols, a combined approach of Channel State Information based recovery for channel estimation with recovery of desired signal is used. Further in this paper Non Orthogonal Multiple Access(NOMA) signal with MIMO system base is used. A major breakthrough in this paper is that without any prior signal processing, the signal can be directly sent to the MIMO-NOMA system for recovery of signal. The system is constituted of three blocks. In the first one NOMA signal generation with training of the model takes place. While in the second block testing of data with an online block takes place. Finally the Deep Neural Network block with fine tuned hyperparameters like the weights, bias, regularization parameter, learning rate, and drop-out is deployed. Following the DNN Model used here, An input layer with fully connected hidden layers using Softmax algorithm and an output layer with divided blocks.

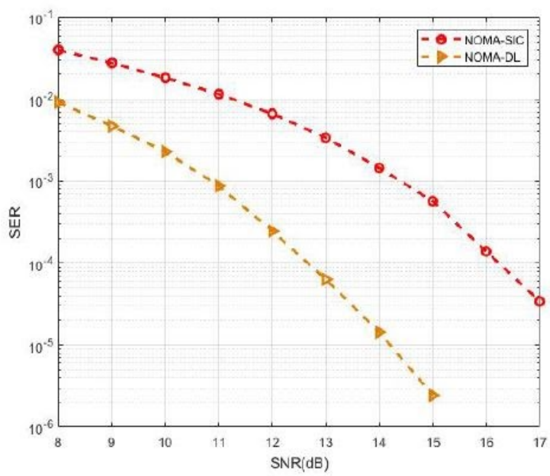

Figure 3.2 MIMO NOMA Deep Learning and traditional models [7]

Chapter *4*

Methodology

4.1 Channel Estimation Techniques : LSE and MMSE

The number of channel estimates that have to be estimated in multiple-input multiple-output (MIMO) system is in general much larger, than in a single antenna communication scheme. This leads to lower signal to noise ratios (SNR) of the channel estimates if a constant pilot power independent from the number of transmit antennas is assumed.The use of long-term spatial channel characteristics can improve channel estimation for MIMO wireless systems. Separating the signal and the noise subspace followed by a dimension reduction can significantly reduce additive noise on channel estimates. This leads to improved channel estimation, especially for MIMO systems with high numbers of antennas, and to lower pilot power requirements. [15] As channel estimation can be either in time domain or frequency domain, the estimators that are used must be relevant to the domain. In this paper most of the mathematics deal with frequency domain,Hence the most widely sought after estimators, LSE(Least Square Estimation) and MMSE(Minimum Mean Square Error) [14] algorithms are used.

Least Square Estimation(LSE) is a relatively simple and easy to implement algorithm.The complexity is also very low.Its calculated by minimising the euclidean square distance between the transmitted and received power signal. It's usually accompanied by a high Mean Square Error.The LSE channel for sub-carrier on which pilot symbols are located is given by -

$$h_p^{LS} = X_p^H y_p \tag{4.1}$$

here,h_p^{LS} is LS estimation channel frequency response and,
X_p^H is transmitted symbols and, y_p is the received signals

Minimum Mean-Square Error Estimator(MMSE) is a relatively complex and difficult to implement algorithm as it involves convolution and inverse matrix mathematics. Usually the normalisation of transmitted data is done to reduce the MMSE. The MMSE channel estimation applies the channel statistics to minimise the MSE estimate of the channel responses is given by -

$$H_p^{MMSE} = R_{HH_P}\left(R_{H_PH_P} + \sigma_\mu^2\left(XX^H\right)^{-1}\right)h_p^{LS} \tag{4.2}$$

here, R_{HH_P} is the cross-correlation matrix between all sub-carriers and sub-carriers with reference signal, $R_{H_PH_P}$ Auto-correlation matrix of sub-carriers of reference signal.

