DATA FUSION USING MIMO SYSTEM

Thesis submitted to Goa University

In partial fulfillment of the requirements for the degree of

Doctor of Philosophy in

Electronics

By

Mr. Panem Charanarur

Under the Supervision of

Prof. Rajendra S. Gad

School of Physical and Applied Sciences

Goa University, Taleigao-Plateau,

Goa - 403 206, India

October 2020

II *Shri Ganesh* **II**

Dedicated to

My Family for their

love and endless support

CERTIFICATE

This is to certify that the thesis titled **"Data Fusion Using MIMO System"** submitted by Mr. Panem Charanarur for the award of the degree of Doctor of Philosophy in Electronics, is based on his original and independent work carried out by him during the period of study under my supervision. The thesis or any part thereof has not been previously submitted for any other degree or diploma in any University or Institute.

(Prof. R. S. Gad)
Research Guide

Place: Goa University

Date:26/10/2020

DECLARATION

I state that the present thesis entitled **"Data Fusion Using MIMO System",** is my original contribution and the same has not been submitted on any occasion for any other degree or diploma of this or any other University or Institute. To the best of my knowledge the present study is the first comprehensive work of its kind in the area mentioned. The literature related to the problem investigated has been cited. Due acknowledgements have been made whenever facilities and suggestions have been availed of.

(Panem Charanarur)
Candidate

Place: Goa University

Date: 26/10/2020

Acknowledgements

I wish to extend my sincerely acknowledgment to all for their endless support and advice to materialize this thesis into a complete document. First and foremost, I would like to express a deep sense of gratitude to my mentor, Prof. Rajendra S. Gad, Department of Electronics, Faculty of Natural Sciences, Goa University for giving me an opportunity to work under his able guidance. His expert guidance and unwavering enthusiasm for research kept me constantly engaged with my research. I could have not imagined having a better supervisor and mentor for my PhD study.

I would like to thank Dr. Vinaya Gad and Dr. Udaysingh Rane for sharing their subject knowledge and valuable suggestions to my research work. Thanks to Prof. Rajendra S. Gad family members for providing nice hospitality and special treatment like a family member. A special thank to, Dr. Vinaya Gad and Prof. Rajendra S. Gad, for being wonderful family at Goa University Campus.

A sincere thank goes to Prof. Gourish Naik, (Former Dean & Head, Department of Electronics), for being a wonderful mentors and his extended support in the lab. He has always been inspired me to take up challenges during my research work.

I appreciate the critics and valuable feedback from Prof. Jyoti Pawar (Subject Expert) during the Departmental Research Committee (DRC) manifestations, which has always given me the step towards the improvements and scope to enhance my knowledge in the field.

I am highly indebted to Dr. Jivan Parab, member of DRC for his advice and timely suggestions that enable me to improve the quality of the thesis.

I am very grateful to the Ministry of Electronics & Information Technology (MeitY), Government of India, for financial assistant thereby granting me fellowship under Vishveshwarya Ph.D. Scheme for Electronics System Design & Manufacturing

(ESDM) and IT/IT Enabled Services (IT/ITES) sectors in the country. I also thank Intrinsic Solutions, Bangalore for training me to work on MIMO system.

My appreciation also extends to all the anonymous reviewer for their constructive criticism that enabled me in improving the quality of my research articles.

I am fortunate to have very good friends at the Goa University and found very pleased being with them all the time. I must be grateful to Dr. Narayan Vetrekar and Aniketh Gaonkar, for being like a brothers and family friend, during this journey, their endless support all the time.

I would like to acknowledge Dr. Noel Tavares, Dr. Ingrid-Anne Nazareth, Dr Sulaxana Vernekar, Dr. Shaila Ghanti, Dr. Niyan Marchon, Dr. Supriya Patil, Mr. Caje Pinto, Mr. Deepak Chodankar, Mr Sameer Patil, Mrs Vilma Fernandes, Mr Anish Prabhu, Mr Daryl Gonsalves, Mr Girish Abhyankar, Mrs.Yogini Prabhu, Mr. Marlon Sequeira, Mr. Abhiraj Pednekar, Mr. Madhusudan Lanjewar for being a wonderful lab members and their valuable suggestions and broad discussions on various scientific talks. I would also like to thank Mr. W. D'Souza, Mr. V. Malik, Mrs. A. Velip, Mrs. P. Andrade, Mr. A. Lopes for supporting and helping me in the laboratory and administrative works.

Finally, I must always be indebted to my parents, my sisters for their support to stand in all the challenges that I have faced in this phase. A soulful and respectful thanks to my parents, who put effort and hard work for making my pathway with full of flowers. I am fortunate to have a little sister Kaivalya, and a cute brother Aditya (Adhi). I would like to thank my wife, Mrs. J. Vijaylaxmi, for her valuable support in all means. I am thankful to all of them for sharing a special part of their life forever with their love, affection and care all the time.

Panem Charanarur

Abstract

In the Modern day of Wireless Communication Systems, there is a direct need for the enhancement of the network capacity to support more and more data. The problem becomes more significant when more sensors data is sent through limited bandwidth in a communication system. The digital data transmitted over low capacity channel results in increased error rate due to which the regression or regeneration of original data at the receiver is difficult. One may note that data communication using the source and channel coding having MIMO channel shows a noticeable improvement in efficiency and reliability. Whereas the efficiency of the bandwidth with a limited communication channel, severely degrades, when the system operates at a high data transmission rate. With the recent advancement in the source coding technique, the modulation schemes, error corrections and detection channel coding codes and MIMO channels, the reliability of the communication system can be enhanced for improved quality of service. This thesis work demonstrates the source and channel coding with Principal Component Analysis (PCA) and Gallagher Codes, at the same time using the state-of-the-art modulations such as M-PSK, M-QAM over 1x1 to 4x4 MIMO channel. The source coding is implemented using PCA for the ensemble of 10 to 50 signals over 100 samples in the time domain. The decoded signal is thus channel coded with LDPC codes and transmitted over the MIMO channel. The evaluations demonstrated with performance measurement matric such as RMSE and BER rate. Specifically, the efficiency of transmission is quantified by RMSE and reliability quantified by Bit Error Rate (BER).

In the experimental evaluation, the performance is computed for efficiency and reliability using different sets of ensemble data. These ensembles (consisting of 10, 20, 30, 40, 50 sets of signals) are the sets of random frequencies signals ranging from 50Hz to

2500Hz. The performance for efficiency and reliability is computed using RMSE and BER over a noisy channel having SNR varying from -15dB to +40 dB with LDPC error correction codes. Further, the PCA was simulated with the BPSK system using Matlab Simulink environment and Xilinx System Generator and synthesized on the Vertex-6 FPGA. The performance of the sensor nodes network over MIMO system 2D, 3D, 4D network on Chip (NoC) mesh topologies is demonstrated more in detail for CBR, and FTP applications in later part of the thesis.

Publications of the Thesis

Journals

1. Charanarur Panem, Vinaya Gad, Rajendra S. Gad "Sensor's Data Transmission over BPSK Using LDPC (Min- Sum) Error Corrections over MIMO Channel: Analysis over RMSE and BER", Materials Today: Proceedings (Elsevier), Vol.27, pp.571-575, 2020, [IF: 0.576].

2. Charanarur Panem, Deepak Chodankar, Rajendra S. Gad "PCA Data Reduction for MIMO based Channel, capacity Enhancement Using 16-QAM Modulation", Universal Journal of Electrical and Electronic Engineering, Vol. 7(2), pp. 118 – 126, 2020, [IF: 0.44].

3. Charanarur Panem, Vinaya Gad, Rajendra S. Gad "Sensor's Data Transmission over QAM Using LDPC over MIMO Channel to enhance the Channel capacity: Performance analysis over RMSE and BER", IOP: Materials Science and Engineering, Vol. 561, pp: 1-9, 2019, [IF: 0.53].

4. Charanarur Panem, Narayan Vetrekar, Rajendra S. Gad "Data Reduction And Recovery in Wireless Communication System: An Extensive Experimental Evaluation Using PSK And QAM Modulations", accepted for review. International Journal of Communication Systems (Willey), 28, August-2020.

5. Charanarur Panem, Rajendra S. Gad, Brajesh Kumar Kaushik "Vertical Traversal Approach Towards TSV's Optimization Over Multi-layer Network on Chip (NoC)", accepted for review. Microelectronics Journal (Elsevier), 2, October-2020.

Published Book Chapter

1. Charanarur Panem, Vinaya Gad, Rajendra S. Gad "Polynomials in Error Detection and Correction in data communication system", Coding Theory, Intec open, 2019, DOI: 10.5772/intechopen.86160.

Conference

1. Charanarur Panem, Deepak Chondakar, Rajendra S. Gad "MIMO Channel Capacity Enhancement Using PCA Data Reduction Methods with QPSK Modulation", IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics (2019 DISCOVER), On August 11-12, 2019. Manipal Institute of Technology, Manipal, Karnataka, India.

2. Charanarur Panem, Rajendra S. Gad, "ICA, PCA and PLSR for data reduction in signal transmission ", 11th Annual Symposium on VLSI and Embedded System VSI-2018, Goa University, Goa, March, 22-23, 2018.

3. Charanarur Panem, Udaysingh V. Rane, Vinaya R. Gad, Rajendra S. Gad, G. M. Naik "Evaluation of 3D &4D Two Layer Mesh NoC for Fault Tolerance over Combinations of Vertical Channel", 2018 IEEE International Conference on Current trends towards Converging Technologies (ICCTCT 2018), Coimbatore on 1 – 3 March 2018.

4. Charanarur Panem, Udaysingh V. Rane, Vinaya R. Gad, Rajendra S. Gad, G. M. Naik "Evaluation of 2, 4, 8 layer 3D Mesh NoC With Virtual Channel for Capacity Enhancement ", 2018 IEEE International Conference on Current trends towards Converging Technologies (ICCTCT 2018), Coimbatore on 1 to 3rd March, 2018.

5. Charanarur Panem, Udaysingh V. Rane, Vinaya R. Gad, Kovendan AKP, D. Sridharan, Rajendra S. Gad, Gourish M. Naik "Performance Analysis Of 16x16, 32x32, 64x64 2D Mesh Topologies For Network On Chip", IEEE International Conference WiSP NET-2017, SSN Engineering College, Chennai, on March, 22 to 24th - 2017.

6. Charanarur Panem, Udaysingh V. Rane, Rajendra S. Gad, "Network On Chip Performance Analysis over 2-D and 3D Mesh Topologies For CBR and FTP applications", 10th Annual Symposium on VLSI and Embedded System VSI-2017, GEC, Goa, April, 28th - 2017.

7. Panem Charanarur, A. A. Gaonkar, U. V. Rane, A. B. Pandit, R. S. Gad "Sensor data fusion architecture over MIMO: Case study of quad copter", IEEE International Conference on Electrical, Electronics and Optimization Techniques-ICEEOT 2016, DMI Engineering College, Chennai, on March, 3rd to March 5th - 2016.

Workshop

1. Attended ACM Summer School on Data Sciences to be held at Goa University from 4th - 16th, June 2018.

CONTENTS

List of Figures

List of Tables

Abbreviation List

ACM Adaptive Coding and Modulation

ADSL Asymmetric Digital Subscriber Line

AWGN Additive White Gaussian Noise

ASK Amplitude Shift Keying

BER Bit Error Rate

BPSK Binary phase-shift keying

BS Base Station

BW Band Width

BWA Broadband Wireless Access

CBR Constant Bit Rate

CDs Compact Disks

CDR Compression Driven Routing

CHs Cluster Heads

CR Compression Ratio

CSI Channel State Information

DC Data Compression

DCT Discrete Cosine Transform

DGN Data-Gathering Node

DoF Degrees of Freedom

DSP Digital signal processing

DVB-C Digital Video Broadcasting-Cable

DVB-S Digital Video Broadcasting-Satellite

DVDs digital versatile disks

FFT Fast Fourier Transform

FEC Forward Error Correction

FIFO First In First Out

FPGA Field programmable Gate Array

FSK Frequency Shift Keying

FTP File Transfer protocol

EGC Equal-Gain Combining

GIS Geographic Information Systems

HSPA High Speed Packet Access

ICA Independent Component Analysis

IoT Internet of Things

ITU International Telecommunication Union

LDA Linear Discriminant Analysis

LDPC Low-Density Parity Check

LLR Likelihood-ratio

LOS Line-Of-Sight

LTE Long-Term evolution

MEMS Micro Electro Mechanical Systems

MFI Multisensory Fusion and Integration

MIMO Multi-Input Multi-Output

MISO Multi-Input Single-Output

MMSE Minimum Mean Square Error

MPEG Moving Picture Experts Group

MRC Maximal Ratio Combining

NASA National Aeronautics and Space Administration

NEMS Nano Electro Mechanical Systems

NoC Network on Chip

NI Network Interface

OFDM Orthogonal Frequency Division Multiplexing

PCA Principal Components Analysis

PLC Power Line Communication

PREMON Prediction-Based Monitoring Mechanism

PSNR Peak Signal to Noise Ratio

PSK Phase Shift Keying

QAM Quadrature Amplitude Modulation

QoS Quality-of-Service

QPSK Quadrature phase-shift keying

RDC Routing Driven Compression

RFID Radio Frequency Identification

RMSE Root Mean Squared Error

SISO Single Input Single Output

SIMO Single Input Multiple Output

SNR Signal-to-Noise Ratio

SoC System-on-Chip

STBC Space Time Block Coding

TB-CCD Tail-biting Convolution Code Decoder

TSV Through-Silicon Via

WiMAX Worldwide Interoperability for Microwave Access

WLAN Wireless Local Area Network

WSN Wireless Sensor Network

Chapter 1

General Introduction

1.1. Introduction

Data fusion combines different data and information into a coherent, precise, and usable representation that reflects the same real-world entity. Data fusion techniques were used in multi-sensor systems to combine and integrate data from different sensors and applicable in various fields, including text processing. The goal is to achieve a low possibility of error detection and higher efficiency by utilizing data from multiple distributed sources via data fusion in multi-sensor environments [1].

In general, activities that include varieties of parameter estimation from multiple sources are advantageous with data/information fusion methods. Usually, the terms information fusion and data fusion are synonyms in some instances. Whereas the word data fusion used for raw data obtained directly from the sensors, and the word information fusion is to describe the data that has been already processed. While the term fusion of knowledge implies a higher degree of semanticity than the fusion of details, certain concepts usually used in the literature relevant to data fusion include decision fusion, data combine, data compression, multi-sensor data fusion, and sensor fusion.

Data come from several sensors that are used to decide on the status of an autonomous mechatronic device. Therefore, robust fusion strategies are required for optimum knowledge fusion for distributed multiple sensors to assemble the knowledge of compress criteria, the technology of multisensory fusion and integration (MFI) is defined. However, the MFI's application involves industrial automation, intelligent robotics production, military applications, biomedical and micro electro mechanical (MEMS)/nano electro mechanical systems (NEMS). The MFI's goal is to have more precise representations of optimum decision-making device. In contrast, the advantage of MFI is that the system can supply high-quality data on some elements of its

environment, which cannot be sensed explicitly by any particular sensor working independently. Therefore MFI has been an essential technology for creating a sophisticated mechatronic system [2].

Multi-sensor fusion and integration is a technique that refers to the synergistic combination of multiple sensor data to make inferences that are not feasible from each sensor working separately. Without the use of multi-sensor fusion techniques, the advancement of sensor technology is inadequate. Since the sensors of different types incorporated in the device have their limitations and perceptive uncertainties, a proper data fusion approach can probably minimize overall sensory possibilities, thereby increasing device output accuracy. The significant advantage of the MFI is to obtain mutual perceptions and provide timely information by parallel sensory data processing [2].

The widespread adoption of cell phones and wireless computer networking in recent years and the extraordinary development of the telecom industry are increasing the demand for wireless communication. The significant drawbacks of this growth are the disadvantages over traditional wireless communication systems, due to the available frequency options, bandwidth, and the efficiency of the networks, complexity, reliability, data rates, and physical areas. The modern multi-input multi-output (MIMO) technology which is future wireless would be more effective in meeting the growing demand for wireless communication in limited frequency resources. MIMO channel is a frequency selective channel (multipath), known to improve high-data transmission improving channel efficiency, low power implementation, and sophisticated signal processing algorithms. The wireless infrastructure designers face various challenges in meeting the demand for wireless communication with higher data levels, improved service quality, and higher network efficiency including restricted radio frequency spectrum coverage and propagation issues caused by several factors such as fading and

multipath distortion. Such specifications include new techniques of multiplexing that boost spectral efficiency and reliability. The MIMO aims to provide this capability [2] cost-effectively.

To enhance communication efficiency, MIMO uses multiple antennas both for the transmitter and the receiver as one types of intelligent antenna technology. MIMO technology has gained an interest in wireless communications due to the data rate and contact range without additional bandwidth, or transmission power, which increasing considerably. MIMO also accomplishes higher spectral efficiency (more bits per second per hertz of bandwidth), reliability and diversity (lower fading). Hence MIMO is considered as a universal wireless research theme [3]. Sensor data fusion techniques are used to obtain higher-level information from multiple sensor data by combining space-time data, leveraging redundant and complementary data, and the background available. This growing branch of applied computer technology seeks to produce complete, accurate, and almost real-time signal data that are keys to further decisions or actions. Significantly applications in the areas of logistics, advanced control systems, medical, public safety, defence, aerospace, robotics, industrial development, precision farming, trace monitoring are possible.

The fusion processes are further classified into medium, intermediate, and high-level fusion models.

• Low-level combining or raw data fusion or association: combines multiple raw data sources to create new data, which is more insightful than the input.

• Fusion at the intermediate level or fusion at the level of features or state estimates blends various features like corners, curves, textures, or positions into a feature map that can then be used to segment and detect it.

• High-level integration or combination of decisions blends multiple expert choices. Decision fusion methods include voting methods, fuzzy-logic methods, and statistical methods.

The technology representing the characteristics of the three designed sensor types are not for the purpose to exclude one another. However, these types of systems which are hybrid are one of the multiple sensors which are generally used in monitoring some specific regions which mostly fall for two or more sensors. At the same time, these may be comparative or cooperative according to the sensors arrangements [4].

Sensing data is obtained as data sources capable of producing vast quantities of real observations from the environment [5]. Such instruments are intended to gather data of only one nature, such as temperature, light, and humidity. These types of data are referred to as univariate data. New tools are designed simultaneously to collect various kinds of data from the field called multivariate data. The nodes in these networks are generally equipped with multiple sensors to collect different types of data simultaneously.

Moreover, an increasing type of data is called an attribute or function in the multivariate data. Besides, measurement of sensing data is said to be anomalous if one or more of its characteristics are anomalous [6]. The anomaly detection is applicable to univariate data by finding the single data attribute, which is abnormal compared with the aspects of other data instances. However, the identification of deviations in multivariate WSN is difficult since there could be unusual activity in the individual attributes [7]. Analyzing multivariate data is computationally costly and anomaly detection on multivariate data provides high precision when the relationships between different quality are exploited carefully [8, 9]. Mostly there are spatial and temporal similarities between sensor readings, even in sensing data. While the temporary correlation is the readings collected during the same time correspond to the readings

obtained during the previous time, and the spatial correlation is the readings of geographically similar nodes should be correlated [10].

Data communication is useful to gather information, including light, precipitation, air quality, wind speed, temperature, and other critical signals. Higher the data transmission, the demand for service quality grows. To support service quality, the use of the MIMO communication channel is growing, because of the bottleneck of channel bandwidth in the conventional data transmission network.

In a monitoring system like healthcare, business, ecology, government, and military applications, it is necessary for continuous transmitting data over a sufficient period. However, it is not easy for WSNs to perform such large-scale tasks for a long time as a power resource has a severe limit. As energy conservation and the durability of WSNs are the key considerations. However, the major part of the collected data is typically redundant and can be derived from other observations. Thus, to minimize or avoid unnecessary transmissions in the communication of data, reducing the total energy consumption is very significant. Thus leads to a lifetime of energy expended on the transfer of data [11-13].

The world of today's rising technological era is the most critical issue in communications systems for data compression. Data compression is used to reduce the amount of storage needed for information like text, images, video, sound, etc. to be transmitted or stored. It is in the process of compressing data or files allow a minimal amount of space than in its original form [14].

The communication system has transferred large-scale data in recent years, and the development of ample storage and retrieval systems has seen tremendous growth. The design of new storage devices should accommodate the rise in database capacity, modems and multiplexers must be continually updated to allow a significant amount of data transmission between the communication network and remote terminals resulting

in a rise in prices and supplies. However, the approach to this problem is compression, with the effective encryption of the data and transmission series. Compression is possible when the data has displayed in a more than appropriate format, while the input data correlates with a certain amount of redundancy [15].

1.2. Background Information

1.2.1. Data Compression

Data compression is the mechanism of eliminating the redundancies to which the number of bits required to display signal content. While the leading technologies for transmitting vast quantities of real-time signal data (sensor level signals), very often small bandwidth channels were data compressed [16].

Furthermore, the data transmission is in the form of a signal, image, graphics, audio, and video from the sensor. In the transmitting process, types of data are compressed, or else the data would need ample storage space and transmitting bandwidth. Enormous volumes of data cannot work because very often, there is insufficient storage capacity. Hence the data must be compressed using some of the analytical methods, which can be quickly transmitted. Data compression is the method of transforming an input data source into another smaller sized data stream.

There are a large number of technical improvements over the last decade as transition involves the ever-present, ever-increasing internet, the booming mobile connectivity growth, and the ever-increasing value of video contact. For the growing part of this multimedia transition, data compression is one of the technology. Without data compression, wireless networks will not be able to have improved consistency of communication. Data compression is the art of compact representation of knowledge.

Despite rapid advancements in mass storage capacity, processing speed, and the efficiency of digital communications, demands for data storage space and network

bandwidth continues to surpass available technology capabilities. Big data files remain a significant bottleneck within applications in a distributed environment. However, data compression is an essential component of possible approaches for the development of usable and transmissible real-time signal data sizes as the portability and efficiency of the application are critical in choosing the technique for compression/decompression.

The digital signals are processed or distributed, and hence, they have become an essential part of our daily life. The dependence on digital media continues to increase, as it is a major concern to find professional ways to store and distribute significant volumes of data. As the amount necessary for carrying unadulterated signal data may be incredibly large in terms of cost and the massive bandwidth needed to transmit such data, researchers are searching for ways to effectively reflect such digital data to facilitate their delivery and save disk space.

The data compression strategy has now become very relevant and widely applicable. The area of data compression involves numerous source coding schemes from conventional lossless technologies and recent solutions of converting the new forms of segmentation coding. The effect of digital data on commercial, science, and computer applications is tremendous [17]. Uncompressed digital signals need considerable bandwidth and storage space. The rapid rise of data intensives and many academic studies on both lossy and lossless data compression [15] have rendered successful data compression solutions.

Besides, other applications, the issue of data compression is more important, especially for progressive transfer, data navigation, multimedia applications, and compatible multibit transcoding. However, the majority of current Internet bandwidth is considered for sensor levels signal data, such as pictures and video transmission. Flexible, energy-efficient deployments capable of performing multimedia functions such as signal processing, coding, and decoding are significant [15].

As far as data quality [18] is concerned, the digitalization of signals is the most important move from the analog world to the digital world. However, the step is expressed in the basic necessity of preserving data quality in two basic quantitative parameters:

• Digital data transmission rate or data speed (megabit per second or MB / s)

• Total digital storage or data capacity required (megabit or MByte).

It will have a high compression rate at the same time so that the uncompressed file does not vary significantly from the original file and retains almost all relevant material.

Signal-to-Noise Ratio (SNR) is the ratio of original data and coded/decoded data. It is expressed in terms of the logarithmic decibel (dB) scale. Root Mean Squared Error (RMSE) is a total error measure that is defined as the square root of the variance value and the square of the bias. While Compression Ratio (CR), the CR is defined as the ratio of original data and compressed data.

Compression Ratio = Original data Size / Compressed data Size.

Data Compression advantage is it reduces computing demands and incurs lower transmission costs. It takes less time to load than their more sluggish predecessors, allowing more data on signals to be displayed in a shorter period. This also reduces the probability of transmitting errors.

Data compression algorithms are divided into two groups:

- Lossless Data Compression
- Lossy Data Compression

Lossless data compression operates by compressing the total data without extracting any information of the signals. As a result, the overall data size is decreased only by half to one third when data is compressed by using a lossless method, signal uncompressed and correlates with the consistency of the original signals [18-19].