Due to the inversion matrix lemma, the complexity of MMSE estimator is high. Every time inversion is needed data changes. By transmitting the average data the complexity of the estimator can be reduced, therefore we replace the term,

$$\left(\mathrm{XX^H}\right)^{-1} = \mathrm{E}\left[\left(\mathrm{XX^H}\right)^{-1}\right] \tag{4.3}$$

Therefore, the simplified MMSE estimator becomes,

$$H_p^{MMSE} = \mathrm{R_{HHp}}\left(\mathrm{R_{HpHp}} + (\beta/\mathrm{SNR})\mathrm{I_p}\right)^{-1} h_p^{LS} \tag{4.4}$$

4.2 Monte Carlo Simulations

Monte Carlo simulations is a tool for simulating the results while taking in account the risk factor of the operation. Its highly useful in quantitative and qualitative scrutiny of data.Its more of a reformative tool that allows the user to gain insight into a wide range of possible outcomes that they may encounter while running simulations.Monte Carlo is specially useful in showing edge cases for a scenario, with best and worst possible outcomes, along with the cases in between available for scrutiny. A major mathematical concept it uses from the field of probability is pdf (or density function).Hence it accounts for the weighted probability of what the outcome may be.In this way, Monte Carlo simulation provides a much more comprehensive view of what may happen. Wen-Long Jin,et al[6], showed that Monte Carlo simulations gave similar results to analytical models built for evaluating results of Inter Vehicle Wireless communication. This improved as the number of iterations were increased. Further Monte Carlo simulations can be of various types as illustrated in the chart.

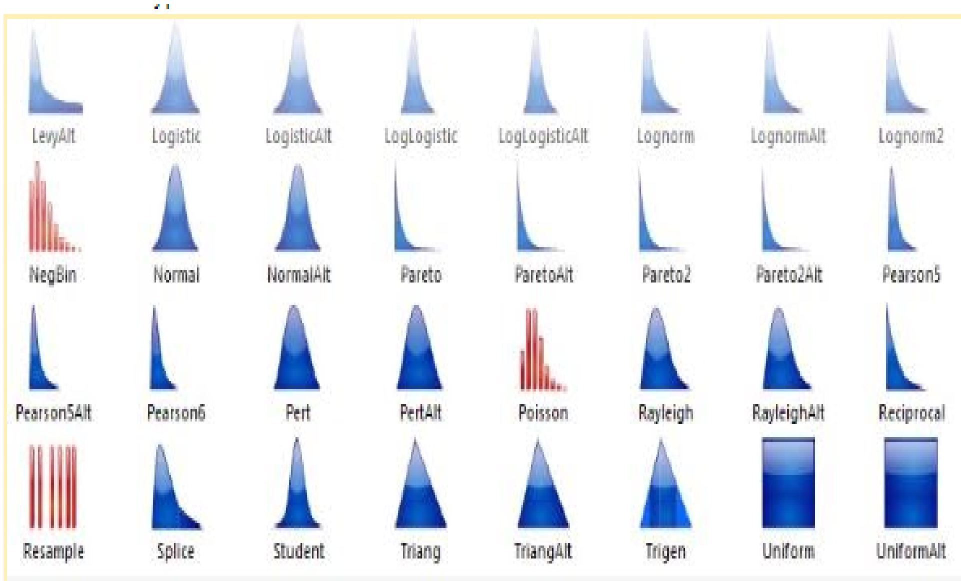

Figure 4.1 Various types of Monte Carlo Simulations[13]

4.3 Hard Decision Decoding

In this paper, We have employed the use of Hard-decision decoding over other techniques such as soft decoding,etc. This technique basically intakes a stream of bits from the receiver end and samples the pulses received. It compares the samples with a fixed threshold value and irrespective of the difference between the sample and the threshold,assign it as 0 or 1.

Following the example and methodology of Ohno,et al.[12],We implemented only one Inverse Fourier Transform(IFFT) at the transmitter and one Fourier Transform (FFT) at Receiver end. This was found to be most computationally efficient and least taxing on the hardware. Further it's the most bandwidth efficiency loss minimising setup.

4.4 Channel Estimation using Deep Learning

We know that the least-squares (LS) estimator has very low complexity in comparison to the MMSE estimator which is very much complex. As a consequence, it is nearly impossible to achieve satisfactory results with the LS estimator while the high complexity of MMSE estimator makes it difficult to achieve satisfactory results in reality with the MMSE estimator. So, in order to deal with these issues with the LS and MMSE channel estimation techniques, we try to solve our channel estimation problem through a deep neural network (DNN). The deep learning based channel estimator can outperform LS (Least Squares) estimator with almost the same complication but when it comes to MMSE estimator, then our deep learning based channel estimation approaches MMSE estimator's performance. It's clearly evident that using a deep neural network is much more beneficial and feasible as it has complexity nearly the same as the LS estimator while performance-wise it is close to the MMSE estimator.

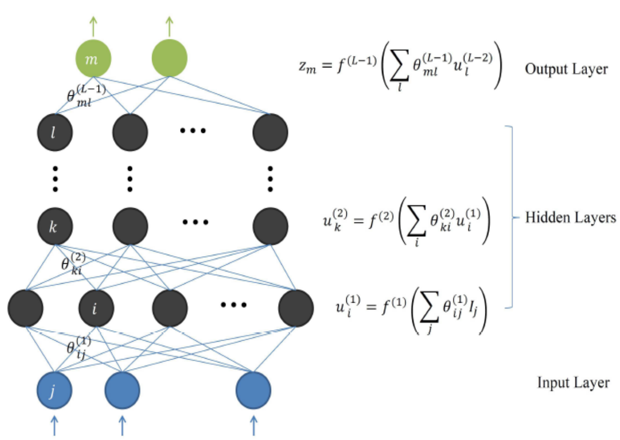

Figure 4.2 An example of Deep Learning Model[17]

4.4.1 Training data generation

The training data and the validation data for our deep learning model in a single-user OFDM system is collected for a single subcarrier which is selected based on a pre-defined metric. From the transmitter side, we send OFDM packets to the receiver side. Here each OFDM packet contains one OFDM data symbol as well as one OFDM pilot symbol. Interesting thing to note is that here data symbols can be interleaved in the pilot sequence. So effectively each training sample contains all the symbols in a received OFDM packet and it is represented by a feature vector.

4.4.2 Deep Learning Model

In the approach of deep learning that we tried to use in this project is a fairly simple algorithm. The model is easy to implement, gets an absolute validation score and has comparable performances to other means of evaluation used. Further the model also provides a SNR vs BER curve similar to the other projects conducted in this field. It builds on other research in terms of robustness as randomness of data is inculcated, Slightly better performance in terms of error rate and time for calculation. This can also be attributed to a comparatively smaller dataset used.

In this paper We used a dataset of 10000 samples out of which the training is done with 8000 samples while the residual 2000 samples are used in validation.The LSTM model architecture is executed for 100 epochs(full data pass) and its progress is monitored. The batch size for training the model in one iteration is 1000 samples which are chosen randomly.

A basic Deep learning model is a three layered Neural network system in which input data is processed and different statistical relations are found to interrelate the data.The three layers are the input layer, hidden layer(s) and the output layers. In this paper we used a Deep Neural Network which is a typical feed forward kind of NN. In this data transfer and flows unidirectionally from one layer to the next one. This implies the links created from one node to another dont touch again and they are unidirectional. We first created the input layer using MATLAB functions which take the input size(Product of number of OFDM symbols and its all cells) as the input. This simple NN as input layer takes in the data and passes it onto the hidden layers for generalisations and storing those trends. The inner layers are made using the LSTM type NN. The number of hidden units is taken as the length of cyclic prefix, ie- 16.Long short-term memory layers are typically Recurrent NN that can overcome the problem of no long term dependencies between time steps in sequential data in the RNN. Then a fully connected layer is added which takes input the previous output and multiplies with a weight and adds a bias.It involves the concept of convolution and pooling.On top of that a softmax layer is added. This is a multi class probabilistic model which issues numerical value of probabilities to each class.These should add to unity and be exhaustive. This layer must be added just before the output layer and must have the same nodes quantity as that of the final layer. Finally an output layer is added whose number of nodes are equal to the outputs required.This label to classification model creates an array of input layer,LSTM layer,softmax layer and output layer.[5]