Lossy Data compression is substantially different from raw binary data compression. The compression algorithms for general purposes may be used to compress signals, but the effect is less than optimal. As the signals have specific statistical properties that encoders specially built for them can use. In order to conserve a little more bandwidth or storage, some of the better information in the data may be sacrificed. It also ensures that the techniques for loss compression can be used in this area. Lossy compressions, operated by extracting data information from signals, are not transparent to the viewer. However, loss compression will limit signal data to one-tenth of its original size without significant data quality changes [18].

Source coding is a compression technique used for the reliability of data communication system. Symbol codes (for example, Huffman codes) or stream code (for example, Arithmetic codes, Lempel-Ziv codes) are few methods used for source coding techniques. The PCA is also a source coding method used to reduce/compress the input sensor data measurements [20]. In this technique, an orthogonal transformation is applied to convert a series of measures of interrelated variable values into a set of variables called Principal Components Analysis (PCA) [21].

1.2.2. MIMO with LDPC and Modulation

MIMO wireless technology significantly increases a channel's capacity. With the increased number of receiving and transmitting antennas, the channel efficiency can be linearly increased with each number of antennas connected to the network. MIMO wireless technology has since become one of the most effective wireless channel coding methods in recent years. Spectral bandwidth is becoming a resource for radio communications systems and is required to allow more effective use of the usable bandwidth [22]. In current wireless protocols such as IEEE 802.11n, 3GPP LTE, and WiMAX (4G) mobile networks, Multiple inputs Multiple outputs (MIMO) antenna

networks are used. The methodology supports better data output even under interference, signal loss, and multipath conditions. WLAN, 3G (long-term evolution), and several other radios, cellular and RF innovations are being listed and used in several emerging applications [23]. The MIMO wireless technology provides expanded communication capability and spectral quality along with better link stability.

It is also essential that the network is built to be energy-efficient. Many researchers are promoting cooperative communications to save energy by making node groups work together to transmit or receive data. The space diversity is produced by transmitting the same signal, or highly clustered versions of it, through many spatially separated antennas, is exploited in cooperative communication and references [24]. The theory of cooperative communication is also closely related to MIMO technology. Under a given budget and fading conditions, MIMO communications deliver a much higher performance (spatial multiplexing gain) or better efficient communications (diverse gain) than single input single output (SISO).

MIMO systems are constructed using multiple transmitting and receive antennas that are under extensive study to support a significant increase in channel capacity. MIMO transmission benefits from the fact that signals on different antennas have a degree of differential attenuation of their signal rates in a scattered setting where antennas are spread and isolated in a multi-transmit fading network and multiple antennas are supplied [25] [26].

The most prevalent error correction code is the Low-Density Parity Check (LDPC) code, has near Shannon limits. The Shannon limit determines the maximum data levels that can be allowed or the total number of users under the defined bandwidth and power constraints. The LDPC code is a linear error correction code, a technique used to relay the data in a noise-filled transmission channel [27]. The Sparse Bipartite

Graph is used to construct the LDPC. LDPC codes are used for maximizing the efficiency of the signal, which causes the noise level for the symmetrical signal network to be set very close to the theoretical limit (Shannon limit).

The type of modulation is one of the most significant characteristics often used in identification as well as classification. M-PSK/M-QAM is both an analog as well as a digital modulation technique, which has to utilize a limited number of at least two phases and two amplitudes. The M-PSK / M-QAM is used continuously as a computerized communications modulation technique. High spectral efficiencies can be obtained using M-PSK / M-QAM by setting the correct constellation scale, constrained only by the associated noise rates and the linearity of the transmission channels [27].

The communication system is designed to get the lowest possible error (bit error) from the sufficient resource usage data with a bandwidth cap, which remains an alternative to the low data rate. LDPC codes can be used as an excellent coding scheme to get better productivity in crowded networks. However, to achieve sufficient channel efficiency, one has to generate a high data rate is with the ideal usage of bandwidth. Therefore it is challenging for low power growth and improved channel capacity over restricted bandwidth for specific SNR. Hence an attempt is made to solve these sensor level problems and method of data reduction over the MIMO system to improvise any aspect in a wireless communication system [28] [29].

Network on Chip (NoC) is introduced as a highly structured and scalable approach for the issue of communication in SoC. On-chip, interconnection network provides advantages over dedicated wiring and buses, i.e., low-latency, low-power consumption, and scalability. Each resource should be contacted to a switch in the network via a network interface (NI). More topologies have been discussed for this switched network includes 2D Mesh, Folded Torus, Ring, Butterfly Fat Tree, Octagon, and irregular connection net-works [30] [31].

1.3. Motivation

Data transmission is primarily motivated by the ease and flexibility of handling digital signals data information instead of the analog information in the communication system. Signals data transmission is required most commonly for educational, business documents, medical images and so on. In a real sensor level, signals consume more amount of storage. As the transmission and storage of every bit incur a cost, the advancement of cost-efficient data compression techniques is of high significance. In addition to the extremely higher data transmission and bandwidth requirements, the use of uncompressed digital signals data adds a high cost to the hardware and systems that process the signals data. At the present state of technology, the only solution is to source coding for compress data before its transmission and to decompress it at the receiver end for playback in the communication system.

Cost per bit is a valuable measure for measuring data transmission rates and evaluating them across systems, service providers, or industries. Nevertheless, estimates of cost per bit cited as absolute values (e.g., $5 / GB) may have a nice punch line, but are also deceptive or unreliable. Ratios of network usage reflect the cumulative use of network services around the footprint and are directly related to averages of cost per bit. As a consequence, wireless networks are built for each position to reach the coverage requirements and the expected peak traffic load. Data per bit has significant consequences for the business model and would become more apparent in the 5G period. When you expect that the cost of providing a gigabyte on the mobile network would decrease 50 percent or more from 4G to 5G, boosted by additional technology, the usage of leased and unlicensed technology and also small cells, so the cost of using a gigabyte of "encrypted" data would be similar to the cost of using a gigabyte of "fixed" connectivity, particularly on a "real" basis. Then the 5G broadband alternative may well become a completely integrated substitute for a cabled connectivity service.

1.4. Contribution

The suggested framework is beneficial when data on compressed sensor level signals are transmitted in the MIMO channel communication network. This research contributes to the compression of data using source code PCA technique, transmitting sensor level data with channel code LDPC and M-PSK, M-QAM modulation techniques, 2x2, 3x3, 4x4 MIMO channels, without losing data during transmission.

Data communication is employed to gather a variety of information, which includes light, humidity, air quality, wind speed, temperature, and other vital signals. A monitoring system is being developed in health care, industries, ecological system, and governmental as well as military applications. These monitoring systems need to transmit data continuously over sufficiently prolonged periods. However, it is not easy to get Wireless Sensor Networks (WSNs) to achieve such a widespread task over a prolonged period due to severe limitations in power resources. Hence, energy efficiency becomes the prime consideration, as well as the lifetime longevity of WSNs. The significant chunk of the collected data usually renders itself redundant and generally can be efficiently mined. Therefore, reducing or preventing altogether unnecessary transmissions in data communication has a very significant influence on reducing the overall energy consumption. Hence, these results in lifelong longevity of WSNs, such that the amount of energy spend on data transmission can be reduced [32].

Source coding is employed in data communication for efficiency, while symbol codes (e.g., Huffman codes) or stream codes (e.g., Arithmetic codes, Lempel-Ziv codes) are employed for source coding methods. On the other hand, Principal Component Analysis (PCA) is one of the well-known source coding method used to reduce or compress the dimensions of incoming sensor data [33]. Fundamentally, PCA is an orthogonal transformation technique that transforms a set of observations of interrelated variable values into a set of un-interrelated variables [34].

Further, channel coding is employed in data communication for reliability. However, Hamming codes, BCH codes, Turbo codes, and Gallager codes are used very often in channel coding. Further, the most dominant error correction code, which has a near Shannon limit presentation, is the Low-Density Parity Check (LDPC) codes. The LDPC coding technique is used for transmitting the data in a broadcasting channel. While the Shannon Limit determines the ultimate limit to the actual number of users or the maximum data rates, which is supported within the given limits of specified bandwidth and the power constraints [35].

Phase Shift Keying (PSK) and Quadrature Amplitude Modulation (QAM) are both analog as well as a digital modulation technique respectively, which in general requires a minimum of two phases and two amplitudes. Moreover, QAM is constantly used as a modulation strategy for computerized telecommunications networks, although high spectral efficiencies can be obtained by setting the QAM for an acceptable constellation scale, constrained only by the associated noise rates and the linearity of the transmission channels [36-37].

The Multiple-Input, Multiple-Output (MIMO) network is constructed utilizing numerous transmit and receive antennas and is currently under intensive development for its exceptional ability to support a substantial increase in capacity. The MIMO transmission takes advantage of the idea that in a scattering setting, in which the antennas are scattered and segregated in a fading network of multiple transmitting and multiple receiving antennas, the signals at various antennas undergo individual attenuation of their signal rates [38].

The modern communication system has several limitations like bandwidth, data rates, data loss due to error probability, spectral efficiency, and original data regression.

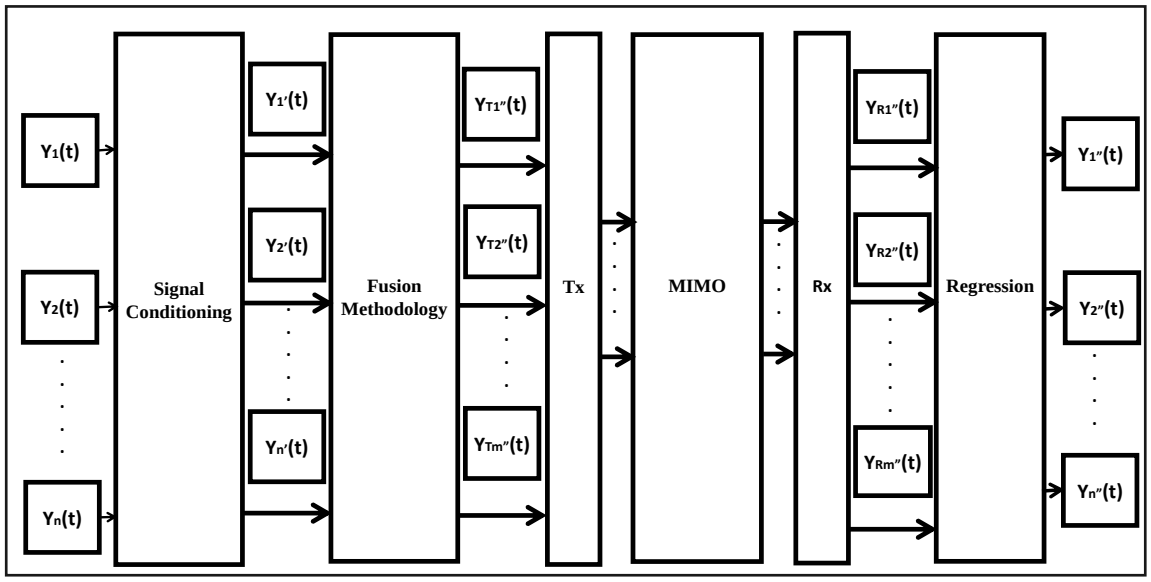


Figure. 1.1 Source coding for data fusion and compression over MIMO.

The thesis proposes data fusion over MIMO, as shown in Figure 1.1, collects the data, i.e., $Y_1(t)$, $Y_2(t)$…..$Y_n(t)$, at a time 't', from various sensors, which is then conditioned for a signal of appropriate formats, to generate $Y_1'(t)$, $Y_2'(t)$…..$Y_n'(t)$ signal. Thus, the formats are further fused and compressed at sensor level for transmission to generate $Y_{T1}''(t)$, $Y_{T2}''(t)$ $Y_{T3}''(t)$……..$Y_{Tm}''(t)$ signals, in which $m \leq n$, depends on the number of sensors (in this work, it is 'n') employed in fusion. The data compression is implemented using Principle Component Analysis (PCA). The compressed data is further modulated using proper modulation techniques, e.g., PSK, QAM, and it is then transmitted over the MIMO system having proper channel coding methods like LDPC, over a guided or unguided channel with Additive White Gaussian Noise (AWGN) characteristics.

The received signal is decoded for channel coding and demodulated to recover the original signal. The demodulated signal is $Y_{R1}''(t)$, $Y_{R2}''(t)$, $Y_{R3}''(t)$ …….. $Y_{Rm}''(t)$ at the receiver is then recovered using PCA reconstruct to regenerate the fused data, i.e., $Y_1''(t)$, $Y_2''(t)$….$Y_n''(t)$.

The data communication system is demonstrated for the data fusion at the sensor level for the number of signals forming an ensemble, which is spread over the range of frequencies from 50Hz to 2500Hz. These signals are fused using the well-known data reduction method, i.e., Principal Component Analysis (PCA) and regressed at the receiver end to generate the original signals. This system is demonstrated for the M-PSK/M-QAM modulation over the LDPC based channel coding error correction Min-Sum algorithm for the Multiple Input Multiple Output (MIMO) system. The performance evaluation of this system is computer over the Signal to Noise ratio from -15dB to +40dB for parameters like RMSE and BER. We have used MATLAB Simulink (2018b) for modeling the said system.

In due course of time, the works carried out in this thesis are listed below:

1. Modeled the sensor level data fusion and reduction using proper source decoding techniques like PCA on the MIMO channel for the data communication system.
2. Studied data consisting of several signals (10, 20, 30, 40, 50) at the sensor levels, using composite data called an ensemble, which is then employed to reduce the data using source coding technique like PCA using M-PSK, M-QAM modulation schemes, and LDPC channel coding with error correction over 2x2, 3x3, 4x4 MIMO Channels.
3. Synthesized the data communication system using for PCA data reduction and recovery, and BPSK modulation on Xilinx Vertex-6, LX240TFPGA.
4. Explored the application of above data transmission system in 2D, 3D, 4D mesh topologies for Network on Chip (NoC) implementation.

1.5. Thesis Outline

The entire work of this thesis is organized into six chapters, and the summary of each of these chapters are described as follows:

Chapter-1. Detail discussion on Data Compression, MIMO Channel, and PCA Source coding for data compression, LDPC channel coding background details are summarized. Further, motivations and contributions based on the previous works are detailed in this chapter, pertaining to the significant findings of this thesis.

Chapter-2. Presents the related work in the research area on data compression source coding, channel coding, modulation techniques, and MIMO.

Chapter-3. Presents the details of the PCA base source coding technique. A system detail of the PCA source coding technique to a data communication with PSK and QAM modulation techniques with MIMO is demonstrated. Verified the experimental results for parameters like RMSE and BER to evaluate the system performance.

Chapter-4. Introduces the LDPC channel coding technique along with the PCA base source coding data reduction methods. Further evaluate and discuss system performance with the modulation of 2, 4, 8, 16, 32-PSK and 4, 8, 16, 32-QAM over 2x2, 3x3, 4x4 MIMO with the parameters like RMSE and BER. Further presents the Xilinx system generator based synthesis of the discussed system with the BPSK modulation on Xilinx Vertex-6, LX240T FPGA.

Chapter-5. Explore the application of proposed data reduction system for the network on chip design. In this chapter 2D, 3D, 4D NoC topologies are introduced to evaluate the performance for throughput and latency over FTP and CBR applications.

Chapter-6. Summarize the thesis, and find future research scope for said work related to the robust data compression technique with the MIMO system.

1.6. References

[1] Castanedo, Federico. A review of data fusion techniques. *The Scientific World Journal*, 2013:1-9, 2013.

[2] Ren C. Luo, Chih Chia Chang, Chun Chi Lai. Multisensor Fusion and Integration: Theories, Applications, and it's Perspective. *IEEE Sensors Journal*, 11(12):3122-38, December 2011.

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

[4] Akyildiz, I.F., Su, W., Sankarasubramaniam, Y. and Cayirci, E. Wireless sensor networks: a survey. *Computer networks*, *38*(4):393-422, 2002.

[5] Gama, J. and Gaber, M.M. eds. *Learning from data streams: processing techniques in sensor networks*. Pages 41–48. Springer Science & Business Media, 2007.

[6] Tan, P. N. Steinbach, M. Kumar, V. *Introduction to Data Mining*; Addison Wesley: Boston, MA, USA, 2005.

[7] Aggarwal, C.; Yu, P. Outlier Detection for High Dimensional Data. In *Proceedings of the ACM SIGMOD International Conference on Management of Data*, Pages 37-46. Santa Barbara, CA, USA, 2001.

[8] Janakiram, D. Adi Mallikarjuna Reddy, V. Phani Kumar, A.V.U. Outlier Detection in Wireless Sensor Networks Using Bayesian Belief Networks, In *Proceedings of the First International Conference on Communication System Software and Middleware (COMSWARE 2006)*, Pages 1–6, Delhi, India, 2006.

[9] Li, Y. Anomaly. *Detection in Unknown Environments Using Wireless Sensor Networks*. The University of Tennessee: Knoxville, TN, USA, 2010.

[10] Jeffery, S.R., Alonso, G., Franklin, M.J., Hong, W. and Widom, J. Declarative support for sensor data cleaning. In *International Conference on Pervasive Computing*. Pages 83-100. Springer, Berlin, Heidelberg. 2006,

[11] Udit Narayana Kar, Debarshi Kumar Sanyal, An overview of device-to-device communication in cellular networks. *ICT Express*, 4(4):203-208, 2018.

[12] Iranmanesh, S. and Rodriguez-Villegas, E. A 950 nW analog-based data reduction chip for wearable EEG systems in epilepsy. *IEEE Journal of Solid-State Circuits*, *52*(9):2362-2373, 2017.

[13] Min, R., Bhardwaj, M., Cho, S.H., Shih, E., Sinha, A., Wang, A. and Chandrakasan, A. Low-power, wireless sensor networks. In *VLSI Design 2001. Fourteenth International Conference on VLSI Design*, Pages 205-210, 2001.

[14] El–said, S.A. Efficient DCT–based image compression technique. *International Journal of Signal and Imaging Systems Engineering*, *6*(4):259-274, 2013.

[15] Brittain, N.J. and El-Sakka, M.R. Grayscale true two-dimensional dictionary-based image compression. *Journal of Visual Communication and Image Representation*,18(1):35-44, 2007.

[16] Abbas, A. and Tran, T.D. Rational coefficient dual-tree complex wavelet transform: Design and implementation. *IEEE Transactions on Signal Processing*, *56*(8):3523-3534, 2008.

[17] Muthulakshmi G. *Neuro statistical space partitioning framework for compression of gray scale images*. Ph.D Thesis (Manonmaniam Sundaranar University, 2016. Shodhganga. Web, 2016).

[18] Nagarajan, A. *Investigation on various compression algorithms for gray scale and color medical images*. Ph.D Thesis (Madurai Kamraj University, 2013. Shodhganga. Web. 2013).

[19] P. Howard. Lossless and Lossy Compression of Text Images by Soft Pattern Matching. In *Proceedings of the 1996 IEEE Data Compression Conference*, Pages 210–219, 1996.

[20] Wu, M., Tan, L. and Xiong, N. Data prediction, compression, and recovery in clustered wireless sensor networks for environmental monitoring applications. *Information Sciences, 329*:800-818, 2016.

[21] K. Keerthi Vasan, B. Surendiran. Dimensionality reduction using Principal Component Analysis for network intrusion detection. *Perspectives in Science*, 8:510 -512, 2016.

[22] Q. Li, G. Li, W. Lee, M. il Lee, D. Mazzarese, B. Clerckx, and Z. Li. MIMO techniques in WiMAX and LTE: a feature Overview. *IEEE Communications Magazine*, 48(5):86–92, 2010.

[23] FPGA Implementation of MIMO Module at https://www.rfwireless-world.com/downloads/FPGA_Implementation_of_MIMO.pdf

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

[25] D. W. Bliss, Keith W. Forsythe, and Amanda M. chan. MIMO Wireless Communication. *Lincoln Labt. Journal*, 15(1):97-126, 2005.

[26] B. Muquet, E. Biglieri, A. Goldsmith, and H. Sari. An analysis of MIMO techniques for mobile WiMAX systems. *Advances in Mobile WiMAX*. New York: Wiley-IEEE Press, 2008.

[27] Mostari, L.and Taleb-Ahmed. A: Simplified a Posteriori Probability Calculation for Binary LDPC Codes, *Pertanika Journal of Science & Technology*, 26 (4): 1751 - 1763 (2018), 24 October 2018.

[28] Hengzhou Xu, Dan Feng, Bo Zhang, Huaan Li, and Hai Zhu. On the girth of tanner (3, 13) Quasi-Cyclic LDPC Codes. *IEEE Access*, 7: 5153–5179, 2019.

[29] Jingjing Wang, Na Li, Wei Shi, Yangyang Ma, Xiulong Liang, and Xinli Dong. Capacity of 60 GHz Wireless Communications Based on QAM. *Journal of Applied Mathematics*, 2014: 1-5, 2014.

[30] Ogras, U.Y., Hu, J. and Marculescu, R. Key research problems in NoC design: a holistic perspective. In *Proceedings of the 3rd IEEE/ACM/IFIP international conference on Hardware/software codesign and system synthesis*, Pages 69-74, 2005.

[31] B. M. K. Potti, M. V. Subramanyam, and K. Satya Prasad. Hybrid genetic optimization to mitigate starvation in wireless mesh networks. Indian Journal of Science and Technology, 8(23):1–10, 2015.

[32] Udit Narayana Kar, Debarshi Kumar Sanyal. An overview of device-to-device communication in cellular networks. *ICT Express*, 4(4):203-208, 2018.

[33] Mou Wu, Liansheng Tan, Naixue Xiong. Data prediction, compression, and recovery in clustered wireless sensor networks for environmental monitoring applications. *Information Sciences*, 329:800-818, 2016,

[34] K. Keerthi Vasan, B. Surendiran. Dimensionality reduction using Principal Component Analysis for network intrusion detection. *Perspectives in Science*, 8:510 -512, 2016.

[35] Hengzhou Xu, Dan Feng, Bo Zhang, Huaan Li, and Hai Zhu. On the girth of tanner (3, 13) Quasi-Cyclic LDPC Codes. *IEEE Access*, 7:5153–5179, 2019.

[36] D. W. Bliss, Keith W. Forsythe, and Amanda M. chan. MIMO Wireless Communication. *Lincoln Labt. Journa*l, 15(1):91-126, 2005.

[37] Jingjing Wang, Na Li, Wei Shi, Yangyang Ma, Xiulong Liang, and Xinli Dong. Capacity of 60 GHz Wireless Communications Based on QAM. *Journal of Applied Mathematics*, 2014:1-5,2014.

[38] Imran Khan, Joel J. P. C. Rodrigues, Jalal Al-Muhtadi, Muhammad Irfan Khattak, Yousaf Khan, Farhan Altaf, Seyed Sajad Mirjavadi, and Bong Jun Choi. A Robust Channel Estimation Scheme for 5G Massive MIMO Systems. *Wireless Communications and Mobile Computing*, 2019:1–8, 2019.

Chapter 2

Related Work in the Research Area

2.1. General Information

The purpose of source coding is to represent information as accurately as possible using a few bits and eliminate redundancy from the source. In a certain way, the purpose of channel coding is to introduce redundancies, which can be used to protect information when transmitted through a non-ideal channel. The combining of these two techniques leads to conventional source and channel coding, enabling the design of a data communication system to generally achieve better performance than when the source and channel codes separately designed. This research explores the source and the channel coding, i.e., MIMO channel PCA and LDPC.

Most people around the world use apps in daily life, arising from what is known today as the fields of source coding and channel coding. For example, the applications may be compact disks (CDs), cell phones, MP3 players, digital versatile disks (DVDs), digital TV, voice over IP (VoIP), video streaming. Hence, leading to considerable interest in the field of collaborative source and channel coding, usually makes it easier to enhance the efficiency of source and channel coding by collectively developing these fundamental building blocks rather than considering separate units. However, these sources and channel coding is of concern when it comes to information delivery, i.e., data. Here $X_n = (X_1, X_2, \cdots, X_n)$ is a sequence of source data, e.g., from the processing of continuous signals, which is to distribute over a channel.

Furthermore, the information in X_n represents the transmit over the channel. It is relevant that the receiver will decode the information received and produce the value of $X'_n = (X'_1, X'_2, \cdots, X'_n)$ from the original details. The encoder's task is to create an X_n representation, and the decoder's task is to provide the X'_n approximation based on the source obtained. It is obvious to know the source and channel encryption, on how the encoders and decoders are designed [1].

2.2. Data Compression (Source Coding)

Besides, dealing with the transmission of information, it is appropriate to interpret using a discrete value. The source coding represents the information as efficiently as possible as few bits, based on ‘compression’. The compression is mainly divided into two categories: lossless coding, which allows a source coded version of the original data to be adequate to replicate the same version of the original data. Hence such data is no longer needed when working with lossy coding, wherein the goal is to restore instead, an approximate version of the original data that is as good as possible [1].