Now the training options are defined for the architecture. MaxEpochs would define the number of full passes of the entire dataset that are to be taken in order to minimise the loss function of the algorithm. MiniBatchSize would define the size of the mini batch that is used for each round of iteration.Shuffle at every-epoch would shuffle the dataset before each training and validation cycle.This is done when the mini batch size won't properly distribute the data which in turn would force the training algorithm to reject the iteration in the epoch.To bypass this rejection of data shuffle option is set to every epoch.ValidationData defines data that is used for validation purposes of data.InitialLearnRate is taken at an optimum default level as a lower rate would be time costly while a high rate would lead to inaccurate results or even complete divergence. GradientThreshold would decide the cutoff for the gradient descent algorithm and according to which clipping takes place. LearnRateDropFactor would be the factor to apply to the learning rate after every fixed number of epochs.This methodology we implemented here uses the Adam Solver. In Adam (Adaptive Moment Estimation) a similar approach to RMSProp is used. However an extra momentum is added.

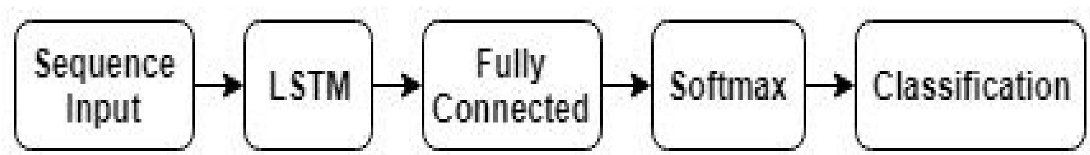

Figure 4.3 Deep Learning Model Architecture [8]

4.4.3 Methodology for training data generation

We first define the basic OFDM system parameters Number of subcarriers, Pilot spacing, Number of OFDM symbols etc. We are basically doing QPSK modulation and we define random labels for the QPSK symbols. We define pilot symbols which will remain fix during the whole transmission. For the Channel generation part we could have used different channel models like the Narrowband Rayleigh fading channel or the 3GPP channel model. Then we do SNR calculation and subcarrier selection. We now choose the subcarrier randomly in such a way that its gain is above the median value. Training data is generated in the sequence of the modulation constellation points. We define the size of the dataset to be the number of packets per modulation symbol. We use the same pilot sequences across all the packets. Now that we have defined some useful parameters, we now loop over constellation symbols. From the OFDM pilot and OFDM data symbols we generate the Transmitted OFDM frame.

The received OFDM frame can be found out from the transmitted OFDM frame using the basic transmission and reception process in OFDM systems, like adding noise, serial to parallel transformation, performing FFT (fast-fourier transform) etc. Then we perform training data collection which is thoroughly explained in the coming paragraph. So, after performing the training data collection, we Reorganize the dataset, then we split the dataset into training set and validation set (testing set). We keep the size of training data for our model to be around 80% while we have kept the validation data size to be around 20%. We finally save the training data for training the neural network.