The source coding (data compression) and channel coding (error control) are built independently of each other. It is explained in Shannon's theory of separation. However, new developments in the world of information technology have resulted in an immense amount of data generated every second. As a result, the data collection and delivery infrastructure expected to increase to an immense extent. According to Parkinson's first rule, the demand for storage and transmission decreases at least twice as storage and transmission capacity decreases. While optical fibers, Blu-ray, DVDs, Asymmetric Digital Subscriber Line (ADSL) and cable modems are available, the pace of data growth is far higher than the rate of technology growth. Hence, refuses to resolve the problem as mentioned above based on the management of massive data storage and transmission. Further, in order to address this problem, an alternative concept of data compression (DC) is proposed to minimize the scale of the data stored or transmitted [1].

Currently, DC techniques are essential for most real-time applications such as satellite imaging, Geographic Information Systems (GIS), visualization, Wireless Sensor Networks (WSN). While the advancement of recent technology significantly improves the quality of data, as it ultimately raises the size of the data. Therefore, reducing the size of the file allows more information to store in the same data space

with less transfer time. Hence, it is complicated and sometimes hard to store or contact a vast number of data files without DC [2].

Moreover, the fundamental DC concepts are lossless compression, loss compression, and text compression, compression of images, audio compression, video compression, bit reduction, and data redundancy. Hence, the performance of the DC system is measured in many ways, by measuring the complexity of algorithms, computational memory, speed, amount of compression, and quality of data recovered. Further, the compression ratio (CR) is the most common metric to quantify the performance of data compression. Hence, the following equation 2.1 shows the ratio of the total number of bits necessary to store uncompressed data to the total number of bits needed to store compressed data [2].

CR=No. of bits in uncompressed data / No. of bits in compressed data (2.1)

The CR is termed as bit per bit (bpb) [2], and the average number of bits requires storing the compressed data as specified. Similarly, in image compression techniques, bpb is called bits per pixel (bpp). In contrast, modern text compression methods use bits per character (bpc), which reflect the average bits required to compress a character. There is also another metric called Space Savings that describes file size reduction compared to uncompressed data and expressed in equation 2.2.

Space savings= 1 – (No. of bits in compressed data / No. of bits in uncompressed data) (2.2)

For instance, when the original data is 10 MB, and the compressed data is 2 MB, then space savings will be (1.0-(2/10)) = 0.8. A value of 0.8 indicates that 80% of storage space saved due to compression. Next, the compression gain is equated in Equation 2.3,

$$Compression\ gain = 100 log_e (original\ data / compressed\ data) \quad (2.3)$$

The compression speed can be measured according to cycles per byte (CPB). It implies, on average, the number of cycles needed to compress one bit. CR and compression factor (CF) are relatively good to calculate lossless compression techniques. As the recovered data in lossy compression significantly differs from the original data, certain additional performance measurements are required to determine the degree of degradation, accuracy, and efficiency. The difference between replicated data of lossy compression from the original data is considered distortion. Other words like fidelity and accuracy also reflect the discrepancy between original and reconstructed data [2]. A common metric used for this entity is peak signal to noise ratio (PSNR), a dimensionless number used to differentiate the obvious blunders which can be detected by the human eye. PSNR usually revealed to the logarithmic decibel (dB) scale. PSNR equated as equation 2.4 for original information X_i and the replicated information Y_i.

$$PSNR = 20 log_{10} (max|X_i|/RMSE) \quad (2.4)$$

Where RMSE (Root mean square error) is the square root of Mean square error (MSE) and formulated in Equation 2.5,

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - Y_i)^2} \quad (2.5)$$

If the initial and reconstructed data are precisely the same, then RMSE is zero, and SNR is infinite. The value of RMSE will be small, and the SNR value will be strong for the improved similarity between the original and reconstructed results. Another formula relevant to this is equation 2.6.

$$SignaltoNoiseRatio(SNR) = 20log_{10}\frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n} X_i{}^2}}{RMSE} \qquad (2.6)$$

SNR is used for quantifying the magnitude of the signal related malfunction. While SNR numerator is the root mean square of the original information. However, PSNR is the most effortless and commonly utilized method, as it disposes the survey conditions of influencing the signal information quality and may not be correlated with human perception [2].

In general, a DC methodology affects the accuracy of the data depending on the specification parameters to some degree. If the DC techniques are used for general purposes such as messaging or internet searching, so the accuracy of recovered data is not taken into account. But, in-text compression, it is not tolerable to modify a single character in a text document, as it changes the whole meaning. The importance of a DC techniques data quality is strongly dependent upon the form of data. The DC techniques can split into lossy and lossless compression, based on the criterion of restored data quality. However, lossless compression applies to no knowledge loss, i.e., the data recovered is consistent with the original data. It is used in applications where information loss is not needed, such as text, medical imaging, law forensics, military imaging, satellite imagery, etc. [3].

In some cases, lossy compression methods are preferred, where the restored data is not precisely identical to the original data, and it is often appropriate to estimate the original data. Compared with lossless compression techniques, this results in higher CR with lossy compression techniques. Near-Lossless compression techniques are another form of compression technique where the discrepancy between the original and restored data is guaranteed to vary by no more than a user-specified sum called Maximal Absolute Distortion (MAD) from the respective values in the original data [4]. Hence, used for the compression of signal data, hyper spectral pictures, videos, etc.

The effective representation, computational complexity, and resource management are the characteristics of every DC technique. Recently, there has been a considerable regular rise in the amount of data production, and different file formats are now being added. Firstly, the correct representation of data during compression often affects the efficiency of a compression process, and there are different types of demonstration, such as matrix format, configuration, chain code, and quantization. Hence, an incorrect data representation may add larger compressed file than the original file. At the same time, the correct way to represent data can give a better CR [2].

Moreover, the advent of emerging technology and analytics in data processing enabled the organization of big data in processes as a revolutionary feature of wireless sensor networks (WSNs). Big data model, combined with WSN technology, implies new problems that need to be tackled in parallel. Hence, data collection is a fast-growing in the field of study. While, big data processing is a critical problem in wireless sensor networks that needs efficient approaches to capture, interpret, store, and compile vast quantities of data. However, the collection of broad bandwidth utilization is essential to allow the use of specialized techniques to tackle this task, so the sensor-gathered data need analysis and storage. Methods of research are required for managing the growing quantities of data at the same time. They do need to be changed, to reduce response time and to save more energy to prolong the lifespan of the network. Further, the aggregation of data in large-scale wireless sensor networks is an essential technique. It is an efficient method for processing high volume scale by integrating related data, removing the duplicate data issue, and thereby growing the resource consumption [5].

In recent years, enormous efforts have been made in the field of research of data reduction in sensor networks. However, among these specific methods, the data anticipation and data fidelity are the most appropriate groups for creative construction. In the data forecast, the sensor hubs do not continuously relay their determined quality

to the base station and enable the base station to predict the predefined values. In this way, the base station regularly holds up its estimation of the announcement rather than the real expense [6].

Most compression algorithms [7-10] are used to encrypt information at source nodes, and then decode it on the sink where consistency is a significant concern. Goel and Imielinski [7] used Moving Picture Experts Group (MPEG) compression techniques for energy-efficient sensor network monitoring. They suggested a prediction-based monitoring mechanism (PREMON), which is best-known for submitting only the difference between the current picture and the previous picture rather than the entire present image. MPEG is used to resolve the significant divergence created by moving objects between successive images in the film. In [8], the authors proposed a multi-target evolutionary algorithm to generate a set of variations of parameters of the quantification mechanism corresponding to various equilibrium interconnections between compression and information loss. Therefore, the user should pick the combination with the best value for the individual user. The methods were introduced in-network data reduction to [9, 10] incorporate data compression and routing. The routing driven compression (RDC) performs the function of data compression on the shortest path to the sink, in which the compression based routing (RDC) does not automatically carry out the highest possible compression on the shortest path to the sink. The combined routing compression strategies RDC and Compression Driven Routing (CDR) will both be highly dependent on the routing algorithm efficiency and the data correlation stage. For the Internet of Things, a simplified sensing-based framework [11] was proposed in which the final nodes calculate, transfer, and store the sampled data. Then an effective cluster-sparse restore algorithm for in-network compression is used with the goal of more reliable data restoration and lower energy consumption.

Reduce spatial redundancy among neighbouring nodes within each cluster can be obtained with the principal component analysis (PCA) technique. The explicit goal behind this strategy is to provide an optimized solution to the data transported from the cluster heads (CHs) to the base station (BS). It requires a linear transformation of a series of associated variables into an arbitrary set of variables, i.e., the aggregation of data while preserving the initial data set properties without any data loss. According to its basic and non-parametric characteristics (parameters can be computed), the PCA has widespread implementations in wireless sensors networks such as outlier detection [12], information processing, and machine learning [13]. Another significant advantage of the PCA is the optimum linear method for projecting a large data set into a lower-size space to expose the hidden data (in terms of mean square error).

In [14], the paper suggests a low cost, loss-based compression solution with an error-bound guarantee for WSN application. However, the approach also indicates the reduction of data compression and minimizes the dissipation of resources. While the use of Spatio-temporal similarity between data samples achieves this, it is based on neural networks, an algorithm in machine learning to simulate human and environmental situations automatically. This approach eliminates constraints on power and bandwidth while gathering data in the tolerable error range. The algorithm is evaluated using meteorological data sets and provides better results than CR, RMSE, for Fast Fourier Transform (FFT), Principal Component Analysis (PCA), and Discrete Cosine Transform (DCT).

In [15], a novel energy-efficient architecture for distributed WSN is incorporated into the data estimation, compression, and regression techniques. However, the PCA technique is used for data compression and regression. Further, the sensor nodes organized into clusters in the clustered WSN and concurrent double prediction process at sensor nodes and cluster heads are introduced, for cluster heads (CHs) which uses

PCA method to filter the main part and to remove redundant data. So, the data at the base station is restored entirely to further check with a sample of open sensors. However, the findings show the approach for continuous control of WSN applications is effective. Also, to test the efficiency, the algorithm is used for prediction precision, convergence, and improving communication speed. The compression algorithm uses the efficiency metric as connectivity costs and increases both the accuracy and the convergence rate.

The motivation for the present research is data compression for the sensor level data using data reduction approaches for data transmission in the communication system. The necessity of a methodical literature survey has been recognized after considering progressive research in data reduction for data compression in the communication system. Therefore, based on broad and methodical search in existing works, the available research is summarized, and research challenges for future research are also presented.

2.3. Channel Coding

When the source coding executed, one aims to delete the highest redundancy in the source data, and redundancy is inserted into the data to protect data from channel errors for channel coding. For example, these channel errors can be continually assessed, packet losses, or bit errors [16].

The channel encoding and decoding are intended to detect and correct noisy signal errors. While the noisy channel errors that occur during data transfer from the source to the destination, there arise the need of mechanism for detecting and remedying these errors. A communication system is a medium by which information can be conveyed from one individual to another, and the digital communication is a device which physically transmits the data. Further, source coding is used to reduce the source of data

and minimize redundancies for effective transmission from source information. While the signal power and channel bandwidth for transmission are key parameters for the configuration of the digital communication system, these parameters calculate the signal energy per bit (Eb) to noise intensity spectral density (N_0) ratio. This ratio is special to the likelihood of a bit error, also called a Bit Error Rate (BER). In reality, an appropriate BER with channel coding is feasible for a given $E_b = N_0$. It can be done by adding additional numbers to the data source. In which, additional digits do not have new information but allow the receiver to identify and correct errors, reducing the overall risk of a mistake [16].

In a Shannon's Noisy Channel Coding Theorem, if any noise-influenced channel has a set channel capacity C, a rate of information transmission can never be surpassed without error, but an error-correcting code still exists such that information can be transmitted at speeds below C with an arbitrarily low BER. The Rayleigh fading channel also tests the BER performance of various diversity techniques such as selective integration, equal-gain combining (EGC), and maximal ratio combining (MRC) [16].

Dr. Hamming introduced the idea of error-correcting codes in 1947 and created the first error-correcting code called Hamming code. Later, revealed a paper [17] in Bell System Technical Journal on error detection and error correction codes. In 1948, C. E. Shannon presented the Mathematical Theory of Communication and published a paper [18] in the scientific journal Bell Method. In 1949, J. E. Golay wrote an article [19] on computer coding in the IEEE proceedings and drew up a Golay Code (23, 12). In their logic design paper, Dr. Reed [20] and Dr. Muller [21] identified a new code for error correction in 1954. Dr. Reed also provided the algebraic ring theory [22], the Galois field, and the polynomials [23] over fields.

In 1957 Dr. Prange discovered cyclic codes at the Cambridge Research Institute for Air Force, and then in 1958, he developed quadratic residue code. Further, in 1960

Alexis Hocquenghem et al. discovered binary cyclic BCH codes [24]. They first implemented a cyclic code word shift, which makes encoder and decoder design easier. In 1960 Irving S. Reed and Gustave Solomon invented another cyclic code called Reed Solomon (RS), which has a large Hamming distance. The encoding method is similar to cyclic codes. Elwyn Berlekamp and James Massey [25] developed an efficient decoding algorithm for RS codes. This code is more common and used in numerous applications, with the advent of powerful computers and a more effective decoding algorithm. However, Reed-Solomon codes are used in satellite communications, transport, and space communication. While these are non-binary cyclic codes, with two GF(2) symbols. RS code (255, 223, 33) is a NASA standard code which has 33 Hamming distance and can correct up to 16 errors.

Similarly, several authors [26] [27] [28] [29] have published research works on this code. Binary Quadratic Residue (QR) codes are the most effective recognized cyclic codes, but it is challenging to decode this code. The block size of the QR code is selected according to $n=8t \pm 1$, where t is the code's capacity to fix errors. Various researchers have recently suggested several decoders [30]-[34].

In 1962, Robert G. Gallager suggested the Low-Density Parity Check (LDPC) series of codes with some technical limitations and, as a result, these codes have not been in operation for 30 years, but later emerge as one of the most effective codes [35]. In 1977, V. D. Goppa proposed a startling new development of codes, called Goppa Codes. This code is surprising because of its asymptotic dependent minimal distance properties. Goppa code is based on algebraic-geometric methods that can measure the minimum distances [36].

In 1993, Claude Berrou, Alain Glavieux, and Punya Thitimajshima developed Turbo codes, which have output quite similar to its channel capacity [37]. David J.C. MacKay and Radford M. Neal later in 1996 or later developed LDPC codes, which are

also similar to channel efficiency output. Further, in 1997 Madhu Sudan developed an algorithm for list decoding [38] with an enormous improvement of linear block codes. This algorithm allows for more than one potential code word as performance but does not transmit more errors than half the code's minimum size. Rafid Mahmood, [39] introduced a new convolution code family with Hamming distance column property. In this, distance property specifies further application distance for the error correction capability. They attempted to raise this distance by streaming across a link loss network. Irina E. Bocharova, [40] developed a low BER and measurement delay low-density parity-check (*LDPC*) convolution codes (LDPC-CC) module. They focused on various deployment problems, such as low encoding or decoding difficulty and fast decoding time.

Mengqi Zhou W. Ross, [41] researched convolution and turbo codes and explored the concepts of encoding and decoding. They built up an Acquisition of Signal (AOS) connectivity simulation model for convolution and turbo code. However, the effects of these codes are contrasted with three aspects: bit error rate, frame error rate, and efficiency. They evaluated and shown that the Turbo code works better for an SNR value than a convolution code.

Mohamed H. Omar, [42] an author, has developed the Long Term Evolution (LTE) device Tail-biting Convolution Code Decoder (TB-CCD). This method used to encrypt the channel on two downlink lines. They focused on time sequencing and thorough analysis. In 1ms, a maximum of 44 decoding attempts per LTE sub frame needed for TB-CCD. In the present study, various facets of the processing time saving and architecture for selected TB-CCD design were identified. The decoding strategy for reducing storage requirements and decoding delays was suggested by Chris Winstead [43]. BER comparisons of LDPC block code and LDPC convolution codes are supported, assuming equivalent hardware difficulty and decoding time. They developed

a terminated LDPC convolution code for broad frame transmission. Hosseinzadeh et al. [44] researched sources and channels coding strategies and lossless compression technology needed for the efficient transmission of digital mammography over an AWGN wireless platform. They addressed schemes with the lowest bit error rates to ensure the effective transference of data over a noisy wireless channel.

In an eminently implementable program with a linear decoding difficulty in terms of block duration, the solution by Berrou was successful. While, the subsequent search to understand the theory behind this puzzling results, the coding of Low Density Parity Checks (LDPC) originally introduced by Gallager in his PhD thesis [45] has been discovered in [46], with comparable features. These are both approaches that have become the workhorses of modern communication practices, with debates over one's technical advantages over the other largely over shadowed by the desires of companies and standardization. The evident and incontestable is that LDPC codes are simpler to clarify and interpret and will thus generally dominate over turbo codes. Currently, it is well established by both LDPC and turbo codes which are viewed as rare graphical codes. As a consequence, they share a variety of features, and design or research process that can be used with one will usually be repeated for the other.

The complexity of turbo and LDPC codes has placed capacity-based output in the scope of implementable systems, and it is not inherently realistic to implement them. However, the design of code that performs well under realistic restrictions such as a short encoding period and a strong spectral quality remains a significant barrier to low power distribution in integrated circuits. Hence, the methods to simplify code design, creation, storage, and decoder implementation are urgently required.

Digital video broadcasting adaptation cable, Digital Video Broadcasting - Cable (DVB-C), and digital video broadcasting-satellite (DVB-S) began in 1994. These systems used the first modulation scheme, which was QPSK, which uses Reed-Solomon

FEC and convolution coding as a hybrid error management system. In 1997, the DVB format for optical satellite news collection was accepted for coverage purposes. While, Broadcast vans were used to relay live satellite signals at big public events, in the studios from the outside. DVB DSNG uses 8PSK and 16QAM. However, in the year 2003, the DVB Project developed DVB-S2, a second-generation satellite broadband specification that improved the DVB-S capacity by approximately 30 percent [47]. The system was designed as a toolkit that enables the implementation of certain satellite applications, such as TV and sound, interactivity (that is, internet access), and professional services such as digital satellite news collection. Hence, three concepts were developed to this framework that included rational receiver complexity, maximum transmission efficiency near to the Shannon threshold, and complete flexibility [48] low-density parity codes (LDPC), channel coding technique, as well as QPSK, 8PSK, 16PSK, and 32PSK modulation schemes.

Further, the framing structure allows maximum flexibility in a versatile system and consistency in worst cases settings (low signal-to-noise ratios). The configuration of each user's transmission parameters in single-to-one links is made possible by adaptive coding and modulation (ACM), depending on path conditions. However, due to the availability of backwards-compatible modes, the new DVB-S integrated receiver – decoder can be used during the transition period [47, 49, 50]. In [51], the authors analyzed the increase in DVB-S2's spectrum efficiency using the hierarchical modulation of QPSK and the 32APSK. Their results show that QPSK modulation increased spectrum efficiency by 12% for a poor channel condition, i.e., 2dB SNR, while was 7% at 11dB SNR for 32APSK. In [48], the efficiency of LDPC codes for the DVB-S2 system was tested with the AWGN and Rayleigh fading channels [52] [53].

The low-density parity-check (LDPC), Gallager codes are defined by sparse matrices, usually with a random design. Shannon performed such codes by decoded

using a probabilistic algorithm. The low-density parity check codes are often used for communication across networks for additive (substitution) errors. However, LDPC codes can also be expanded through Davey et al. non-binary channel coding sources [54]. Besides the codes showed that there was a 0.6 db signal-to-noise increase for a given bit error rate. Liveris et al. proposed the use of LDPC protocols for open-source coding [52]. They also developed a distributed source code scheme in the log domain for binary sources.

Current LDPC implementations can be regarded as solving the problem of capacity for elementary point-to-point networks. In recent years, information analysis has progressed on many multi-user networks such as multiple access, streaming, transmission, and intervention platforms. The theory has also grown to involve network compression and shared source and channel coding, distributed storage, net computing, and quantum channels and protocols, outside direct communications.

2.4. M-PSK/M-QAM Modulation

Modulation is a mechanism in which digital data cannot be directly sent through transmission networks as the modulation makes it possible to modify the message signal into a form suitable to the characteristics of the channel [55].

The modulation is a remarkable technique of transmitting data on the radio network, which is essential to all wireless communications [56]. Most wireless communications today are digital, and its low bandwidth allows the modulation process to be essential than ever. The fundamental aim of modulation is to squeeze as much data as possible onto the lowest bandwidth available. However, it is defined as a spectral efficiency, which determines how quickly the data can be transmitted within an allotted bandwidth.

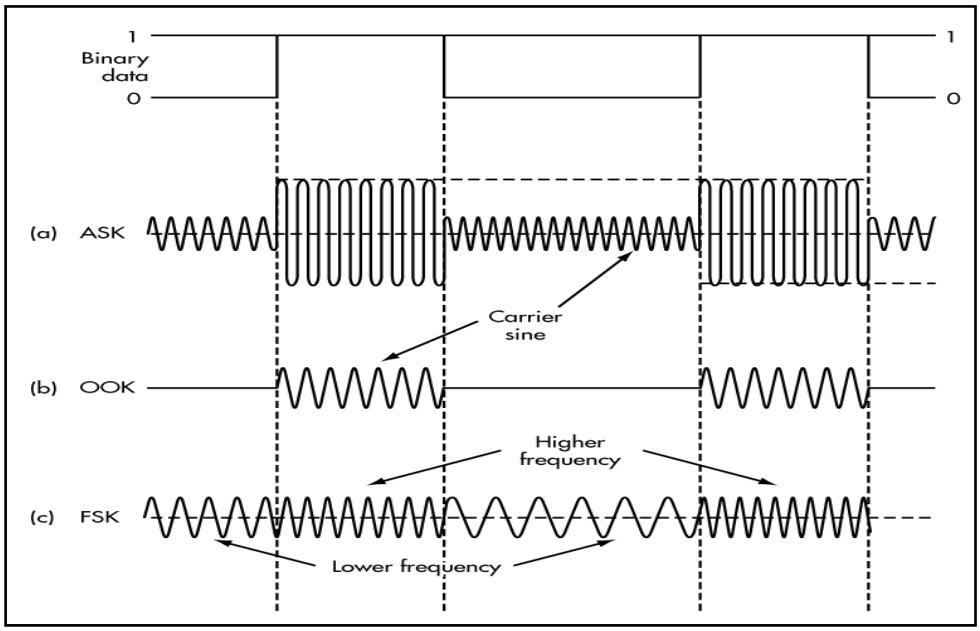


Figure. 2.1 The low-data short-range wireless systems also use three simple optical modulation formats: (a) ASK. (b) on and off. (c) FSK.

And the unit is calculated as bits per second per Hz (b / s / Hz). Hence, the multiple approaches to achieve and improve spectral performance have arisen. However, there are three essential ways to change the amplitude, frequency, and phase of a sinusoidal wave radio signal. More specific approaches are combined with increasing spectral efficiency by two or more of these combinations. These simple modulation types still are used in digital signals today.

The output waveforms are consistent as shifts arise at carrier zero-crossing points in binary configuration. Figure 2.1 shows simple serial digital signals for binary and transmitted zeros and associated modulation-based amplitude (AM) and frequency (FM) signals. There are also two forms of AM signals, on-off keying (OOK) and ASK (Amplitude Shift Keying). In Figure 2.1 (a), the carrier amplitude is adjusted between two amplitude stages. Also, the binary signal switches off the carrier and starts to generate OOK in Figure 2.1 (b). AM which produces sidebands above and below the carrier equivalent to the maximum frequency output of the module. The required

bandwidth is, therefore, twice the maximum frequency, including any harmonics for binary pulse modulation [56].

The FSK transmits the carrier between the two separate frequencies called the frequencies mark and space, fm and fs (Figure 2.1 (c)). These changes are referred to as FSK phase. Nevertheless, FM creates several sideband frequencies above and below the carrier point, as the produced bandwidth is centered on the highest frequency of modulation, including harmonics and the modulation index given by equation 2.7 [56].

$$m = \Delta f * T \qquad (2.7)$$

$$\Delta f = fs - fm \qquad (2.8)$$

Δf is the frequency deviation or shift between the mark and space frequencies, as per equation 2.8. T is the bit time interval of the data or the reciprocal of the data rate (1/bit/s). Furthermore, there are two approaches to increase spectral performance in both ASK and FSK by switching from one binary state to another, choose data rates, carrier frequencies and change frequencies. Hence, there is no discontinuity within the sinusoidal network as these discontinuities produce the bandwidth and harmonic output glitches [56].

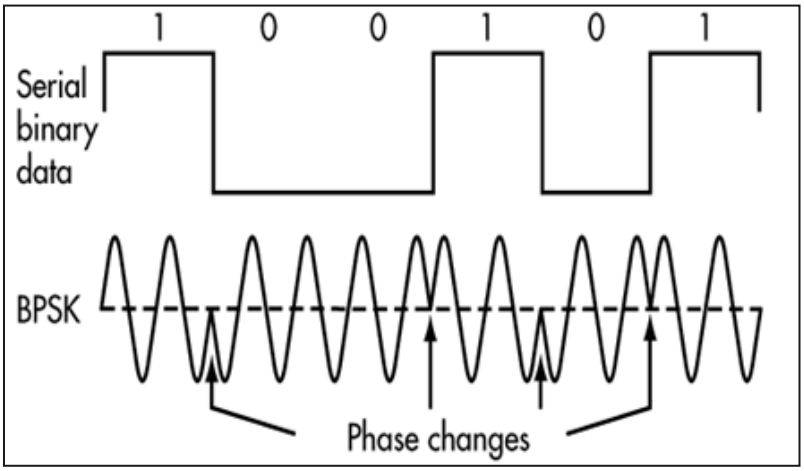


Figure. 2.2 Binary phase-shift keying modulation.

The theory is based on the stop and start times of the binary data which should be synchronized at zero-crossing points in amplitude or frequency before the sine carrier phase. However, it is also called a continuous phase or a consistent operation as all the cohesive ASK / OOK and secure FSK have reduced harmonic bandwidth and fewer than opposing signals. A second approach is the pre-modulation filtering of binary data which also rounds off the signal, by decreasing the intervals of rising and fall, and also reducing the harmonic output. Hence, the unique Gaussian and elevated low pass cosine filters have been used.

Binary phase-shift keying (BPSK) shifts the carrier sine wave 180° at each binary state transformation (Figure 2.2). BPSK is reliable since phase changes take place at zero-crossing points. In conjunction with a sine carrier of the same phase BPSK must be demodulated correctly. Thus, it includes carrier recovery and other complex circuits [56]. BPSK is very spectrally efficient as the bandwidth data rate, or 1 bit / Hz, can be transmitted. Further, a common BPSK variant, which is the quadrature PSK (QPSK), the modulator, generates two sine carriers separated by 90 degrees. As, the modulated binary data produces four different sine signals in each phase, which are shifted by 45° from each other. Then they merge the two phases to create the final signal while every single pair of bits produces a different phase carrier.

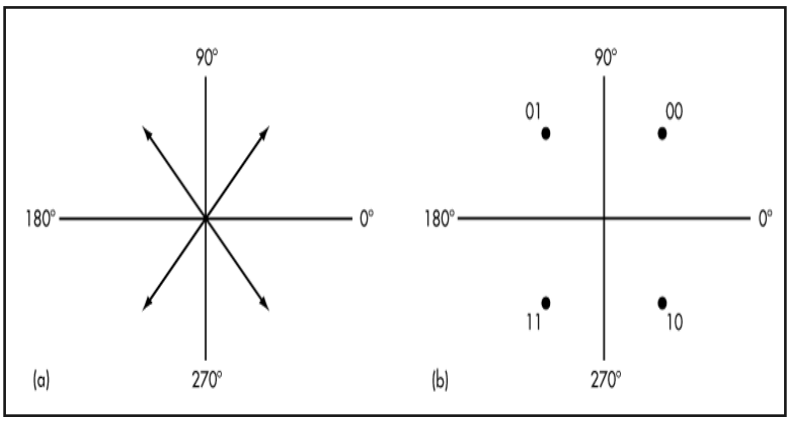


Figure. 2.3. (a) QPSK constellation diagram, (b) QPSK phases and amplitudes.

Figure 2.3 (a) displays the QPSK in a phaser diagram, where a phaser is the transmitter's sine amplitude value. The same description is observed in Figure 2.3 (b), which reveals in a constellation diagram. The QPSK is spectrally efficient, representing two bits of data in each carrier phase. And the spectral efficiency is 2 bits / Hz, indicating the data rate can be obtained twice as high a BPSK.

The maximum possible data rate or channel capacity (C) in bits / s is a feature of the channel bandwidth (B) in Hz, and the signal-to-noise ratio (SNR) is given,

$$C = B \log_2 (1 + SNR)$$

It is called the Shannon-Hartley Law. The standard data rate is directly proportional to the SNR bandwidth. Noise significantly decreases a specific bit error rate (BER) in the data stream.

The important aspect is the baud rate or the number of modulation symbols emitted per second. However, the term modulation symbol refers to a distinctive feature of a sine carrier signal. It can be a group, intensity, duration, phase. The BPSK uses a binary shift from 0° to 0, and a 180 ° to 1. Data rate in bits/s is calculated as the reciprocal of the bit time (tb):

bits/s = 1/tb

The baud rate is equivalent to the bit rate, with a symbol per line. Nevertheless, the propagation of more bits per symbol will slow the baud rate down by a factor proportional to the number of bits per symbol. QPSK produces two bits per symbol, making it very spectrally efficient. Also, the two current versions are 8-PSK and 16-PSK, 8-PSK, which utilizes eight symbols of frequent 45° adjustments between them, allowing for the transfer of three bits for each symbol, 16-PSK also utilizes the constant-width carrier signals of 22.5°, this design conveys 4 bits per symbol [56]. While Multiple Phase Shift Keying (M-PSK) is more effective in bandwidth than

increasing the amount of smaller phase changes, the easier it is to demodulate the signal in the presence of noise. The downside of M-PSK is that it provides more robust nonlinear power amplification that can be used with the constant amplitude of a load.

The development of symbols, which balance amplitude with duration, may help endorse the idea of more bits per symbol being transmitted. This approach is called quadrature amplitude modulation (QAM) For example, 8QAM uses three-bit transmission of symbols with four carrier phases, plus two amplitude stages. Other popular versions include 16QAM, 64QAM, and 256QAM, conveying 4, 6, and 8 bits per symbol (Figure 2.4) [56] respectively.

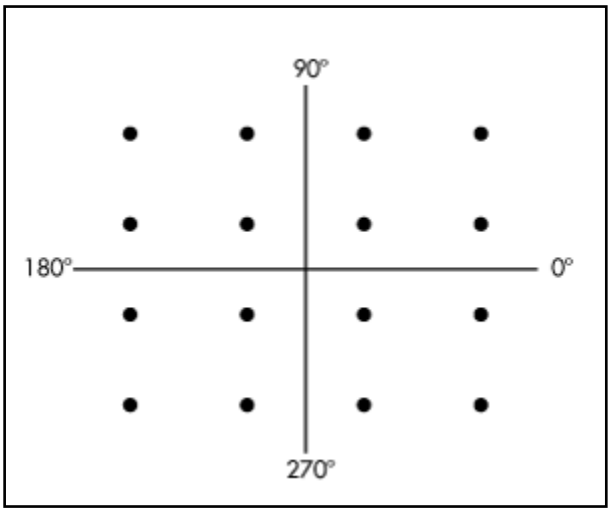


Figure. 2.4. 16-QAM constellation diagram.

Moreover, QAM is incredibly spectral efficient, in the presence of noise, which is mainly random amplitude shifts, it is more difficult to demodulate as the amplification of linear power is also necessary. The QAM is commonly used in cable TV, Wi-Fi, LAN, mobile and cellular telecommunications networks to achieve the maximum speed data rate [56].

Wireless communication has become a new area in our modern lives that is rapidly growing and has a significant impact on nearly every aspect of our daily lives.

Digital modulation aims to improve mobile connectivity by increasing wireless network coverage, speed, and capacity. In communication, the idea of modulation is a primary factor because, without an appropriate modulation scheme, a planned flow could not be achieved. While, in the construction of communications networks, it is essential to take note of the bandwidth, the allowable power, and the amount of the system's inherent noise. Due to the error-free capability of digital modulation, it is preferred over analog modulation techniques. The WiMax utilizes combinations of various modulation schemes, including M-PSK, M-QAM, and offers high bandwidth, multimedia, and data services [56].

To meet the increased demand for quality services, the latest generation of wireless communications systems needs a higher data transmission rate [57]. Scientists and engineers have also faced the difficulty of communicating effectively over a great distance. While the change to analog and digital modulation offers increased information capacity, improved network synchronization protection for digital data services, better transmission efficiency, and a quicker device uptime [58]. Due to the modern communication network that is more consistent than an analog method [59], however, a significant transition from analog to digital technology has occurred during the past decades and can be seen in all technology fields. In addition, data protection, sharing of the RF spectrum and higher quality communication capacity to provide additional services are supported by digital modulation [60-62]. The digital modulation is preferred over analog modulation due to digital modulation, while the requirement over wireless audio, video, and data on cell telecommunications network or the internet network defined as third-generation cell communication presents a crucial problem with bandwidth such that current modulation schemes need to be changed to accommodate a new bandwidth performance implies whether capabilities are used for the available bandwidth or the ability of a modulation scheme to take data into account within a

limited bandwidth [63-64]. Digital modulation provides significant advantages over analog modulation due to the distinct propagation condition. The analog signal may be clearly sensed in a noise detector, which means an infinite number of values.

An analog signal is modulated with binary code in the optical modulation techniques in which, the optical modulator interface between the transmitter and the screen. The optical modulation can be primarily graded based on its compaction bandwidth capabilities. While, the key requirements for the right method of modulation depend on the signal to noise ratio (SNR), power supply performance, usable bandwidth, better service quality, bit error rate (BER), and usability. Every modulation method's efficiency is determined by calculating the likelihood of error on the basis that the device operates on the additive White Gaussian Noise [65]. However, the modulation schemes that are capable of distributing extra bits per signals are an external immune error caused by channel noise and interruption [66]. Hence, the delay distortion may be a significant element in deciding the modulation process for digital radio [67].

The pulsed PSK, ASK and FSK form an integral technique of digital modulation with the Nyquist pulse type [68-70]. However, the approaches are often possible by combining two or more digital modulation techniques with or without the inclusion of pulse formation in the database. Further, the modulation may be built to hybrid and suit the signal types and their implementations. Implementing ASK is easy, but it is limited to having low power and a low data transmission rate. And the PSK modulation has a normal transformation step from symbol to symbol, but non-continuously. However, the PSK modulation methods derive from DPSK, QPSK, and MSK. Hence, the type of M-array modulation scheme [71] is Binary Phase Shift Keying, Quadrature Phase Shift Keying, 8-PSK, and 16-PSK. The modulation methods BPSK, QPSK, 16-QAM have been tested for the accuracy of their BER values [72]. While the methods of power

efficiency, including Binary Phase Shift Keying or Quadrature Phase Shift Keying, are used to boost channel output requirements, 16-QAM or 64-QAM [73].

Furthermore, the efficiency assessment of a WiMax device under various configurations including Binary Phase Shift Keying, Quadrature Phase Shift Keying, 4-QAM, 16-QAM, and different communication channels and discolouration makes for additive White Gaussian noise results [74] [75]. Hence, the efficiency of several channel modulation schemes is studied in order to make White Gaussian Noise additive results [76].

The MSK, ASK, PSK, DPSK design, and demodulation devices were proved to be cost-effective, as the QAM methods are used exclusively in modems for digital television, digital microwave radios, set-top boxes, and broadband [76]. In mobile communication, QAM demonstrated its highest results with the maximum spectral output over QPSK and MSK [77-84].

The sensor nodes, tiny cubicles, usually have minimal battery power. In certain situations, the battery shift is just too expensive, causing energy consumption, if node concurrently transmits all unprocessed measurements obtained from sensors to a sinking node. At the same location, the sensor measurements of adjacent nodes are strongly correlated, so it is important to lower the data until communication systems are accurate. Network sensor nodes are responsible for four main tasks: data collection, data transmission & receiving, and data processing in-network. Hence, managing their resources to maximize their existence and efficiency, including memory manage, processor capacity, and most significant energy. In addition to the accumulation of energy, the enhancement of sensor life by energy usage in the network has been one of the key challenges of the use of underground contact in realistic applications. To mitigate energy usage as much as possible, the energy usage *Econ* of the underground communications physical layer can be indicated, as seen in equation 2.9[85].

$$E_{con} = ((E_{send}/\sigma) + E_{amp} + Q_{CS}+Q_{CA})*Q_{all} \quad (2.9)$$

Where, E_{send} refers to the transmitted capacity, which is described by the signal-to-noise ratio σ and by the particular error rate $(E_{accept}/2BN0)$, E_{accept} refers here to the received power, B to the signal bandwidth, N0 to the additive white Gaussian noise's spectral power density. There are various coding methods for the relationship between the signal-to-noise ratio and the bit error rate. The $E_{send} = Q((4\sigma) * 1/2)$ can be used to denote the relation between the frame error and the signal-to-noise limit, among which σ refers to the amplification efficiency of the receiving signal amplifier, Q specifies the phase number in QAM, E_{amp} refers to power amplifier usage, $E_{amp} = E_{send}$ refers to power consumption of the circuit at the transmission end, implies that the energy usage is equal to the data size, i.e., the bits. The problem of transmitting higher data (bits) to the digital communication system thus immediately improves the capacity of transmission. The present innovation focuses on rising transmission power when transferring the bits in a communication network. Q_{CS} refers to the circuit electricity used by the circuit on the transmitting end, Q_{CA} means the circuit power consumption at the receiving end while the data is being transmitted, and Q_{all} refers to the time necessary for data transmission to be completed, the completion of data transmission, which is the time needed for each date to be transmitted.

The M-FSK and M-PSK signals reliably envelope [86] (i.e., the carrier amplitude stays constant). The only distinction is that the symbol mapping for QAM produces non-constant amplitudes of the symbols. For both M-ary modulation schemes, bit error rate efficiency may be increased at the cost of improvement in signal-to-noise ratios per E_b / N_0 line. M-ary modulation schemes may, therefore, compared to the necessary E_b / N_0 to achieve a specific bit error rate. However, it is essential to include the evaluation of modulation schemes that affect the modulation efficiency (δ), transmission bit rate

(α_b), and bandwidth requirement (B). Hence, the bandwidth efficiency ω is specified, which is an essential measure to connect α_b in bits per second to the channel bandwidth requirement B in Hz. Further, the M-ASK signaling scheme may use one sideband communication, the channel bandwidth needed is therefore about one-half the symbolic limit given by equation 2.10 [86],

$$\omega = \frac{\alpha_b}{B} = \frac{\alpha_b}{\alpha_s/2} = \frac{\alpha_b}{(\alpha_b/log_2\,\delta)/2} = 2\,log_2\,\delta \qquad (2.10)$$

For the M-PSK signaling scheme, since double-sideband transmission is employed, the required channel bandwidth is equal to the symbol rate equated by equation 2.11.

$$\omega = \frac{\alpha_b}{B} = \frac{\alpha_b}{\alpha_s} = \frac{\alpha_b}{(\alpha_b/log_2\,\delta)} = log_2\,\delta \qquad (2.11)$$

QAM modulation employs double-sideband transmission, and the transmitted signal consists of two independent orthogonal quadrature carriers, its transmission rate is therefore twice that of ASK as given by equation 2.12

$$\omega = \frac{\alpha_b}{B} = \frac{\alpha_b}{\alpha_s/2} = \frac{\alpha_b}{(2\alpha_b/2log_2\,\delta)} = log_2\,\delta \qquad (2.12)$$

In the orthogonal MFSK signaling scheme, there are δ orthogonal carriers with a minimum frequency separation of 1/2Ts, where an orthogonal signal carries $\log_2 \delta$ bits of information given by equation 2.13.

$$\omega = \frac{\alpha_b}{B} = \frac{\alpha_b}{\delta/2T_s} = \frac{\alpha_b}{\delta\alpha_s/2} = \frac{\alpha_b}{\delta(\alpha_b/log_2\,\delta)/2} = \frac{2\,log_2\,\delta}{\delta} \qquad (2.13)$$

Bandwidth efficiency ω is an increasing function of δ for M-ASK, M-PSK, and QAM modulation schemes (i.e., for a fixed bit rate α_b, when δ increases, channel bandwidth B decreases logarithmically). Also, QAM and MPSK are commonly used in wireless telecommunications networks that use high-speed voice band modems and for

the delivery of mobile multimedia communication on digital cellular wireless channels [86].

In terms of the signal space, it is often helpful to evaluate δ–ary (δ=M) modulation schemes. The dimensionality of the signal space stays unchanged as further signals for M-ASK, M-PSK, and QAM. In other words, as the number of signal points is increased, the signal points are packed closer together, resulting in a reduction of the symbol error rate, while the necessary bandwidth remains relatively constant, irrespective of the value of δ. With the number of signals increased in M-ASK, M-PSK, and QAM, bandwidth efficiency, and the bit error rate performance can be increased. Although M-PSK and QAM have similar bandwidth efficiencies, the average bit error rate for QAM is better than for M-PSK for a given peak transmitted capacity (i.e., QAM has a better power output than M-PSK). However, increasing the efficiency of the channel, which is linear, and the transmitting amplifier can be worked in a linear region. The M-FSK is an improvement in δ, resulting in a significant rise in dimensionality and, in turn, in modem complexity. If more signals are applied to the dimensionality, the signal points do not have to be crowded closely together [86].

Modulation is a silent requirement for all pass band channels, including all wireless communication systems. The primary motivation for the widely-used M-ary modulation schemes, such as MPSK, is bandwidth conservation, but at the expense of an increase in transmits power and/or increases in bit error rate. On the other hand, M-QAM, with its significant expansion of bandwidth, allows a reduction in transmit power and/or a decrease in bit error rate.

2.5. Multiple Input and Multiple Output (MIMO)

MIMO solutions have been one of the most popular in recent years to ensure a high data rate approach and more reliable wireless communication networks. While,

several antennas are mounted on both the source (transmitter) and the target (receiver), by increasing the number of transmissions and receiver antennas, the efficiency of the MIMO network increased linearly. Further, the multiplexing gain (data throughput), the increase in complexity, as well as the coding gain (link reliability) of MIMO systems are generally higher than that of traditional single input single output systems (SISO) [87].

MIMO wireless technology can increase the efficiency of a given channel considerably while keeping with the principle of Shannon's. However, by changing the number of antennas received and broadcast, the channel efficiency can be linearly increased for each pair of antennas connected to the network. In recent years, MIMO wireless technology has become one of the most popular wireless techniques. As spectral bandwidth becomes an asset for radio communication systems, strategies are needed for more efficient use of the usable bandwidth. One such technique is wireless MIMO technology [87-89].

Multi-user MIMO (MU-MIMO) systems have received impressive popularity in recent times, primarily due to their potential to improve the efficiency and reliability of wireless data transfer. For MU-MIMO, the Base Station (BS) is wired to various antennas and also wired to multiple Mobile Stations (MS). In turn, each of these MS is fixed for multiple antennas. In a one-time resource initiative, multi-user diversity in the field of space is distributed to multiple users, which leads to substantial increases compared to Single-user MIMO (SU-MIMO), especially in space-related channels [90].

Transmitter and receiver antennas have been designed to increase the potential for signal transmission and reception. The use of multiple elements in antennas to form beams and direct nulls at a given point has a significant positive impact over decades in the received power and signal to noise ratio (SNR) for maximum signal transmission and reception. However, since modulation of the antenna components can be achieved

at both base station and unit terminal, this gives rise to an important method of using multiple antennas for sending and receiving signals with an increased ability to channel capacity from space. It is called the multiple-input and multiple-output (MIMO) principle [91].

The increasing demand for data transmission through wireless channelled through multiple input multiple output (MIMO) technology being developed. The further use of several antennas on each end of a wireless network makes it possible to access several spatial data pipes within the frequency operating range between the transmitter and the receiver with no extra power expenditures, leading to a spectral efficiency significantly improved, known as space multiplexing gain. The pioneering work of Foschini and Gans [92] and Telatar [93] has well illustrated the potential of MIMO Technology of delivering high information rates without additional spectral requirements. Further, Rayleigh fading has a substantially large amount of literature regarding only non-line-of-sight (NLOS) elements. However, Line-Of-Sight (LOS) components are best represented by the Rician fading distribution, between the transmitter and the receiver. In [94], after the dissemination of Rician, the author explores the limitations on the efficiency of the MIMO communication system. In [95], when the fading coefficients are equal, but not the same, as the approach implemented an exact expression for the MIMO Rician fading channels average rate of Mutual Information (MI).

The research [96] indicates that the presence of large LOS components contributes to channel sparsity, thereby reducing the number of Degrees of Freedom (DoF). The presence of NLOS components decreases the correlation between the signals, thus increasing the channel matrix range [97] [98]. The author analyzes the ergodic capabilities of MIMO channels with mean grade 1 matrix. The upper and lower limits were set on ergodic capacity. However, the upper limits on the ergodic capacity of the

Rician fading system are obtained from the arbitrary Signal to Noise ratio (SNR) and matrix ranking [99].

M. Surendar, P. Muthuchidambaranathan [100], demonstrated that conventional detection of MIMOs such as zero force (ZF) and minimum mean square error (MMSE) is seen to have a substantial output loss, thereby exponentially raising the difficulty of maximum likelihood (ML) detection. The decoding algorithm for the system using STBC is also proposed, and the complexity is further minimized. It is also seen that the diversity order ranges according to the number of antennas broadcast are received. However, simulation tests indicate that NCP-OFDM systems with multiple antennas do better than standard SISO systems. This also highlighted the benefit of adding STBC to the MIMO system and its contribution to the rise in diversity.

Yang Zhang et al. [101], proposed multi-user MIMO Systems interference alignment. Point-to-point MIMO systems the imperfect channel state information (CSI) only causes the SNR capacity versus SNR curve. However, the precision of the CSI decreases the cure slope, that is, Degree of freedom (DoF), in the MIMO systems with multiple users. Hence, the efficiency of feedback can be enhanced by means of channel properties such as temporal correlation, spatial correlation, and heterogeneous loss of path. Feedback systems must be specifically configured to match the system performance with overhead feedback. Hence, Interference synchronization (IA) has been widely studied as an emerging technique to reach the optimum degree of Freedom (DoF) in wireless multi-user communication systems.

Frederick W Vook, Amitava Ghosh, Timothy A. Thomas [102], suggests multi-antenna innovations such as beam shaping, MIMO, which are expected to play a crucial role in '5G' networks in 2020 and beyond. Further, a class of 5G systems, which is expected to be installed in both cm-waves (3-30 GHz) and waves (30-300 GHz), has prompted a rethink of the design and the performance of the existing multi-antenna

techniques to determine which framework is preferred for MIMO technology in 5G systems. The main design problems surrounding the introduction of MIMO communication capacity for 5G networks are addressed.

Kassim Khalil et al. [103] suggested the Indoor Power-Line Communication MIMO Random Channel Generator. MIMO technology has already been used in power line communication (PLC) used to build and test digital connectivity algorithms. The detailed definition of MIMO channels for PLC network topologies is of great significance and importance. However, the enormous complexity of PLC networks makes the task very difficult. Hence, a random channel generator has been presented in the literature for a single-input single-output (SISO) system in this respect.

L. Yang et al. [104], suggested a spectrum sharing capability study of a spatial multiplexant MIMO network. It is a MIMO spectrum sharing (SS) system running in a Rayleigh fading environment. Initially, the capacity of a single-user SS spatial multiplexing system is investigated in two scenarios, assuming separate receivers. Some estimated expressions of capacity for two scenarios that obtained to display the capacity scaling law of SS MIMO systems specifically. Further, they extend the study under spatially independent scheduling to multiple user programs of zero-forcing receivers (ZF) and evaluate the sum-rate. It also provides an asymptotic sum-rate analysis to analyse the impact of various parameters on the gain in diversity from multi-users.

Lu Theyi, Zhong Zheng, Jukka Corander, and Giorgio Taricco, [105] have suggested a useful model for pico-cellular MIMO networks for the exit capacity of orthogonal Space-time Codes over multi-cluster scattering MIMO channels. The space-time coded communication is a channel where the effective channel corresponds to a set of complex Gaussian matrices. An exact, closed-form approximation to the channel outage capability is obtained using orthogonal space-time block codes.

X Gu, X-H Peng & G C Zhang [106], presented MIMO systems for wireless broadband communications. The advantage of broadband wireless access (BWA) and connectivity raises critical obstacles for the future environment of telecommunications. Whereas, MIMO is a potential technology that can be used to improve system performance for broadband wireless network coverage, capacity and data rate. However, certain findings of studies interest to both current and future systems, yet such benefits can potentially be accomplished with a MIMO system is unclear. The basic issues related to capacity building and efficiency limitations need to be dealt with before operators are able to deploy MIMO systems on their networks.