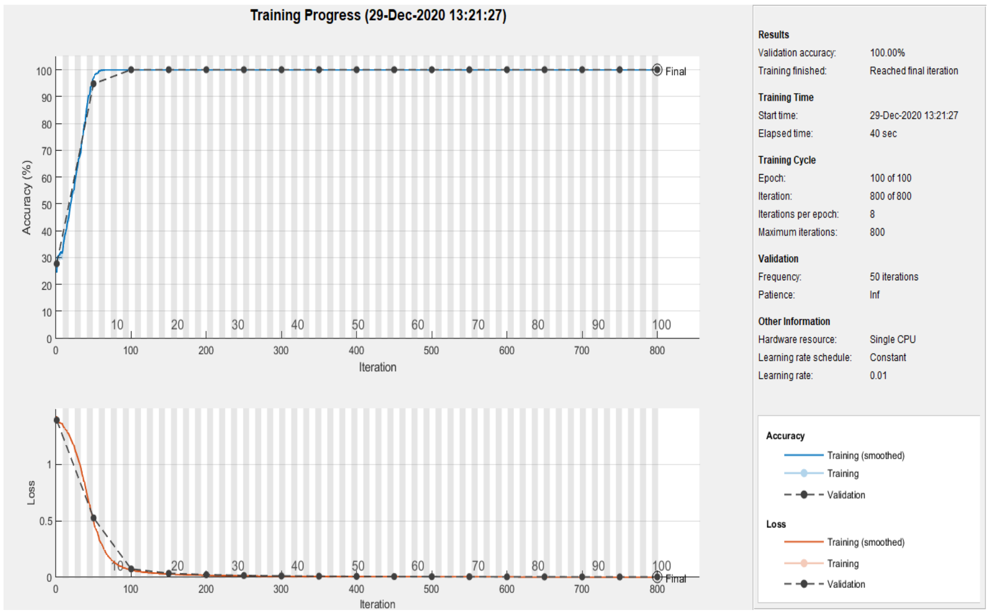

Figure 4.4 Training Data Progress in MATLAB

4.4.4 Training Data Collection

For training data collection we transform the OFDM packets which we received into feature vectors. Then we collect the corresponding labels.

4.4.5 Methodology for Training Data Collection

First of all, for training data collection, we need to determine the dimensions of our feature vector. For doing so, Number of Symbols and Number of subcarriers are required which is nothing but the size of the real part of our received packet. Then feature vector dimension (DimFeatureVec) = Number of symbols * Number of subcarriers * 2. Then we find packets of the target label and also get the number of packets. Then for data collection we find out the real part and the imaginary part from the real and imaginary parts of the received packet respectively which in turn will help us in getting our feature vector. The real part and imaginary part will be present alternatively in our feature vector. Finally we do label collection using the target label and the number of packets.

Chapter 5

Results

Following are the step by step results/simulation gathered from the execution and implementation of the code. All the code was implemented on MATLAB.