Mohammad Rakibul Islam et al. [107], proposed the energy-efficient cluster-to-cluster transmission over the wireless sensor network on the MIMO connectivity cooperative. However, the reasons for energy-efficient WSN are the energy-efficient data transfer. The cooperative MIMO explores wireless communication schemes between multiple sensors highlighting the structure of MIMO. In WSN, an energy-efficient cooperative technique is proposed in which selected numbers of sensors are used at the transmitting end to form a wirelessly connected MIMO structure with a selected number of sensors at the receiving end. The selection of nodes in the transmitting end is based on a selection method, which is a mixture of channel state, residual capacity, inter-sensor distance in a cluster, and geographic position. Hence the selection on the receiving side is performed on the basis of channel state. And the data is sent to a data-gathering node (DGN) utilizing a multi-hop transfer by the sensors in a cluster.

Fengbiao Zan [108] introduced a wireless sensor network in the energy-saving virtual MIMO transmission scheme. Vijaykumar. V. R. et al. [109] used compressive sensing to present an effective image delivery through MIMO channels. An image of both source coding, and channel coding techniques are employed. While the image is

source coded using the Compressive Sensing (CS) algorithm that transforms the input image to numerical results, such numerical data are then coded as space-time block and can be concurrently distributed over the four antennas. Then the unequal power allocation scheme is implemented over such data streams where each transmitting antenna gets a different amount of power based on the data that it transmits.

Furthermore, Yi Gai et al. [110] suggested MIMO co-operative energy conservation with data aggregation in wireless sensor networks. An energy model is developed for wireless sensor networks based on the mutual MIMO method, considering both the energy transfer and data aggregation. As, the model contrasts contain two mutual WSN systems, namely the MIMO approach and the SISO solution. Further, the total energy consumption in the systems is seen to be related not only to the spectrum of propagation but also to the connection between the raw sensor data, which can be solved as a nonlinear programming problem. Also, a critical value is evaluated over which MIMO solution outperforms the SISO solution.

Songfu Cai et al. [111] proposed MIMO pre-coding with energy harvest sensors for networked control systems. At the sensor, they concentrate on energy harvesting and MIMO pre-coding architecture to balance the fragile MIMO complex plant subject to the sensor's energy availability restriction. However, the proposed closed-form dynamic energy harvesting and dynamic MIMO pre-coding solution has an event-driven control system, using the Lyapunov optimization method. The MIMO pre-coding solution has its own value water-filling structure, where the water level depends on the covariance of the state estimate, the energy queue, the state of the channel, and the depth of the sea bed depends on the covariance of the state estimation.

L. Gortschacher et al. [112] implemented SIMO UHF RFID reader for tag localization in a selected area by means of sensor fusion. Here the channel-based approach of deploying RFID readers is used to achieve efficient, customizable UHF

RFID systems in any chosen setting in a favourable manner. Hongli Xu et al. [113] proposed a virtual MIMO joint and wireless sensor network data set, reveal Virtual multi-input multi-output vMIMO technology which allows for spatial flexibility by requiring external antennas in wireless networks and decreases power consumption by operating with multiple nodes. As data collection is one of the most critical activities in many sensor network systems, energy-efficient data collection is provided by vMIMO in wireless sensor networks.

Mojtaba Radmard et al. [114] proposed the fusion of data in a coherent passive position based on the MIMO DVB-T. It is expected that a successful combination of MIMO and PCL (passive coherent location) ideas would boost the efficiency of localization schemes. The authors further suggest a scheme that efficiently tackles the association issue. Here an interaction methodology has been suggested based on concepts inspired by multi-sensor-multi-object monitoring to solve this problem in an effective way.

Further, for the fifth-generation technology (5G), there is a need for evolution in MIMO. The new field of work currently underway for this 5G system is named massive MIMO. Simply put as the technology which will supply base stations with a vast number of antennas (hundreds or even thousands) compared to traditional MIMO [115, 116, 117]. The MIMO may also be referred to as very large MIMO, ARGOS, full MIMO or hyper MIMO [116, 118] so as to exploit the spatial diversity.

In addition, 5G networks are intended to support a very complex system of large numbers of machine-to-machine and machine-to-person contact devices [119]. Therefore, the spectrum of frequencies for large MIMO is required to vary from centimeter/millimeter, from 6 GHz to around 100 GHz, to help these ever searching consumers of data. The traffic is likely to go up tremendously. However, one billion new mobile subscribers planned to reach the current 3.6 billion and make it 4.6 billion

by 2020, and to cope with this growth, and massive MIMO may be a needed to offer data rates of up to 100 Mbit / s and data peaks of about 10 Gbit / s [120].

2.6. Objectives

The literature review and objectives of the research work. A literature review on the study of source coding technique, LDPC channel coding technique for robust forward error corrections, State of the Art modulation schemes i.e., PSK/M-QAM, with MIMO channel is extensively performed for the data communication system. Thereafter in view of the above, the significance of the present study to develop a PCA base source coding for data reduction using a MIMO system having application in a modern communication system has been presented. The objectives of this thesis are as follows:

1. To model the source coding technique for the sensor level signals to fused data and transmit over the data communication system.
2. To study the performance of the source coding technique using the MIMO channels.
3. To test the channel coding LDPC forward error corrections algorithm for the above data communication system over AWGN channels.
4. To explore the above data communication system for Network on Chip (NoC) application.

2.7. References

[1] Niklas Wernersson. *Source-channel coding in network*. Doctoral thesis, KTH - Royal Institute of Technology, 2008.

[2] J. Uthayakumar, T. Vengattaraman, P. Dhavachelvan. A survey on data compression techniques: From the perspective of data quality, coding schemes, data type and applications. *Journal of King Saud University - Computer and Information Sciences*, 2018.

[3] Drost, S.W., Bourbakis, N. A hybrid system for real-time lossless image compression. *Microprocess. Microsystem*, 25:19–31, 2001.

[4] Ansari, R., Memon, N., Ceran, E. Near-lossless image compression techniques. J. Electron. *Imaging System*. 7, 486-494, 1998.

[5] S. Boubiche, D. E. Boubiche, A. Bilami and H. Toral-Cruz. Big Data Challenges and Data Aggregation Strategies in Wireless Sensor Networks. *IEEE Access*, 6:20558-20571, 2018.

[6] Anastasi, G., Conti, M., Di Francesco, M. and Passarella, A. Energy conservation in wireless sensor networks: A survey. *Ad hoc networks*, *7*(3):537-568, 2009

[7] Goel, S. and Imielinski, T. Prediction-based monitoring in sensor networks: taking lessons from MPEG. *ACM SIGCOMM Computer Communication Review*, *31*(5):82-98, 2001.

[8] Marcelloni, F. and Vecchio, M. Enabling energy-efficient and lossy-aware data compression in wireless sensor networks by multi-objective evolutionary optimization. *Information Sciences*, *180*(10):1924-1941, 2010.

[9] Pattem, S., Krishnamachari, B. and Govindan, R. The impact of spatial correlation on routing with compression in wireless sensor networks. *ACM Transactions on Sensor Networks (TOSN)*, *4*(4):1-33, 2008.

[10] A. Silberstein, G. Filpus, K. Munagala, J. Yang. Data-driven processing in sensor networks. In *Proceedings of the Biennial Conference on Innovative Data Systems Research (CIDR)*, Pages 10–21, 2007.

[11] S. Li, L. Da Xu, X. Wang. Compressed sensing signal and data acquisition in wireless sensor networks and internet of things. *IEEE Transactions on Industrial Informatics*, 9(4): 2177–2186, 2013.

[12] Y. Zhang, N. Meratnia, P. Havinga. Outlier detection techniques for wireless sensor networks: survey. *IEEE Communications Surveys & Tutorials*, 12(2):159-170, 2010.

[13] Bishop, C. M. *Pattern recognition and machine learning*. Springer, 2006.

[14] Alsheikh, M. A., Lin, S., Niyato, D. and Tan, H. P. Rate-distortion balanced data compression for wireless sensor networks. *IEEE Sensors Journal*, *16*(12):5072-5083, 2016.

[15] Wu, M., Tan, L. and Xiong, N. Data prediction, compression, and recovery in clustered wireless sensor networks for environmental monitoring applications. *Information Sciences*, *329*:800-818, 2016.

[16] K. D. Rao. *Channel Coding Techniques for Wireless Communications*. Springer, 2015.

[17] Hamming, Richard W. Error detecting and error-correcting codes. *Bell System Technical Journal*, 29 (2): 147–160, 1950.

[18] Shannon, C.E., 1948. A mathematical theory of communication. *The Bell system technical journal*, *27*(3):379-423, July 1948.

[19] M. Golay. Complementary Series. *IRE transactions on information theory*, 7(2): 82-87, 1961.

[20] Reed, I. A class of multiple-error-correcting codes and the decoding scheme. *Transactions of the IRE Professional Group on Information Theory*, *4*(4):38-49, 1954.

[21] Muller, D. E. Metric properties of Boolean algebra and their application to switching circuits report no. 46. *Digital Computer Laboratory, Univ. of Illinois, Tech. Rep*, 1953.

[22] E. R. Berlekamp. *Algebraic coding theory*. World Scientific, 2015.

[23] Reed, I.S. and Solomon, G. Polynomial codes over certain finite fields. *Journal of the society for industrial and applied mathematics*, *8*(2):300-304, 1960.

[24] Bose, R.C. and Ray-Chaudhuri, D.K. On a class of error correcting binary group codes. *Information and control*, *3*(1):68-79, 1960.

[25] Massey, J. Shift-register synthesis and BCH decoding. *IEEE transactions on Information Theory*, *15*(1):122-127, 1969.

[26] T.C. Lin, T.K. Truong, and P.D. Chen. A Fast Algorithm for the Syndrome Calculation in Algebraic Decoding of Reed-Solomon Codes. *IEEE Transactions on Communications*, 55(12):2240- 2244, 2007.

[27] Chang, Y.W., Truong, T.K. and Jeng, J.H. VLSI architecture of modified Euclidean algorithm for Reed–Solomon code. *Information Sciences*, *155*(1-2):139-150, 2003.

[28] Truong, T. K., Jeng, J. H. and Cheng, T. C. A new decoding algorithm for correcting both erasures and errors of Reed-Solomon codes. *IEEE Transactions on communications*, *51*(3):381-388, 2003.

[29] Truong, T. K., Jeng, J. H. and Reed, I. S. Fast algorithm for computing the roots of error locator polynomials up to degree 11 in Reed-Solomon decoders. *IEEE Transactions on Communications*, *49*(5):779-783, 2001.

[30] Truong, T.K., Lee, C.D., Chang, Y. and Su, W.K. A new scheme to determine the weight distributions of binary extended quadratic residue codes. *IEEE transactions on communications*, *57*(5):1221-1224, 2009.

[31] Truong, T.K., Shih, P.Y., Su, W.K., Lee, C.D. and Chang, Y. Algebraic Decoding of the $(89, 45, 17) $ Quadratic Residue Code. *IEEE transactions on information theory*, *54*(11):5005-5011, 2008

[32] Lee, C.D., Chang, Y. and Truong, T.K. A result on the weight distributions of binary quadratic residue codes. *Designs, Codes and Cryptography*, *42*(1):15-20, 2007.

[33] Chen, Y.H., Truong, T.K., Chang, Y., Lee, C.D. and Chen, S.H. Algebraic decoding of quadratic residue codes using Berlekamp-Massey algorithm. *Journal of information science and engineering*, *23*(1):127-145, 2007.

[34] Truong, T.K., Chang, Y., Chen, Y.H. and Lee, C.D. Algebraic decoding of (103, 52, 19) and (113, 57, 15) quadratic residue codes. *IEEE transactions on communications*, *53*(5):749-754, 2005.

[35] Gallager, R. Low-density parity-check codes. *IRE Transactions on information theory*, *8*(1):21-28, 1962.

[36] Goppa, V.D. A new class of linear correcting codes. *Problemy Peredachi Informatsii*, *6*(3):24-30, 1970

[37] Berrou, C., Glavieux, A. and Thitimajshima, P. May. Near Shannon limit error-correcting coding and decoding: Turbo-codes. In *Proceedings of ICC'93-IEEE International Conference on Communications*, 2:1064-1070, 1993.

[38] Sudan, M. List decoding: Algorithms and applications. In *IFIP International Conference on Theoretical Computer Science*, Pages 25-41. Springer, Berlin, Heidelberg, 2000.

[39] Mahmood, R., Badr, A. and Khisti, A. Convolutional codes with maximum column sum rank for network streaming. *IEEE Transactions on Information Theory*, *62*(6):3039-3052, 2016.

[40] Bocharova, I.E., Kudryashov, B.D. and Johannesson, R.LDPC convolutional codes versus QC LDPC block codes in communication standard scenarios. In *2014 IEEE International Symposium on Information Theory*, Pages 2774-2778, 2014.

[41] Hong, W. Development of microwave antennas, components and subsystems based on SIW technology. In *2005 IEEE International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications*, 1:1-14, 2005.

[42] Omar, M.H., El-Mahmoudy, A., Seddik, K.G. and Elezabi, A. November. On the Tail-Biting Convolutional Code Decoder for the LTE and LTE-A Standards'. In *2013 Asilomar Conference on Signals, Systems and Computers*, pages510-514, 2013

[43] Kashyap, M. and Winstead, C. Decoding LDPC convolutional codes on Markov channels. *EURASIP Journal on wireless communications and networking*, *2008*(1):729180, 2008.

[44] Hosseinzadeh, D., Khademi, A. and Krishnan, S. *Source and Channel Coding Techniques for faithful transmission of Digital Mammograms.* IEEE Canadian review, 2004.

[45] R. G. Gallager. *Low Density Parity Check Codes*. PhD thesis, Massachussetts Institute of Technology (MIT), Cambridge, Mass., 1963.

[46] D. J. C. MacKay. *Good error-correcting codes based on very sparse matrices*. IEEE Trans. Inf. Theory, 45:399-431, 1999.

[47] Cataldi, P., Gerla, M. and Zampognaro, F. Rateless codes for file transfer over DVB-S. In *2009 First International Conference on Advances in Satellite and Space Communications*, Pages 7-12, 2009

[48] Azarbad, B. and Sali, A.B. DVB-S2 Model in Matlab: Issues and Impairments. In *MATLAB-A Fundamental Tool for Scientific Computing and Engineering Applications*, 2:217-234, 2012.

[49] Morello, A. and Reimers, U. The DVB-S2 standard for broadband satellite systems. *International journal of satellite communications and networking (Print)*, *22*(3), 2004.

[50] Méric, H. and Piquer, J.M.DVB-S2 spectrum efficiency improvement with hierarchical modulation. In *2014 IEEE International Conference on Communications (ICC)*, Pages 4331-4336, 2014.

[51] Hussien, H., Shehata, K.A., Khedr, M. and Hareth, S. Performance study on implementation of DVB-S2 low density parity check codes on additive white Gaussian noise channel and Rayleigh fading channel. In *2012 IEEE International Conference on Electronics Design, Systems and Applications (ICEDSA)*, Pages 179-182, 2012.

[52] Kolawole, M.O. *Satellite communication engineering*. CRC Press, 2017.

[53] Liveris, A.D., Xiong, Z. and Georghiades, C.N. Compression of binary sources with side information at the decoder using LDPC codes. *IEEE communications letters*, *6*(10):440-442, 2002.

[54] Gallager, R.Low-density parity-check codes. *IRE Transactions on information theory*, *8*(1):21-28, 1962.

[55] Ali Grami. *Introduction to Digital Communications*, Academic Press, Pages 1-10, 2016.

[56] Lou Frenzel. Understanding Modern Digital Modulation Techniques. *Electronicdesign*, at www.electronicdesign.com/technologies/communications article/21798737/understanding-modern-digital-modulation-techniques.

[57] Taha, H.J. and Salleh, M.F.M. Multi-carrier transmission techniques for wireless communicationsystems:a survey. *Wseas transactions on communications*, *8*(5):457-472, 2009.

[58] B.W. Hiroshi, W.M. Ramjee. *Digital Communication*. Artech House, London, Pages 5-8, 2005.

[59] Tharakanatha, G., Mahaboob, S.K., Chanda, V.B. and Hemalatha, I. Implementation and Bit Error Rate analysis of BPSK Modulation and Demodulation Technique using MATLAB. *International Journal of Engineering Trends and Technology (IJETT), 4*(9):4010-4014, 2013.

[60] Chauhan, V., Arora, M. and Chauhan, R.S. Comparative BER performance of PSK based modulation techniques under multipath fading. *Advances in Applied Science Research*, *2*(4):521-524, 2011.

[61] Bagga, J. and Tripathi, N. Study and comparison of various modulation classification techniques under noisy channel conditions. *International Journal of Emerging Technology and Advanced Engineering*, 2(4):216-221, 2012.

[62] Pandey, R. and Pandey, K., 2014. An Introduction of Analog and Digital Modulation Techniques in Communication System. *Journal of Innovative Trends in Science Pharmacy & Technology*, *1*(1):80-85, 2014.

[63] Bhambare, R.R. and Raut, R.D. A survey on digital modulation techniques for software defined radio applications. *International Journal of Computer Networks and Wireless Communications*, *3*(3):292-299, 2013.

[64] Mongre, R. and Kapoor, M. Compative Analysis of Digital modulation Techniques on the Basis of Their Bit Error Rate in VHDL, *International Journal of Engineering and Technical Research*, 5(5):1205-1215, 2013.

[65] K. Hamid, I. Khider, A. Babiker. Performance Evaluation of Digital Modulation Techniques on DS-WCDMA, *International Journal of Computer Applications*, 74(8):1-4, 2013.

[66] J.D. Oetting.A Comparison of Modulation Techniques for Digital Radio. *IEEE Transactions on Communications*, 27(12):1752-1762, 1979.

[67] B.Sklar, P.K.Ray. *Digital Communications, Fundamentals and Applications*. Second Edition, Pearson Education, Inc, 2001.

[68] Ziemer, R.E. and Tranter, W.H. *Principles of communications*. John Wiley & Sons, 2014.

[69] Tao, M. Principles of communication: Chapter 8-Digital modulation techniques. *Shanghai Jiao Tong University*, Pages 1-79, 2012.

[70] Adeleke, O.A. and Abolade, R.O. Modulation methods employed in digital communication: an analysis. *International Journal of Engineering and Computer Science, 12*(03):85-93, 2012.

[71] R. Singh. MIMO System using Space-Time Block Code with Digital Modulation Techniques, *International Journal of Electronics Engineering Research*, 1(1): 26-31, 2013.

[72] Rehman, A., Khan, T. and Chaudhry, S. Study of WiMAX Physical Layer under Adaptive Modulation Technique using Simulink. *International Journal of Scientific Research Engineering & Technology*, *1*(5):05-11, 2012.

[73] Gupta, S. and Sharma, H. Performance Investigation for different modulation techniques in WCDMA with multipath fading channels. *International Journal of*

Advanced Research in Computer Science and Software Engineering, *2*(7):20-23, 2012.

[74] A. Islam, Z. Hasan. Performance Evaluation of Wi-Max Physical Layer under Adaptive Modulation Techniques and Communication Channels. *International Journal of Computer Science and Information Security*, 5(1):1-5, 2009.

[75] S.A. Oyetunji, A.A. Akinninranye. Performance Evaluation of Digital Modulation Techniques in AWGN Communication Channel. *International Journal of Engineering Research & Technology*, 2(5):34-39, 2013.

[76] V.Tilwari, A.S. Kushwah. Performance Analysis of Wi- Max 802.16e Physical Layer using Digital Modulation Techniques and Code Rates. *International Journal of Engineering Research and Applications*, 3(4):1449-1454, 2013.

[77] Q. Guo, P. Nan, X. Zhang, Y. Zhao, and J. Wan. Recognition of radar emitter signals based on SVD and AF main ridge slice. *Journal of Communications and Networks*, 17(5):491–498, 2015.

[78] D. Zeng, X. Zeng, H. Cheng, and B. Tang. Automatic modulation classification of radar signals using the Rihaczek distribution and Hough transform. *IET Radar, Sonar & Navigation*, 6(5):322–331, 2012.

[79] B. Feng and Y. Lin. Radar signal recognition based on manifold learning method," *International Journal of Control and Automation*, 7(12):399–406, 2014.

[80] S. Huang, Y. Yao, Z. Wei, Z. Feng, and P. Zhang. Automatic Modulation Classification of Overlapped Sources Using Multiple Cumulants. *IEEE Transactions on Vehicular Technology*, 66(7):6089–6101, 2017.

[81] L. Wang and Y. Ren. Recognition of digital modulation signals based on high order cumulants and support vector machines. In *Proceedings of the 2009 ISECS International Colloquium on Computing, Communication, Control, and Management (CCCM)*, Pages 271–274, 2009.

[82] H. Bai, Y.-J. Zhao, and D.X. Hu. Radar signal recognition based on the local binary pattern feature of time-frequency image. *Yuhang Xuebao/Journal of Astronautics*, 34(1):139–146, 2013.

[83] M. W. Aslam, Z. Zhu, and A. K. Nandi. Automatic modulation classification using combination of genetic programming and KNN. *IEEE Transactions on Wireless Communications*, 11(8):2742–2750, 2012.

[84] M. Chen and M. L. Fowler. Data compression trade-offs in sensor networks. In *Proceedings of SPIE*, SPIE-5561:96-107, 2004.

[85] Chunliang Zhou, Ming Wang, Weiqing Qu, and Zhengqiu Lu. A Wireless Sensor Network Model considering Energy Consumption Balance. *Mathematical Problems in Engineering*, 2018:1-8, 2018.

[86] Guimarães D.A. *Passband Digital Transmission. In: Digital Transmission*. Signals and Communication Technology. Springer, Berlin, Heidelberg, 2010.

[87] Dr. Doron Ezri. A quick introduction to MIMO technology. *2009 IEEE International Conference on Microwaves, Communications, Antennas and Electronics Systems*, Tel Aviv, Pages 1-23, 2009.

[88] S Alamouti, S.M. A simple transmit diversity technique for wireless communications. *IEEE Journal on selected areas in communications*, *16*(8):1451-1458, 1998.

[89] Ismail, A., Sezginer, S., Fiorina, J. and Sari, H. A simple and robust equal-power transmit diversity scheme. *IEEE communications letters*, *15*(1):37-39, 2010.

[90] Narasimhan, T.L., Raviteja, P. and Chockalingam, A. Large-scale multiuser SM-MIMO versus massive MIMO. In *2014 Information Theory and Applications Workshop (ITA)*, Pages 1-9, 2014.

[91] Paulraj, A.J., Gore, D.A., Nabar, R.U. and Bolcskei, H. An overview of MIMO communications-a key to gigabit wireless. *Proceedings of the IEEE*, *92*(2):198-218, 2004.

[92] Foschini, G.J. and Gans, M.J. On limits of wireless communications in a fading environment when using multiple antennas. *Wireless personal communications*, *6*(3):311-335, 1998.

[93] Telatar, E. Capacity of multi-antenna Gaussian channels. *European transactions on telecommunications*, *10*(6):585-595, 1999.

[94] Yan, S. and Yerong, Z. Maximum MIMO capacity of Rice channel. In *2009 International Conference on Wireless Communications & Signal Processing*, Pages 1-5, 2009

[95] Kang, M. and Alouini, M.S. Capacity of MIMO Rician channels. *IEEE Transactions on Wireless Communications*, *5*(1):112-122, 2006.

[96] Matthaiou, M., Sayeed, A. M., and Nossek, J. A. Sparse Multipath MIMO Channels: Performance Implications based on Measurement Data. In *Proceedings of IEEE Workshop on Signal Processing Advances in Wireless Communications, Perugia, Italy*, Pages 364-368, 2009.

[97] McKay, M. R., and Collings, I. B. General capacity bounds for spatially correlated Rician MIMO channels. *IEEE Transactions on Information Theory*, 51(9):3121-3145, 2005.

[98] Jin, S., Gao, X., and You, X. On the Ergodic capacity of rank-1 Rician Fading MIMO channels. *IEEE Transactions on Information Theory*, 53(2):502-517, 2007.

[99] Matthaiou, M., Kopsinis, Y., Laurenson, D. I. and Sayeed, A.M. Ergodic capacity Upper Bound for Dual MIMO Rician Systems: Simplified Derivation and Asymptotic Tightness. *IEEE Transactions on Communications*, 57(12):3589-3596, 2007.