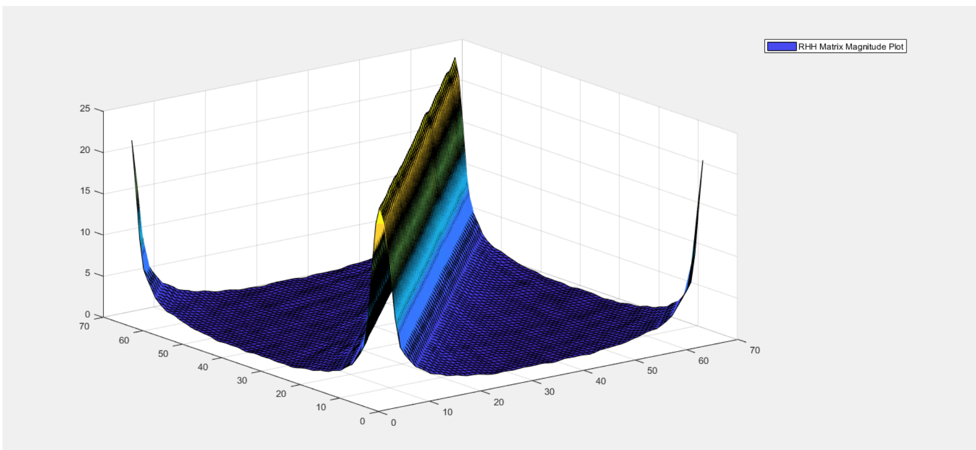

Figure 5.1 Auto Covariance of the Channel Frequency response

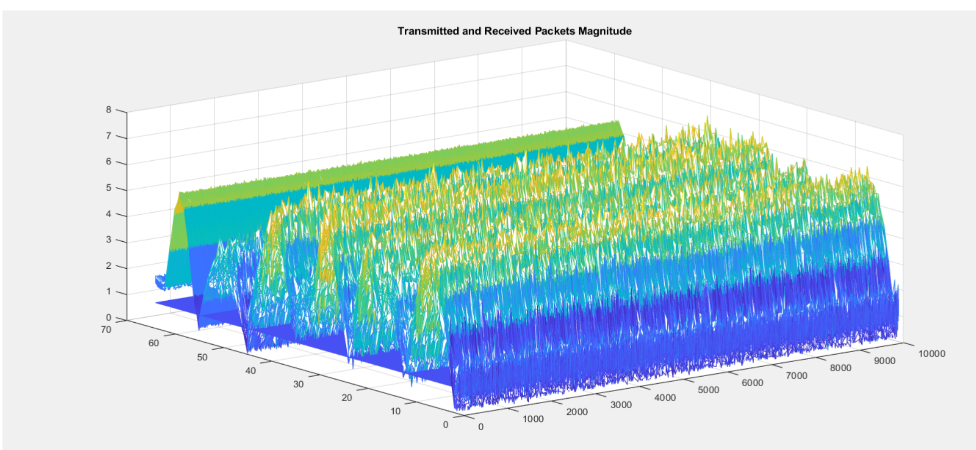

Figure 5.2 Transmitted and Received packets Magnitude

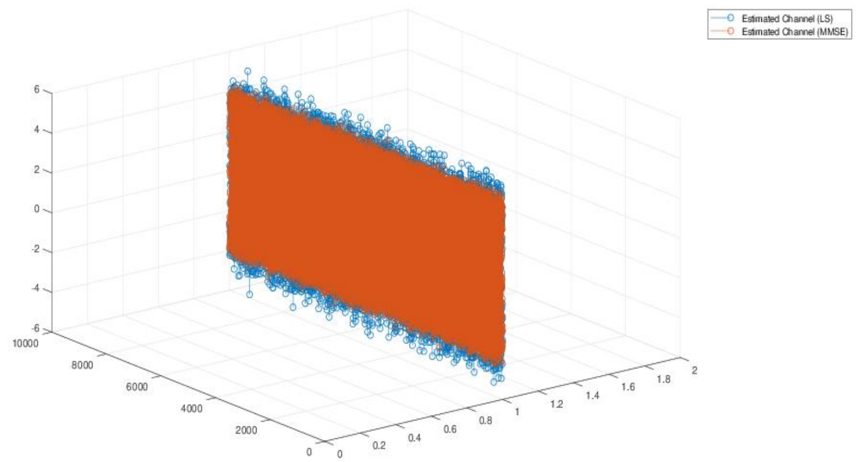

Figure 5.3 3D plot of Channel Estimation in LS and MMSE

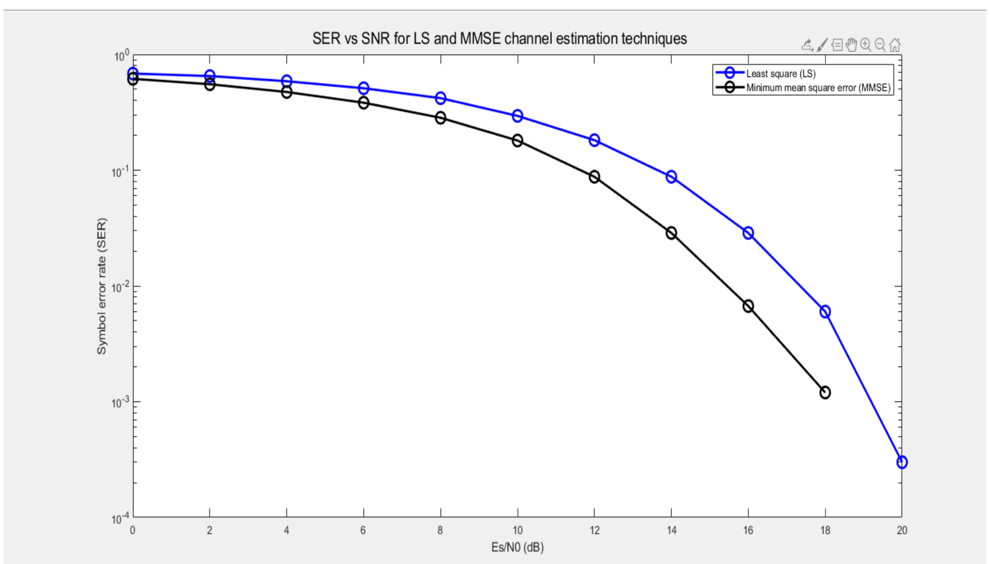

Figure 5.4 Channel Estimation plot of LS and MMSE

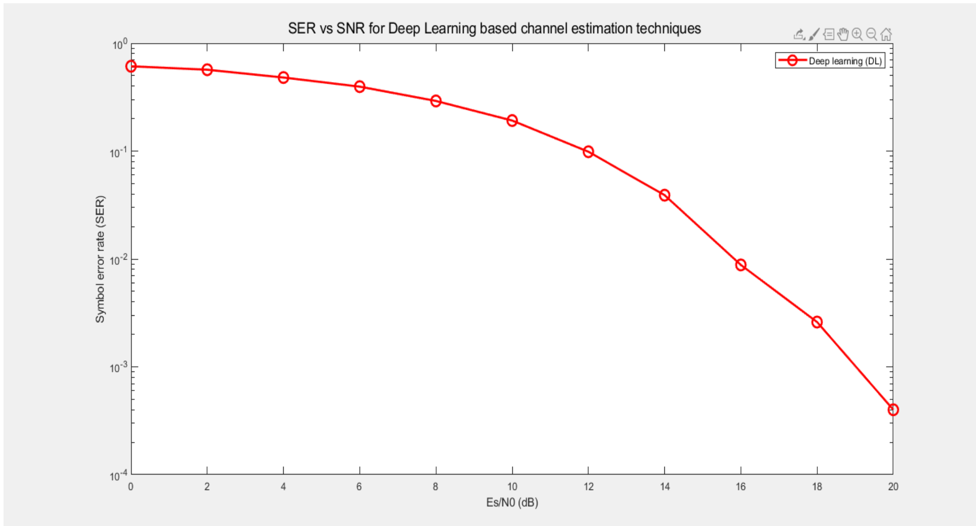

Figure 5.5 Channel Estimation plot using Deep Learning

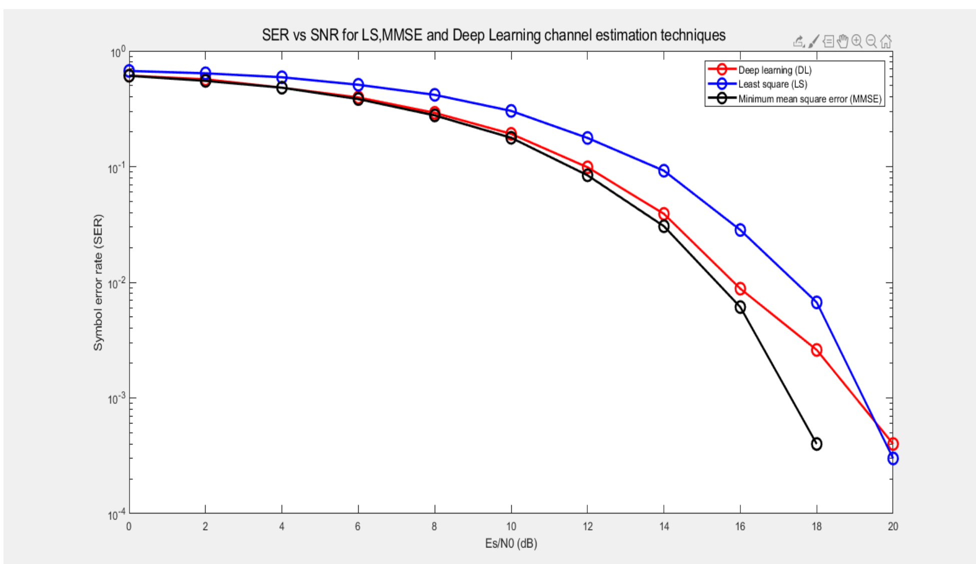

Figure 5.6 Channel Estimation plot using LS, MMSE and Deep Learning Compared

Chapter *6*

Conclusions

In this section we discuss the aforementioned results. We need efficient channel estimation techniques that can combat the dynamically changing channel.The RHH 3D plot gives the shape of a (type of value RHH is). It's a typical mathematical curve with a peak in center along the axis. The transmitted and received packet are also plotted in 3D for all n=10000 so as to emphasise on the difference in their values on a macro level. Then we developed a plot for the Estimation by LSE and MMSE.This provides a rough picture and insight into the results we can expect. Finally a plot for BER