[100] Surendar, M. and Muthuchidambaranathan, P. Low complexity and high diversity gain non-linear constellation pre-coded MIMO-OFDM system with subcarrier grouping. *AEU-International Journal of Electronics and Communications*, *70*(3):265-271, 2016.

[101] Zhang, Y., Gu, C., Shu, R., Zhou, Z. and Zou, W. Interference alignment in multi-user MIMO systems. In *2014 14th International Symposium on Communications and Information Technologies (ISCIT)*, Pages 564-568, 2014.

[102] Vook, F.W., Ghosh, A. and Thomas, T.A. MIMO and beam forming solutions for 5G technology. In *2014 IEEE MTT-S International Microwave Symposium (IMS2014)*, Pages 1-4, 2014.

[103] Khalil, K., Gazalet, M.G., Corlay, P., Coudoux, F.X. and Gharbi, M. An MIMO random channel generator for indoor power-line communication. *IEEE transactions on power delivery*, *29*(4):1561-1568, 2014.

[104] Yang, L., Qaraqe, K., Serpedin, E. and Alouini, M.S. Capacity analysis of spectrum sharing spatial multiplexing MIMO systems. *Physical Communication*, *13*:109-119, 2014.

[105] Wei, L., Zheng, Z., Corander, J. and Taricco, G. On the outage capacity of orthogonal space-time block codes over multi-cluster scattering MIMO channels. *IEEE Transactions on Communications*, *63*(5):1700-1711, 2015.

[106] Gu, X., Peng, X.H. and Zhang, G.C. MIMO systems for broadband wireless communications. *BT Technology Journal*, *24*(2):90-96, 2006.

[107] Islam, M.R. and Kim, J. On the cooperative MIMO communication for energy-efficient cluster-to-cluster transmission at wireless sensor network. *Annals of telecommunications-annales des télécommunications*, *65*(5-6):325-340, 2010.

[108] Zan, F. Wireless sensor network in the virtual MIMO energy-saving transmission scheme. *Int. J. Future Gener. Commun. Netw*, *7*(5):75-84, 2014.

[109] Vijaykumar, V.R., Ashok, K. and Dhivya, K. Efficient image transmission over MIMO channel using compressive sensing. In *2012 Third International Conference on Computing, Communication and Networking Technologies (ICCCNT'12)*, Pages 1-4, 2012.

[110] Gai, Y., Zhang, L. and Shan, X. Energy efficiency of cooperative MIMO with data aggregation in wireless sensor networks. In *2007 IEEE Wireless Communications and Networking Conference*, Pages 791-796, 2007.

[111] Cai, S. and Lau, V.K. MIMO pre-coding for networked control systems with energy harvesting sensors. *IEEE Transactions on Signal Processing*, *64*(17):4469-4478, 2016.

[112] Goertschacher, L., Grosinger, J., Khan, H.N., Auinger, B., Amschl, D., Priller, P., Muehlmann, U. and Bösch, W. SIMO UHF RFID reader using sensor fusion for tag localization in a selected environment. *e & i Elektrotechnik und Informations technik*, *133*(3):183-190, 2016.

[113] Xu, H., Huang, L., Qiao, C., Dai, W. and Sun, Y.E., 2014. Joint virtual MIMO and data gathering for wireless sensor networks. *IEEE Transactions on Parallel and Distributed Systems*, *26*(4):1034-1048, 2014.

[114] Radmard, M., Karbasi, S.M. and Nayebi, M.M. Data fusion in MIMO DVB-T-based passive coherent location. *IEEE Transactions on Aerospace and Electronic Systems*, *49*(3):1725-1737, 2013.

[115] J. Korhonen. *Introduction to 4G mobile communications*. Boston: Artech House, 2014.

[116] Larsson, E.G., Edfors, O., Tufvesson, F. and Marzetta, T.L. Massive MIMO for next generation wireless systems. *IEEE communications magazine*, *52*(2):186-195, 2014.

[117] Gao, X., Edfors, O., Rusek, F. and Tufvesson, F. Massive MIMO in real propagation environments. *IEEE Transaction Wireless Commun*, 2014.

[118] Anderson, A.L. and Jensen, M.A. A generalized sum-rate optimizer for cooperative multiuser massive MIMO link topologies. *IEEE Access, 2*:1040-1050, 2014.

[119] B. ALASTAIR. 5G mobile network features at www.unwiredinsight.com/5g-mobile-networkfeatures .

[120] GSMA. GSMA Press Release. At http://www.gsma.com/newsroom/press-release/one-billionnew-unique-mobile-subscribers-by-2020-gsma/.

Chapter 3

PCA base source coding

3.1. Introduction

Over the limited bandwidth in most of the cases, the data in real-time cannot be sent. However, the repeated transmitting of these raw data creates a major loss of data due to the contention and congestion of the channel. Therefore, the sensor signal data should be able to reduce or compress before transmission autonomously. Hence the present study model the source coding for data reduction over MIMO channel, as shown in Figure 3.1.

Also, to enhance the quality of the sensor node monitoring and performance, the quantity of data transmitted must be reduced by predicting future data behaviour, detecting and classifying important events locally within the sensor node, and transmitting only essential data. By way of Principle Component Analysis (PCA), the present study proposes to reduce the amount of data transmitted between sensor nodes and end-users.

In continuous monitoring of the most slow-speed data transition, the process results in a high amount of spatial or time redundancies and is consequently a waste of limited energy in communication between sensor nodes and end users. In theory, the increase in network life is proportionate to the decrease in the number of data packets sent. According to this theory, the reduction of data has become one of the most sophisticated methods of reducing data transmission [1–4]. An alternate solution to data reduction is the use of compressing techniques [5] [6] that minimize the amount of data transmitted as data size is limited. In general, data compression schemes can be divided into two categories, lossless and lossy compression. While in lossless compression of data involves the complete reconstruction of the original data from the compressed data.

Moreover, the data compression allows specific characteristics of the original data to be lost after decompression. The WSN with a high resource limitation, lossless algorithms are typically not required, due to the improved data recovery performance.

Further, the lossy compression reduces the amount of data to be transmitted through the WSN. Besides for loss compression, the compression amount and original reconstructed data are the essential parameters for determining the quality of compression algorithms. Compressing original data using the Principal Component Analysis (PCA) approach has proven to achieve satisfactory results.

Data Reduction (DR) is a method that can increase storage capacity and reduce cost through pre-processing of data mining. While DR aims at eliminating redundant data during transmission, throughout this effect, specific data management techniques are applied according to the WSN situation. However, there are three main groups of data reducing techniques that is, the network processing approaches, data compression approaches and data prediction approaches. Further, the data obtained by sensor nodes is first coded in data compression techniques [7] using other coding algorithms before being sent to the storage resource. While the cycle is reversed on the sink, the decoding is applied to retrieve the original sensor node data. However, the computational costs of coding algorithms are especially high in energy-restricted sensor nodes.

The use of PCA for WSN data processing is not recent and was adopted by Bontempi and Le [8] for mining sensor data. Seo et al. [9] performed a preliminary comparative analysis on multivariate stream data reduction for WSN, based on wavelets, filtering, hierarchical clustering, and Singular Value Decomposition (SVD) techniques. Results revealed that the strategies introduced differed in terms of efficiency and relative error ratio. Since the study was exploratory, no details were provided on the techniques.

The evolution of compression, measure in different parameter like Compression ratio (CR), peak signal to noise ratio (PSNR), and mean square error (MSE). CR determines the level of compression achieved to calculate the efficiency of the PCA technique.

Another important problem in data compression is to establish the original input data, recover or reconstruct. Therefore the noise or distortion in the compression process was introduced. The CR is increased, and then usually the quality of resulting data is decreased. So, a parameter noise or distortion degree, the measure is introduced in peak signal to noise ratio (PSNR). The PSNR and RMSE are used for measuring the signal data quality of reconstructed data after PCA compression. Furthermore, Le Borgne et al. [8], worked on distributed PCA model for data reduction, where the key components were in a distributed manner. The power iteration method by Z. Bai et al. [10] model has been used to quantify key components that have a high measurement rate iteratively.

There is a common restriction to current works such as [11], [12], [13] which are built to tackle one data stream variable, hence they are known as univariate data reduction models. While PCA is used in these models on the basis that different nodes obtain different readings for the same component at the same time. Many stream data variables (i.e., temperature, humidity, light, and voltage) can be used to customize the model, instead of one (i.e., temperature) variable.

Fenxiong et al. [13] suggested a multi-level PCA data reduction model in each layer of the clustered WSN system. Carvalho et al. [14] revealed the multivariate sensing results used to improve prediction precision for WSN data reduction. A multiple linear regression method, which produced high computational complexity and communication costs, has been adopted in this model. Furthermore, dependence on variable like time to forecast certain variables contributed to poor approximation accuracy. A multivariate approach for air quality control applications has been introduced by Silva et al. [15]. However, the detailed design of the network and reduction models is not elaborated.

The present work, obtain compressed signal data from input signal data using the PCA data reduction technique in a general data communication system. Also, the

present proposed system has been fixed for higher sensor data transfer and recovery of the original data on the receiver side. In this proposed model, mostly condensed input sensor data is reduced using PCA, and then transmitter and receiver using PSK, QAM modulation, the channel coding used here are MIMO with AWGN. Figure 3.1 shows PCA based data reduction communication system.

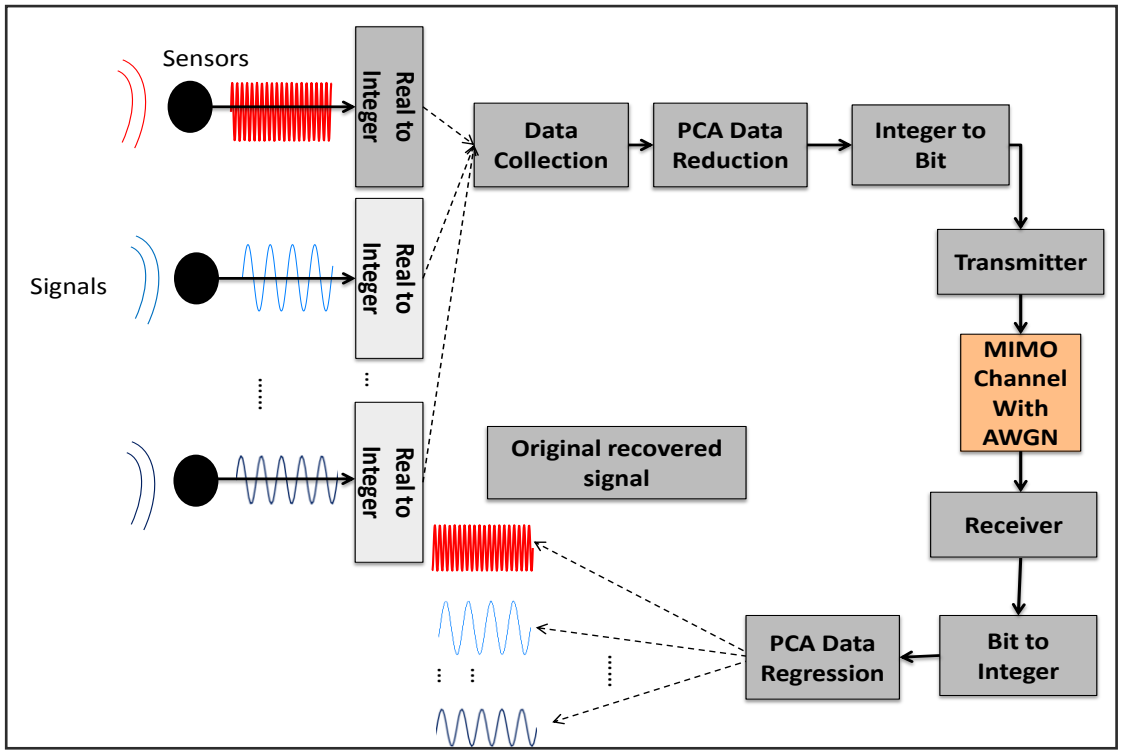


Figure. 3.1 Proposed PCA data reduction for digital data communication systems.

The principal component analysis (PCA) is a conventional data analysis source coding technique. It may also be used to reduce higher-dimensional data sets to lower-dimensional data for processing, simulation, pattern extraction, compression of data. The most important of these is to eliminate mean-square data, and then to locate mutually orthogonal directions in data that have large variances, and to compare the data with orthogonal transformations. The following steps are PCA data reduction or compression process.

Step 1: Get some data

To quantify the collection of data (H), calculate the mean of the data, for instance.

$$\bar{H} = \frac{\sum_{i=1}^{n} H}{n} \qquad (3.1)$$

$\bar{H}$- Indicates the mean of the set.

Step 2: Subtract the mean (Data adjust)

To ensure that PCA works properly, subtract the mean from each from the data dimensions. The mean subtracted is the average for each dimension. Therefore, all the averages (the sum of all the data points) have been subtracted. This produces a data set with a mean of zero. Here, the standard deviation (S.D.) or variance correction is a data set, which is a function of how the data is spread out. The way to do it is to measure the distance squares from each data point to the set mean, sum them all together, divide by, and take the square root positively. As shown in the equation3.2.

$$D = \sqrt{\frac{\sum_{i=1}^{n}(H_i - \bar{H})^2}{(n-1)}} \qquad (3.2)$$

Step 3: Calculate the covariance matrix

In this covariance matrix, the non-diagonal elements are positive, and it should be expected that both the variable will grow together. Variance is yet another indicator of data distribution in a data set. Also, the standard deviation is almost identical, hence displayed in equation 3.3.

$$D^2 = \frac{\sum_{i=1}^{n}(H_i - \bar{H})^2}{(n-1)} \qquad (3.3)$$

The formula for covariance is very similar to the formula for variance. The formula for variance could also be written, as shown in Equation 3.4.

$$var(H) = \frac{\sum_{i=1}^{n}(H_i - \bar{H})(H_i - \bar{H})}{(n-1)} \quad (3.4)$$

Where expanding the square term to show both parts as shown in Equation 3.5 for covariance:

$$cov(H, K) = \frac{\sum_{i=1}^{n}(H_i - \bar{H})(K_i - \bar{K})}{(n-1)} \quad (3.5)$$

Step 4: Calculate the eigenvectors and eigenvalues of the covariance matrix

The matrix of covariance is square, which determines that, the matrix its eigenvectors and eigenvalues. It is important to note that the individual vectors are unit eigenvectors, are of the same lengths. It is also significant for PCA which can also express whether the pattern of the data is very strong. However, the variables simply increase together as expected from the covariance matrix. Data is perpendicular to each other. The eigenvector tells how such sets of data are connected along the axis. And the second eigenvector offers the other, less significant, pattern in the results. So, by this process of taking the eigenvectors of the covariance matrix, it has been able to extract the lines that characterize the data. The rest of the steps involve transforming the data that it is expressed in terms of lines.

Step 5: Choosing components and forming a feature vector

This is where the concept of reduction of the data and reduced dimensionality comes in. From the previous part, it is found that the eigenvectors and eigenvalues are of different values. Also, it turns out that the data set equation 3.6 is the principle component of the eigenvector with the highest eigenvalue.

$$Feature\ vector\ (V) = (eig_{1,}\ eig_{2,}\ eig_{3\ldots.}eig_{n}) \quad (3.6)$$

The eigenvector with an eigenvalue is the one that points to the center of the data, and it is the most significant interaction between the dimensions of data which is usually 75% to 80 % of the total contribution. The said criteria are used to select the eigenvector to separate them from the list of eigenvector, and constructing a matrix in the columns of the eigenvector.

Step 6: Deriving the PCA data reduction

This is the final step in PCA data reduction and is, therefore, the easiest step. After choosing the components required to retain and to form a feature vector, it is necessary to simply transpose the vector and replicate it to the left of the original data set as given in equation 3.7.

$$S = V^T x D \qquad (3.7)$$

Final data is the S set, with data items in columns, and dimensions along rows. This is a format of reduced data or compression.

3.1.1. Regression of original Data

To obtaining the original data regression to retrieve the original data, it requires all the eigenvectors during our transformation i.e., the regression of the original data process shown in Figure 3.1. Note that the final transform is equation 3.7, i.e., $S = V^T x D$, which can be turned around to get the original data back as in equation 3.8.

$$D = V x S \qquad (3.8)$$

Where transpose of the feature vector is the inverse of feature vector. However, it turns out that the inverse of the feature vector is equal to the transpose of the feature vector. Hence it is true only because the matrix elements are all the unit vectors of the present data set. Therefore, it enables the return of our data as the equation is Equation 3.8.

$$D = V x S$$

But, to get the actual original data back, it is necessary to add on the mean of that original data (the data was subtracted right at the start). So, for completeness, equation 3.9 and 3.10 is given here.

$$\textit{Original recovered data } (R) = D + \bar{H} \qquad (3.9)$$

$$R = (V \times S) + \bar{H} \qquad (3.10)$$

This formula also applies, if all the vectors in the function variable are not in place. And when certain vectors are selected, the equation above always allows the correct transform. Further, the PCA data reduction or compression process is shown in Figure.3.2 in term of flowchart format.

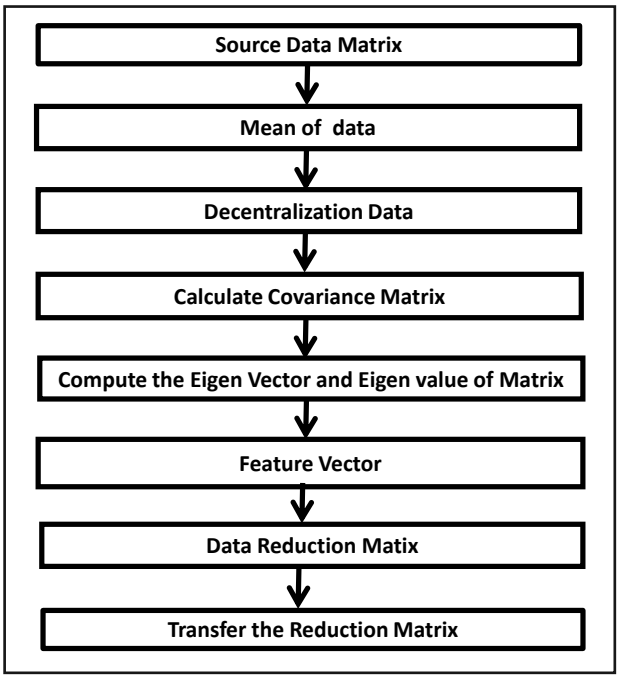


Figure.3.2. Flowchart for the PCA data reduction algorithm.

3.2. Experiments and Results

The present research evaluates system accuracy using the RMSE and BER varying SNR over the AWGN channel. The analysis is divided into different scenarios according to the number of signals transmitted to make an ensemble, different

modulation techniques, fading effects, and MIMO (1x1, 2x2, 3x3, and 4x4) channels. Table 3.1 is shown Different scenarios of PCA based communication system.

The root-mean-square error (RMSE) as expressed earlier, is used frequently to measure the differences between values predicted by a model or an estimator, and the values being observed. The RMSE serves to aggregate the magnitudes of the errors in reconstruction for various times into a single measure of predictive power, and it is a measure of the accuracy of reconstruction. While RMSE has been calculated by taking the square root of M.S.E. equation 3.11,

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} \left(\overline{y_i} - y_i\right)^2}{n}} \qquad (3.11)$$

Where, y_i is the observed value for the ith observation, and $\overline{y_i}$ is the corresponding predicted value. However, it can be positive or negative, as the predicted value is under or over estimates the actual value. Squaring the residuals, averaging the squares, and taking the square root gives us the RMSE. Further, RMS error is a measure of the spread of the y_i values about the predicted $\overline{y_i}$ value

Table 3.1 Different scenarios of PCA based communication system.

Scenario I	RMSE, BER v/s SNR for an ensemble of 10, 20, 30, 40, 50 sets of sensor signals over QPSK
Scenario II	RMSE, BER v/s SNR for an ensemble of 10, 20, 30, 40, 50 sets of sensor signals over QAM
Scenario III	RMSE, BER v/s SNR for an ensemble of 20 sets of sensor signals over 2, 4, 8, 16, 32 QPSK
Scenario IV	RMSE, BERv/s SNR for an ensemble of 20 sets of sensor signals over 2, 4, 8, 16, 32 QAM
Scenario V	RMSE, BER v/s SNR for an ensemble of 20 sets of sensor signals over QPSK, 4-QAM with Rayleigh fading at 0 to 5e-7 range.
Scenario VI	RMSE v/s SNR for an ensemble of 20 sets of sensor signals over QPSK, 4-QAM with Rayleigh fading at 0, 5e-7 range for 1x1, 2x2, 3x3, 4x4 MIMO channel.

Scenario I:

Here we consider, a wireless sensor network consisting of an ensemble of 10, 20, 30, 40, and 50 sets of sensor signals for 100 data points to be transmitted over QPSK modulation.

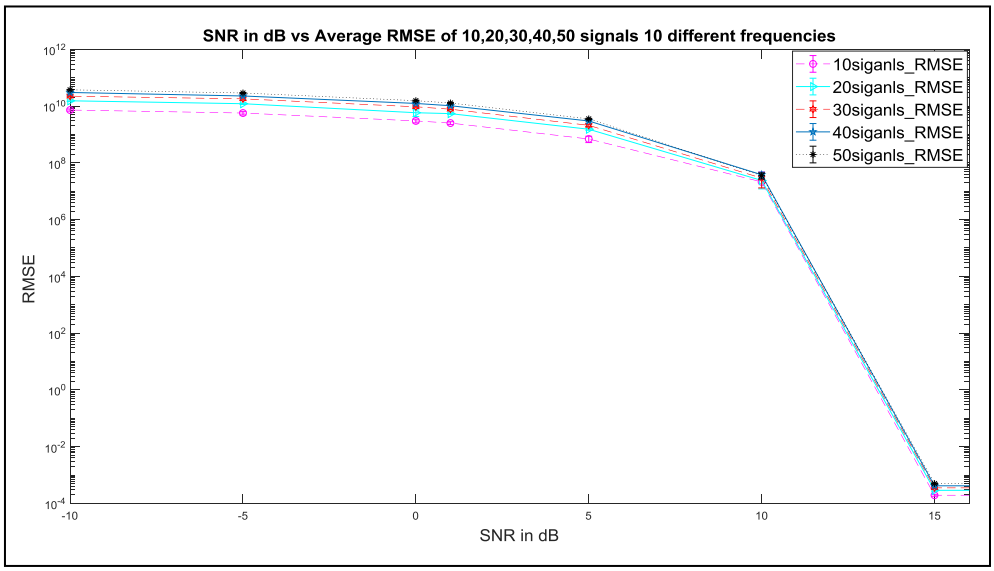


Figure. 3.3 RMSE v/s SNR for an ensemble of 10, 20, 30, 40, 50 sets of sensor signals over QPSK.

The communication channel is modeled over AWGN. Further, the RMSE of 10, 20, 30, 40, 50 signals ensemble is shown in Figure 3.3, and the BER analysis is shown in Figure 3.4. The study shows that the RMSE value is 10^{-3} at the SNR value of 15 dB, and below that SNR value, the RMSE increases. While, BER is found to be 10^{-3} at the SNR value of 10dB, and BER increases below 10 dB. Here, the lower order ensemble shows better RMSE and BER as compared to higher-order ensemble as a greater number of data points are not regressed satisfactory as PCA is a lossy compression technique.

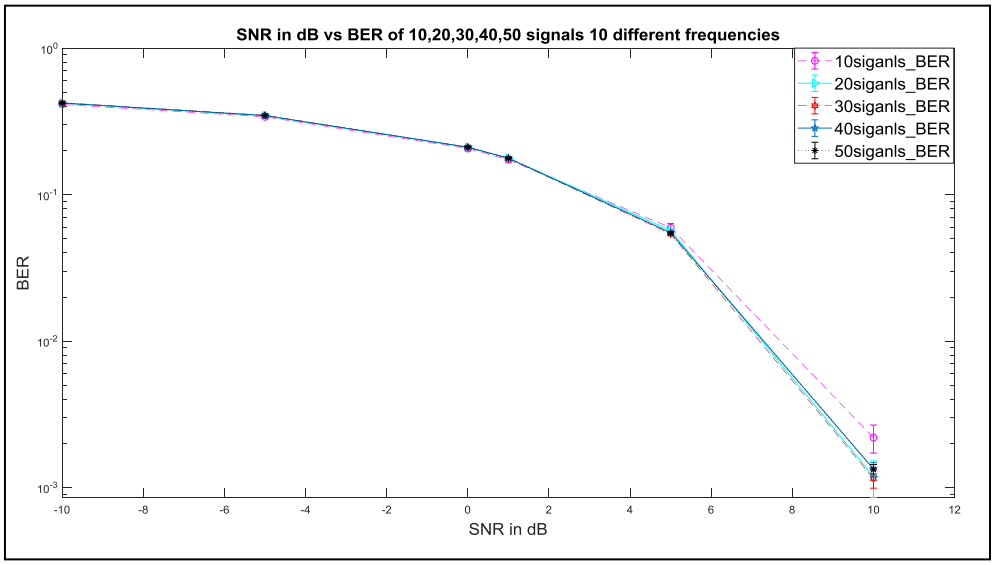


Figure. 3.4 BERv/s SNR for an ensemble of 10, 20, 30, 40, 50 sets of sensor signals over QPSK.

Scenario II:

Here, a wireless sensor network consisting of an ensemble of 10, 20, 30, 40, and 50 sets of sensor signals for 100 data points to be transmitted over QAM modulation. The communication channel is modeled over AWGN. Whereas the RMSE of 10, 20, 30, 40, and 50 signals ensemble is shown in figure 3.5 and the BER analysis is shown in Figure 3.6. The study shows that the RMSE range is 10^{-3} up to the SNR of 15 dB, and below that SNR, the RMSE increases drastically. However, BER is found to be $5x10^{-4}$at

the SNR value of 10dB, and it increases below 10 dB. While the lower order ensemble shows better RMSE and BER as compared to higher-order ensemble due to more number of data points, that are not regressed satisfactory as PCA is a loss compression technique.

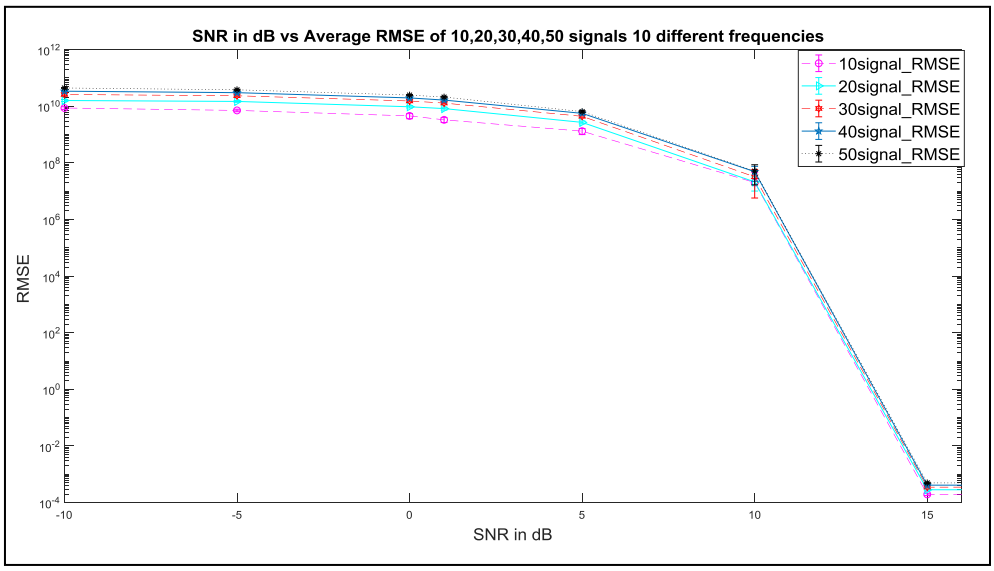


Figure. 3.5 RMSE v/s SNR for an ensemble of 10, 20, 30, 40, 50 sets of sensor signals over 4-QAM.

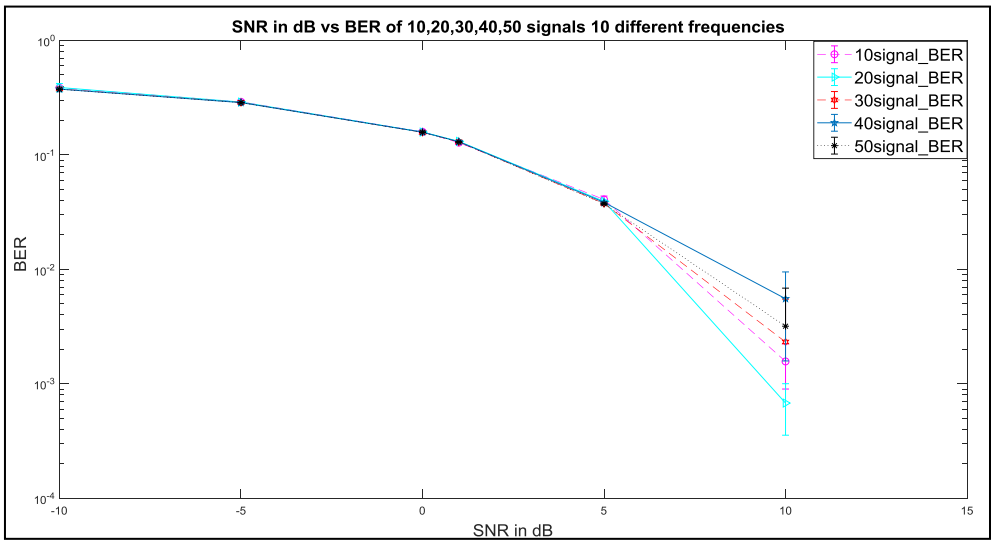


Figure. 3.6 BER v/s SNR for an ensemble of 10, 20, 30, 40, 50 sets of sensor signals over 4-QAM.

Scenario III:

A wireless sensor network consisting of an ensemble of 20 sets of sensor signals for 100 data points to be transmitted over 2, 4, 8, 16, 32 QPSK modulation. Here, the communication channel is modeled for AWGN. Further, the RMSE of 20 signals ensemble is shown in Figure 3.7, and the BER analysis is shown in Figure 3.8. However, the study shows that the RMSE range is from $3x10^{-3}$ for 32 PSK at the SNR value of 30 dB to 2PSK at the SNR value of 10dB. Whereas, the lower order modulation levels show better RMSE and BER as compared to higher-order modulation levels at the value of M (M= 2, 4, 8, 16, 32). Hence increase in the value results into more number of combined bits, wherein to make a symbol and as these bits are packed more closely in the signal constellation, which demands higher SNR for transmission. Similar observations are seen with reference to BER values, which state that lower-order PSK shows better performance. Hence the BER can be below 0.005 for 2PSK at 5dB SNR and for 32PSK at 25dB SNR. And, also for 2PSK and other versions of PSK for lower SNR, the BER will increase.

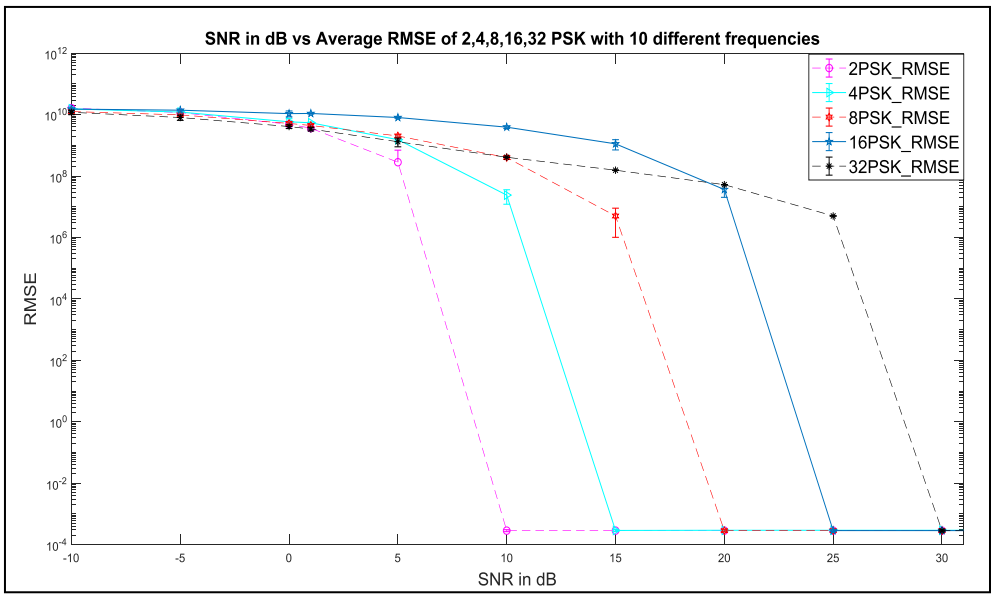


Figure. 3.7 RMSE v/s SNR for an ensemble of 20 sets of sensor signals over 2, 4, 8, 16, 32 QPSK.

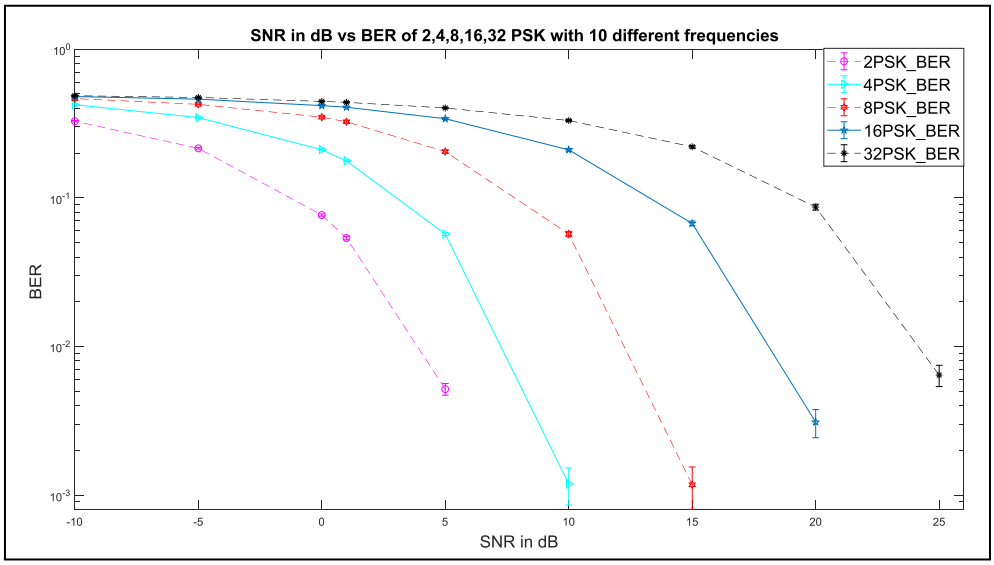


Figure. 3.8 BER v/s SNR for an ensemble of 20 sets of sensor signals over 2, 4, 8, 16, 32 QPSK.

Scenario IV:

Here the sensors network consisting of an ensemble of 20 sets of sensor signals for 100 data points to be transmitted over 2, 4, 8, 16, 32 QAM modulation. Also, the communication channel is modeled for AWGN. The RMSE of 20 signals ensemble is shown in Figure 3.9, and the BER analysis is shown in Figure 3.10. Further, the study shows that the RMSE range is from $2x10^{-4}$ for 32 QAM at the SNR value of 25 dB to 2QAM at the SNR value of 5dB. However, it performs better as compared to QPSK modulation, as seen in Figure. 3.7 and 3.8. While, the lower order modulation levels show better RMSE and BER as compared to higher-order modulation levels due to the increase in value of M (M= 2, 4, 8, 16, 32), which results in more number of bits which combined to make a symbol, and these bits are packed more closely in the signal constellation, which demands higher SNR for transmission. Similar observations are seen with reference to BER values, which states that lower order QAM shows better performance. This reveals that BER can be below 0.0008 for 2QAM

at 5dB SNR value and 32QAM at 20dBSNR value. Also, for 2QAM and other versions of PSK for lower SNR, the BER will increase.

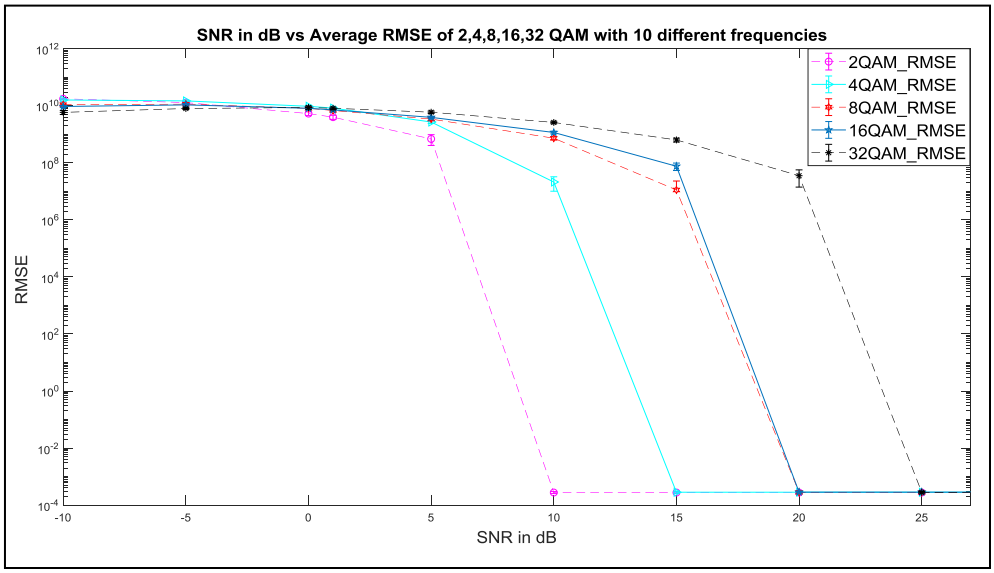


Figure. 3.9 RMSE v/s SNR for ensemble of 20 sets of sensor signals over 2, 4, 8, 16, 32 QAM.

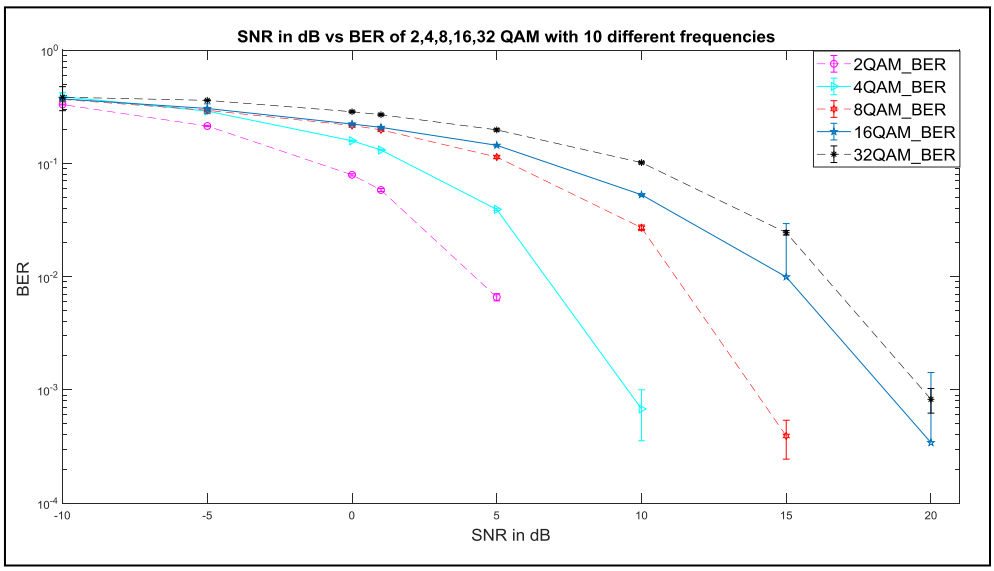


Figure. 3.10 BER v/s SNR for an ensemble of 20 sets of sensor signals over 2, 4, 8, 16, 32 QAM.

Scenario V:

Here the sensor networks consisting of an ensemble of 20 set of sensor signals for 100 data points to be transmitted over QPSK and 4-QAM modulation with Rayleigh fading ($5e^{-7}$ delay).

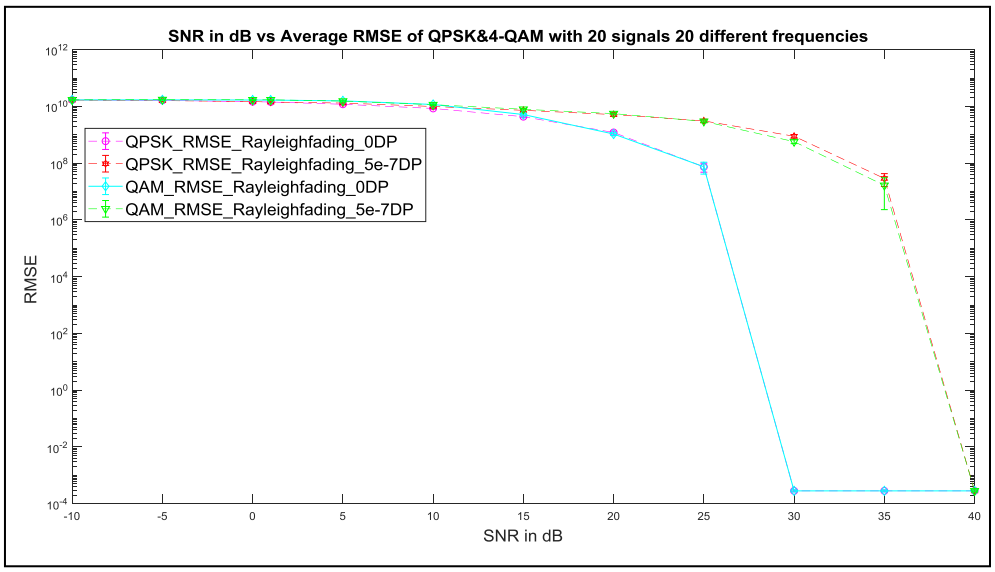


Figure. 3.11 RMSE v/s SNR for an ensemble of 20 sets of sensor signals over QPSK, 4-QAM with Rayleigh fading at 0, 5e-7 delay path.

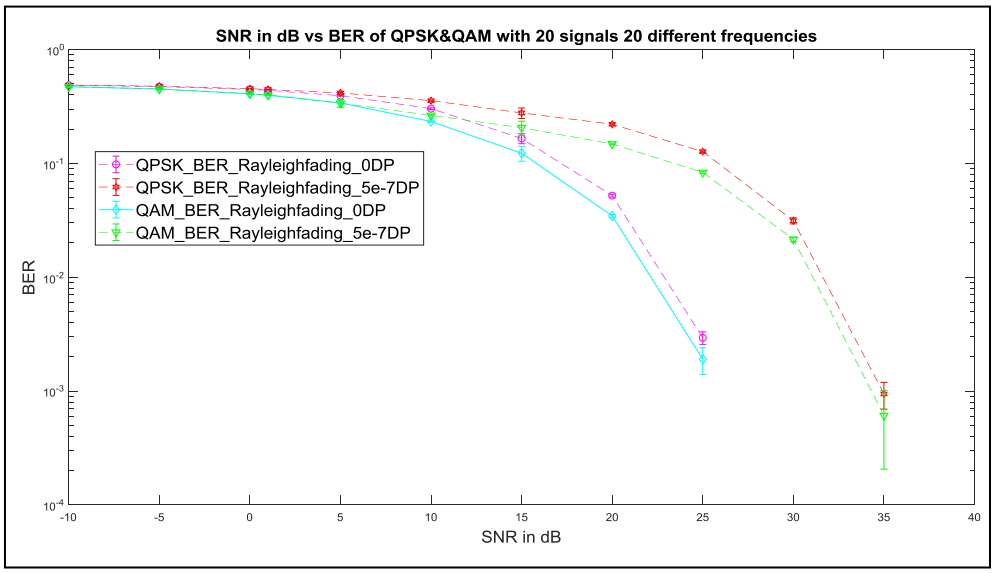


Figure. 3.12 BER v/s SNR for an ensemble of 20 sets of sensor signals over QPSK, 4-QAM with Rayleigh fading at 0, 5e-7 delay path.

The communication channel is modeled over AWGN. The RMSE of 20 signals ensemble is shown in Figure 3.11 and the BER analysis is shown in Figure 3.12. The study shows that the RMSE range is $3x10^{-4}$ up to the SNR of 30dB SNR value without Rayleigh fading and 40dB SNR value with Rayleigh fading. For lower SNR values, modulation with Rayleigh fading RMSE increases. However, a similar trend is seen for the BER. Here, the BER is 0.001 at the 25dB and 35dB, respectively, without and with Rayleigh fading for QAM and QPSK modulation.

Scenario VI:

Here a sensor network consisting of an ensemble of 20 sets of sensor signals for 100 data points to be transmitted over QPSK and 4-QAM modulation with Rayleigh fading ($5e^{-7}$ delay path Rayleigh fading) over 1x1, 2x2, 3x3 and 4x4 MIMO configuration.

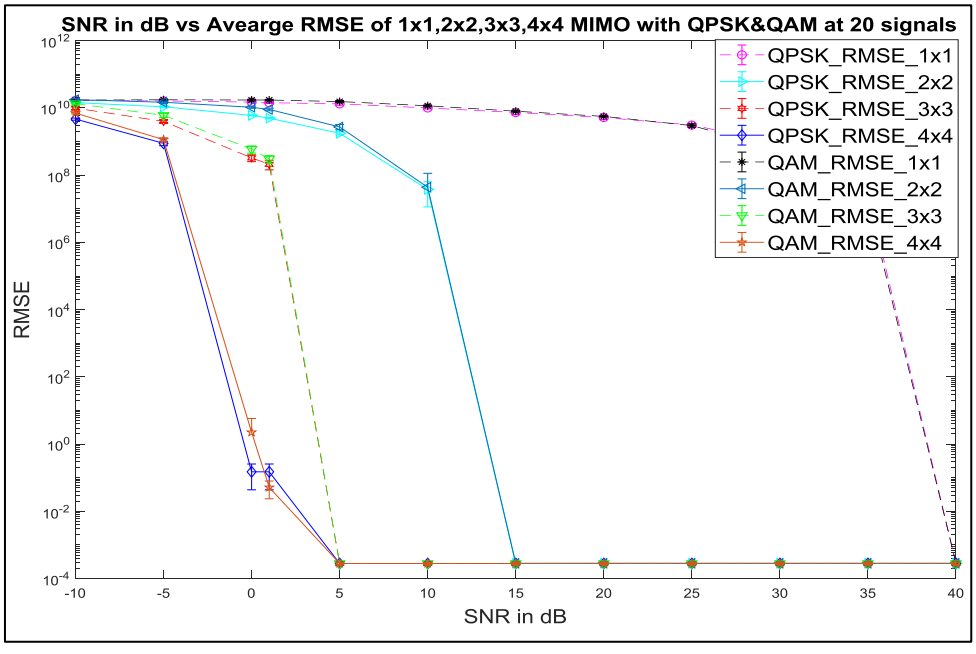


Figure. 3.13 RMSE v/s SNR for an ensemble of 20 sets of sensor signals over QPSK, 4-QAM with Rayleigh fading at 0, 5e-7 delay path at channel 1x1, 2x2, 3x3, 4x4 MIMO.

The communication channel is modeled for AWGN, the RMSE of 20 signals ensemble is shown in Figure 3.13, and the BER analysis is shown in Figure 3.14. Here, the study shows that the RMSE range is $2x10^{-4}$ up to the SNR of 40dB for 1x1 QPSK and 4-QAM, which is identical with the Figure 3.11 and 3.12. As we increase the MIMO configuration for higher-order, the system performance improves from 40dB to 0.0dB for similar RMSE of $2x10^{-4}$. This reveals that, low power transmission can be performed at the higher order of MIMO configuration. Similarly, BER is found to be 10^{-3}at the 35dB SNR value 1x1 QPSK and 4-QAM. Therefore, the BER performance is improvised at the higher order of MIMO configuration, supporting lower power transmission.