vs SNR has been generated of the LSE and MMSE for various SNR's. It's evident from the last plot that even though LSE is easier to implement than MMSE and easy to fit, the results provided by MMSE are more accurate. It can further be seen that for lower values of SNR, LSE and MMSE values map each other. When the SNR is augmented, LSE performs poorly in comparison to MMSE. The deep learning model performs significantly better than the LSE estimator as it takes into account system interferences and better noise handling capabilities. As MMSE has high complexity problems deep learning models provide comparable yet less taxing solution to channel estimation problem. It would be interesting to see the inverting points of these three estimators as at high snr values the graphs start to diverge. It would be interesting to find the inversion point of this trend.

Bibliography

[1] S. ahmed Ghauri, S. A. Khan, and M. F. Sohail. Implementation of OFDM and channel estimators using LS and MMSE estimators.

[2] Y. S. Cho, J. Kim, W. Y. Yang, and C. G. Kang. *MIMO - OFDM Wireless Communications with Matlab*. Wiley, 2010.

[3] Dorra Ben Cheikh Battikh, Jean-Marc Kelif, Marceau Coupechoux and Philippe Godlewski. Dynamic System Performance of SISO, MISO and MIMO Alamouti Schemes.

[4] R. Gate.

[5] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath. Capacity Limits of MIMO Systems.

[6] W.-L. Jin and W. W. Recker. Monte Carlo Simulation Model of Intervehicle Communication.

[7] C. Lin, Q. Chang, and X. Li. A Deep Learning Approach for MIMO-NOMA Downlink Signal Detection.

[8] Mathworks.

[9] E. Notes.

[10] C. of Orthogonal Frequency Division Multiplexing (OFDM) and . WLAN.

[11] H. H. H. U. of Technology. An Overview of Diversity Techniques in Wireless Communication Systems.

[12] S. Ohno and G. B. Giannakis. Optimal Training and Redundant Precoding for BlockTransmissions With Application to Wireless OFDM .

[13] Palisade.

[14] K. A. Parmar and S. M. Patel. Performance and Analysis of Channel Estimation Techniques for LTE Downlink system under fading with mobility.

[15] M. Stege, P. Zillmann, and G. Fettweis. MIMO channel estimation with dimension reduction.

[16] Werner Henkel and Georg Taubock and Per Odling and Per Ola Borjesson and Niklas Petersson. The Cyclic Prefix of OFDM/DMT – An Analysis.

[17] H. Ye, G. Li, and B.-H. F. Juang. Power of Deep Learning for Channel Estimation and Signal Detection in OFDM Systems.