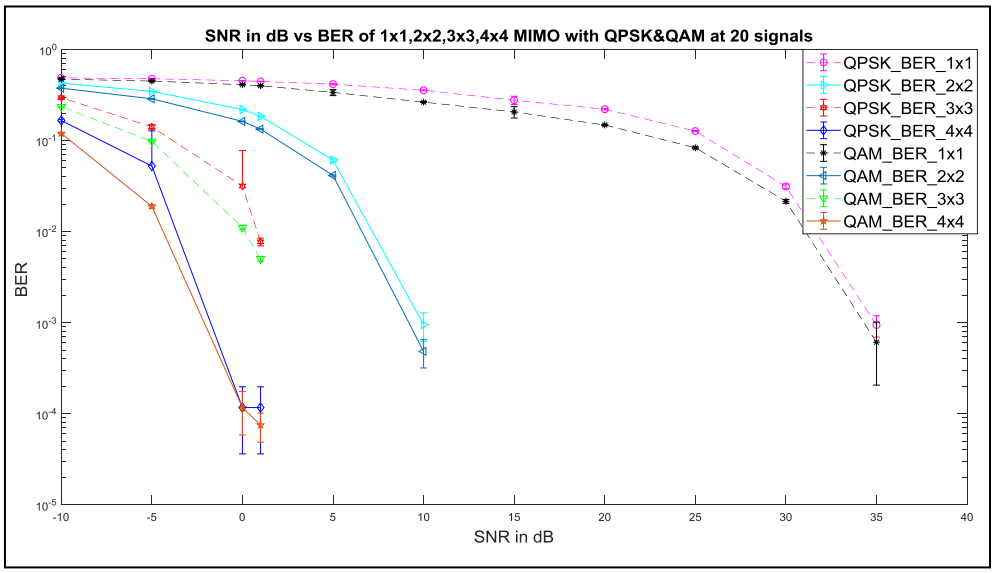


Figure. 3.14 BER v/s SNR for an ensemble of 20 sets of sensor signals over QPSK, 4-QAM with Rayleigh fading at 0, 5e-7 delay path at channel 1x1, 2x2, 3x3, 4x4 MIMO.

3.3. Discussion

Here, in the data organization, the 1-D sensor data is first formatted to a 2-D matrix. However, the use of matrix decomposition is to transform the said data into a new space, in which certain feature coefficients can accumulate most of the information.

Whereas, certain compression cases have been learnt to use PCA [16]-[17], or a method focused on sparsity [18]-[22].

Linear Discriminant Analysis (LDA) utilizes the label information in finding an informative projection. It is also used to find a linear combination of features which characterizes or separates two or more classes of objects or events. LDA seeks to reduce dimensionality while preserving as much of the class discriminatory information as possible. Therefore LDA is also closely related to Principal Component Analysis (PCA) and factor analysis in that both look for linear combinations of variables that best explain the data. It explicitly attempts to model the difference between the classes of data. PCA, on the other hand, does not take into account any difference in class, and factor analysis builds the feature combinations based on differences rather than similarities. However, discriminant analysis is also different from factor analysis as it is not an interdependent technique. Hence, a distinction between independent variables and dependent variables (also called criterion variables) must be made [23].

In PCA, the factor analysis builds the feature combinations based on differences rather than similarities in LDA. The discriminant analysis as done in LDA is different from the factor analysis done in PCA where eigenvalues, eigenvectors and covariance matrix are used. PCA tends to outperform LDA in almost all cases and hence PCA can be adopted as an effective tool for dimension reduction [23]. That means PCA used for feature classification and LDA used for data classification. Secondly the present research showed the PCA successfully reconstruct the data after compression or data reduction process as show in results with remarkable performance.

Dwivedi et al. presented the work on the 2 Demixed PCA (DPCA) concepts for color image compression. While the approach is checked on several standard images and demonstrates, the image quality is higher than traditional PCA-based data compression. Other performance measures, such as time compression, are improved.

Therefore, a comparative analysis is performed using 2DPCA for data compression [24-25]. Further, L. Vasa et al. provide an expanded PCA-based dynamical mesh compression strategy, while evaluating the total efficiency of a compression algorithm, the functions on its scale cannot be underestimated. In this connection, the problem is addressed, and the most advanced algorithms such as LPC are provided, and they proved this by offering a new approach, which reduces the basis size by 90% with regard to direct encoding, which can increase the compression algorithm performance by around 25% [26].

Furthermore, K.N Abdul, Kader Nihal et al. have researched that an image can be reconstructed with different data compression techniques, i.e., multiple redundancies. Lossless and Lossy, are the compression method which can be improved by using PCA in two separate ways, through the mathematical and the PCA Neural approach. Generally, the PCA is used for data reduction and interpretation [27]. However, Maryam Imani et al. found that PCA is one of the most traditional non-controlled methods for extracting the most efficient features [28].

Sunita S Biswal et al. reveals, PCA for the compression of images with four types of PCA algorithms, i.e., 2D-PCA, 3D-PCA, 2D- Kernel PCA, 3DKPCA. The efficiency of the reconstructed 3DKPCA image has been found to be higher than other kinds of PCA-based compression [29]. Sanjay Kumar et al. researched on PCA, LDA, and ICA as methods for the analysis of patterns of different data forms, i.e., images and collections of data. Here, these techniques are also useful to classify the object according to its extracted pattern [30] [31].

The regular PCA operates on 1D vector, which presents a problem of managing high dimensional space data such as images, and 2DPCA operates specifically on matrices, i.e., in 2DPCA data reduction is applied directly to the original image, without converting the image into a 1D. However, 2DPCA function has the advantage in terms

of handling higher dimensional vector space data over regular PCA. Moreover, there are also several other variants of 2DPCA, which effectively combine the proposed method with two 2DPCA-based techniques [32] [33]. From the above discussion, it is concluded that PCA is a technique that is used to increase efficiency, while LDA is used as a classifier for better image quality.

In this chapter, the study on data reduction techniques for types of signals having different frequencies signals has been carried out. However, similar findings can be considered with an appropriate methodology with an emphasis on the PCA algorithm. The study also reveals that the PCA is an advantageous computational method with almost unlimited potential for future use. Further, it demonstrates the signal extraction and data reduction, regression of original signal data capabilities. Furthermore, the study reveals that the PCA technique is more efficient than the ICA, LDA technique for dimension reduction due to the supervised nature of its algorithm. However, the research in the field of PCA is still very active currently, giving hope for further improvements of the method and its algorithmic implementations in data communication systems [34] [35].

The performance for efficiency and reliability has been evaluated for data reduction and regression using five ensemble sets consisting of number signals (10, 20, 30, 40, 50) having signal frequencies ranging from 50-2500Hz, using the RMSE and BER by varying the SNR values from -10dB to +40 dB over.

The data reduction using PCA is studied using five ensemble sets consisting of the number of signals (10, 20, 30, 40, 50) for QPSK and QAM (2, 4, 8, 16, 32 constellation) modulation. Further, the RMSE and BER studies for various SNR in the AWGN channel are performed. This study is carried out for 10 fold iterations to compute the average RMSE and BER performance. Also, the performance of the AWGN channel with Rayleigh fading for the QPSK and QAM modulation is performed. Further, the

studies are extended for the various configuration of MIMO, i.e., 1x1, 2x2, 3x3, 4x4, and demonstrated that the data transmission could be performed at very low SNR to enhance the capacity of channel. The workflow of the source coding with PCA data reduction, having MIMO for the communication system is discussed.

In conclusion we can say that the PCA dimensionality reduction is applied to the sensor level signal data compression, and parameters like BER and RMSE have been evaluated. It was also observed that the amount of compression columns used in PCA improves the spatial consistency of the signal data. Therefore, PCA compression is similar to the most widely used compression technique. The study concludes that PCA technique can be used in data reduction for an ensemble of sets of signals for efficient data communication over MIMO configurations. It can be noted that the present research work is useful for future WSN, IoT, LTE, 5G, 6G wireless communication platforms.

3.4. References

[1] Zheng, J., Wang, P. and Li, C. Distributed data aggregation using Slepian–Wolf coding in cluster-based wireless sensor networks. *IEEE Transactions on Vehicular Technology*, *59*(5):2564-2574, 2010.

[2] He, Z., Lee, B.S. and Wang, X.S. Aggregation in sensor networks with a user-provided quality of service goal. *Information Sciences*, *178*(9):2128-2149, 2008.

[3] Jiang, H., Jin, S. and Wang, C. Prediction or not? An energy-efficient framework for clustering-based data collection in wireless sensor networks. *IEEE Transactions on Parallel and Distributed Systems*, *22*(6):1064-1071, 2010.

[4] H. Li, K. Lin, K. Li. Energy-efficient and high-accuracy secure data aggregation in wireless sensor networks. *Compute Communication*, 34(4):591–597, 2011.

[5] Caione, C., Brunelli, D. and Benini, L. Distributed compressive sampling for lifetime optimization in dense wireless sensor networks. *IEEE Transactions on Industrial Informatics*, *8*(1):30-40, 2011.

[6] Quer, Y.G., Masiero, R., Rossi, M. and Zorzi, M. Sensing, compression and recovery for wireless sensor networks: monitoring framework design. *Wireless Communications, IEEE Transactions*, *11*:3447-3461, 2012.

[7] Hua, G. and Chen, C.W. Distributed source coding in wireless sensor networks. In *Second International Conference on Quality of Service in Heterogeneous Wired/Wireless Networks (QSHINE'05)*, Pages 1-7, 2005.

[8] Le Borgne, Y.A., Raybaud, S. and Bontempi, G. Distributed principal component analysis for wireless sensor networks. *Sensors*, *8*(8):4821-4850, 2008.

[9] Seo, S., Kang, J. and Ryu, K.H. Multivariate stream data reduction in sensor network applications. In *International Conference on Embedded and Ubiquitous Computing*, Pages 198-207. Springer, Berlin, Heidelberg, 2005.

[10] Bai, Z., Demmel, J., Dongarra, J., Ruhe, A. and van der Vorst, H. eds.*Templates for the solution of algebraic eigenvalue problems: a practical guide*. Society for Industrial and Applied Mathematics, 2000.

[11] Rooshenas, A., Rabiee, H.R., Movaghar, A. and Naderi, M.Y. Reducing the data transmission in wireless sensor networks using the principal component analysis. In *2010 Sixth International Conference on Intelligent Sensors, Sensor Networks and Information Processing, Pages* 133-138, 2010.

[12] Rassam, M.A., Zainal, A. and Maarof, M.A . Principal component analysis–based data reduction model for wireless sensor networks. *International Journal of Ad Hoc and Ubiquitous Computing*, *18*(1-2):85-101, 2015.

[13] Fenxiong, C., Fei, W. and Hongdong, J. Algorithm of data compression based on multiple principal component analysis over the WSN. *6th International Conference on Wireless Communications Networking and Mobile Computing (WiCOM)*, Chengdu, China, September, pp.1–4, 2010.

[14] Carvalho, C., Gomes, D.G., Agoulmine, N. and de Souza, J.N.Improving prediction accuracy for WSN data reduction by applying multivariate spatio-temporal correlation. *Sensors,* 11(11):10010–10037, 2011.

[15] Silva, O., Aquino, A.L.L., Mini, R.A.F. and Figureueiredo, C.M.S. Multivariate reduction in wireless sensor networks. In *IEEE Symposium on Computers and Communications*, Pages 726–729 *ISCC 2009*, Sousse, Tunisia, 2009.

[16] C. Kwan, R. Xu, and L. Haynes. A New Data Clustering Technique and Its Applications. *Data Mining and Knowledge Discovery: Theory, Tools, and Technology III, Proc. SPIE* , 4384:1-5, 2001.

[17] H. Hotelling. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24:417–441, and 498–520, 1933.

[18] M. Elad. *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. Springer, 2010.

[19] Y. Qu, H. Qi, B. Ayhan, C. Kwan, and R. Kidd. Does Multispectral/Hyperspectral Pansharpening Improve the Performance of Anomaly Detection. In *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Pages 6130-6133, 2017

[20] Dao, M., Kwan, C., Koperski, K. and Marchisio, G. A joint sparsity approach to tunnel activity monitoring using high resolution satellite images. In *2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON)*, Pages 322-328, 2017.

[21] Kwan, C., Budavari, B., Dao, M. and Zhou, J. New sparsity based pansharpening algorithms for hyperspectral images. In *2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON)*, Pages 88-93, 2017.

[22] 7Z, at http://www.7-zip.org/

[23] Shereena, V.B. and David, J.M. Comparative Study Of Dimensionality Reduction Techniques Using Pca And Lda For Content Based Image Retrieval. *Computer Science & Information Technology*, Page 41-55, 2015

[24] Deng, H.B., Jin, L.W., Zhen, L.X. and Huang, J.C. A new facial expression recognition method based on local Gabor filter bank and PCA plus LDA. *International Journal of Information Technology*, *11*(11):86-96, 2005.

[25] Ashutosh Dwivedi, Arvind Tolambiya, Prabhanjan Kandula, N Subhash Chandra Bose, Ashiwani Kumar, Prem K Kalra. Color Image Compression Using 2-Dimensional Principal Component Analysis (2DPCA). *Proceedings of ASID*, Pages 488-491, 2006.

[26] L. Vasa and V. Skala. COBRA: Compression of the Basis for PCA Represented Animations. *Computer Graphics Forum*, 8(6):1529– 1540, 2009.

[27] K.N. Abdul Kader Nihal, Dr. A.R. Mohamed Shanavas .A Survey and Study of Image Compression Methods. *IOSR Journal of Computer Engineering (IOSR-JCE)*, 16(4):11-16, 2014.

[28] Imani, M. and Ghassemian, H. Principal component discriminant analysis for feature extraction and classification of hyperspectral images. In *2014 Iranian Conference on Intelligent Systems (ICIS)*, Pages 1-5, 2014.

[29] Biswal, S.S., Kalpita, K. and Swain, D.R., 2014. Comparative study on image compression using various principal component analysis algorithms. *International Journal of Scientific and Research Publications*, *4*(5):1-4, 2014.

[30] Kumar, S. and Chauhan, E.A., 2014. A survey on image feature selection techniques. *International Journal of Computer Science and Information Technologies (IJCSIT)*, *5*(5):6449-6452, 2014.

[31] Neethu Mohan. Removal of PCA Based Estimated Noise in Processed Images. *International Journal of Science and Research (IJSR)*, 3(10): 897-899, 2014.

[32] Ms. Pallavi M. Sune Prof. Vijaya K. Shandilya.Image Compression Techniques based On Wavelet and Huffman Coding. *International Journal of Advanced Research in Computer Science and Software Engineering*, 3(4):1-5, 2013.

[33] Bontempi, G. and Le, B. An adaptive modular approach to the mining of sensor network data. *Proceedings of 1st International Workshop on Data Mining in Sensor Networks as part of the SIAM International Conference on Data Mining*, Pages 3-9, CA, USA, 2005.

[34] A. R. Shinwari, A. Jalali Balooch, A. A. Alariki and S. Abduljalil Abdulhak. A Comparative Study of Face Recognition Algorithms under Facial Expression and

Illumination. *2019 21st International Conference on Advanced Communication Technology (ICACT)*, Pages 390-394, 2019.

[35] J. Yang, D. Zhang and J. Yang. Constructing PCA Baseline Algorithms to Reevaluate ICA-Based Face-Recognition Performance. In *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 37(4):1015-1021, 2007.

Chapter 4

MIMO Channel with LDPC channel coding

4.1. Introduction

In the modern era, wireless communication systems enhance the network capacity to support a wide range of data. The present work describes the data reduction methodology so as to enhance the channel capacity based on the PCA over MIMO for untether communication system. At the same time, the communication system has employed M-PSK/M-QAM (M=2, 4, 8, 16, 32) schemes with LDPC channel coding method for error correction over AWGN noisy channel. However, the proposed method here utilizes the PCA for data reduction, which transmits only the feature vectors, i.e., Eigen Vector, for the set of composite data as an ensemble for large numbers of signals over time-stamped. The effectiveness of the proposed work is evident through the simulation results of the systems incorporating 1x1, 2x2, 3x3, 4x4 MIMO networks. Hence, the RMSE analysis is carried out for the PCA regression to verify the signal fidelity after regression and the BER for confirmation of the recovery of the data over the MIMO channel and LDPC channel coding.

The present chapter is committed to the study of data at the sensorlevels, using composite data transmitted using PCA features vectors employing M-PSK/M-QAM (M=2, 4, 8, 16, 32) modulation schemes and LDPC error correction over MIMO Channel.

To present the sensor signal level data fusion of 45 signals originating from various sources were selected. These PCA encoded signals are modulated with M-PSK/M-QAM and transmitted over the MIMO channel. The noisy channel is introduced with the help of AWGN noise. The evaluation of the model represents the RMSE analysis for the PCA regression of the original composite signals, for verification of the signal fidelity. Also, the BER for confirmation of the recovery of the data over the MIMO channel over LDPC coding and decoding.

There are four types of MIMO communication channel used here having multiple antenna systems, i.e., SISO, SIMO, MISO, and MIMO [1, 2].

4.1.1. SISO

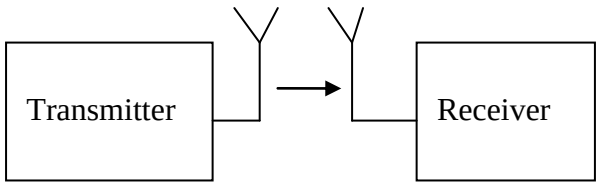


Figure. 4.1 Single Input Single Output (SISO).

The existing technology shown in Figure 4.1 is Single Input Single Output (SISO). Here, the transmitter has one antenna, and the receiver uses no diversity technique. Both the receiver and the transmitter have one RF chain (coder and modulator). The SISO is relatively easy and affordable to introduce and hence, used from several years since the advent of radio technology. It is used in radio and television transmissions and personal wireless (e.g. Wi-Fi and Bluetooth) technologies, under power constraints and BW restrictions [2].

4.1.2. SIMO

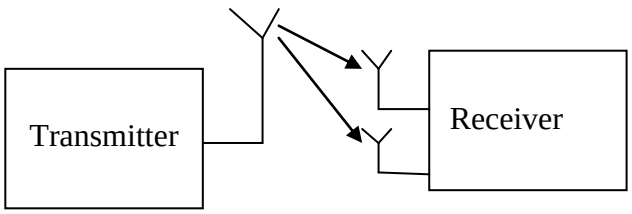


Figure. 4.2 Single Input Multi Output (SIMO).

Here, one antenna on the transmitter and two antennas on the receiver are used as a diversity technique, and the multiple antenna techniques have been developed to

improve the performance. A network that uses a single transmitter antenna and many antennas are referred in Figure 4.2 as a single input multiple outputs (SIMO). However, the receiver is either chosen as the best antenna for stronger signal receiving or combine signals from all antennas to maximize SNR. Therefore, known as switched diversity or diversity of selection hence, referred to as the maximal ratio combining (MRC) [2].

4.1.3. MISO

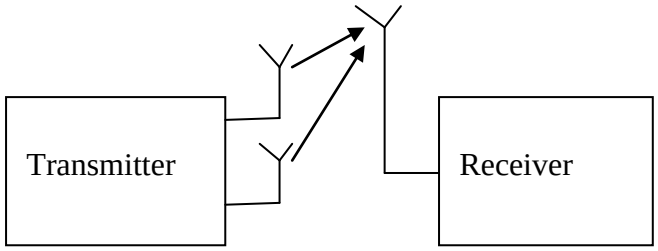


Figure. 4.3 Multi-Input Single Output (MISO).

Figure.4.3 describes the multiple-input single-output (MISO) system in the use of multiple antennas in the transmitter and a single antenna in the receiver. Here, Alamouti STC technique (Space-Time Coding) is used with two antennas on the transmitter. Whereas STC allows the transmitter to transmit signals both in time and space such that the information is transmitted two times in a row by two antennas. However, the SIMO or MISO antennas are usually placed at base stations (BS) with multiple antennas (each with an RF chain). Therefore, all BS subscriber stations (SSs) can share the cost of providing either receiving diversity (in SIMO) or transmitting diversity (in MISO) [2].

4.1.4. MIMO

In the radio, multiple input and multiple-output, or MIMO, multiply the network link capacity by multiple transmissions and receive multi-path antennas. The multiple antennas (correspondingly multiple RF chains) are added to both the transmitter and the

receiver to increase the radio link throughput. Therefore this system is called Multiple Input Output (MIMO) as shown in Figure 4.4. However, a MIMO system with a comparable count of antennas in a point-to-point (PTP) communication will linearly multiply the network throughput by each additional antenna, both for the transmitter and the receiver. For example, a 2x2 MIMO doubles the output. Both the transmitter and the receiver are transmitting and diversified by two antennas [2].

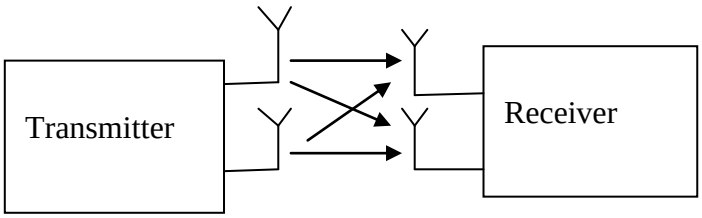


Figure. 4.4 Multi Input Multi Output (MIMO).

The MIMO function is to make wireless communications more reliable and use diversely. Diversity is the way to transmit the same data through different networks in order to achieve greater reliability. Even when one channel is unusable, the data can still be recovered from the redundant transmission via the other channels. Therefore the overall efficiency of the communication network is also increased, with redundant information being transferred. However, MIMO systems can achieve remarkable efficiency and power enhancements by taking advantage of the range of multiple channels between transmitting and receiving antennas.

4.1.5. Modeling the data fusion with MIMO system

There are many ways of data fusion, including the Kalman Filter (KF) with two broad approaches in KF fusion methods, i.e., fusion measurement and fusion of state-vector. Each sensor uses an estimator that calculates the state vector and its

corresponding covariance matrices from the sensor data. However, these state vectors are further transferred to the fusion centre via a data link. Vector fusion procedures use a group of Kalman filters to extract individual sensor-based status assessments that are fused into an enhanced joint status measurement. For any set of measurements, the KF is the algorithm used for each sensor (data) independently and produces state estimates.

In most cases, it is not possible to give the data in a one-hop real-time because of the limited bandwidth. Therefore the frequent transmission of these raw data results in a significant loss of data due to containment and congestion of channels. Hence the sensor nodes should be able to reduce data before transmission autonomously. Thus, suggested and modeled the Data Fusion over MIMO Sensor as shown in Figure. 4.5.

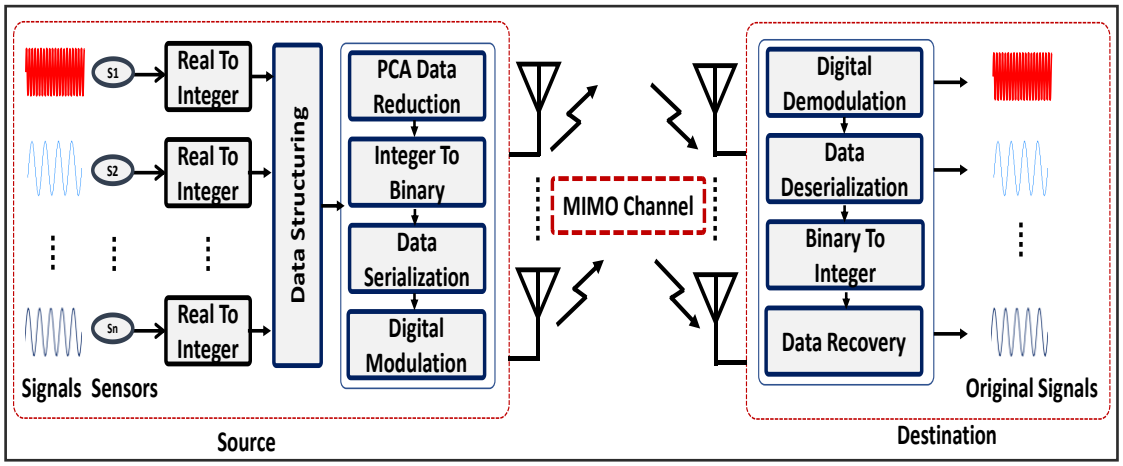


Figure. 4.5 Block diagram of source coding and data Fusion over MIMO.

Furthermore, the quantity of data transmitted must be minimized to enhance the monitoring quality and increase the capacity of the sensor node by predicting future data behavior, while detecting and classifying essential events in the sensor level signals, transmitting only necessary data. Thus the proposed model is to reduce the amount of data transmitted by way of PCA between sensor level signals and the end-user.

The channel capability of the wireless communication system and the dynamic transmitting environment can be interpreted mathematically using complex numbers, on the assumption that the channel is a flat fading path. Here , in-phase component is the real part of the complex signal data, and the quadrature component is the imaginary component [3]. However, for a SISO device, the whole transmission process may be reduced to a single complex number, as shown in the equation 4.1.

$$y = h\,x + e \qquad (4.1)$$

Where h, is the complex channel number, x is the input signal, e is a complex thermal noise modeling number at the receiver. Therefore the factors have the same value as the case of SISO. According to Shannon's theorem, equation 4.2 gives the capacity of a channel to transmit information free of errors.

$$C = B\,\log_2(1 + SNR) \qquad (4.2)$$

Where B is transmission bandwidth and SNR is a channel's signal-to-noise ratio. However, the calculation includes the channel's total maximum capacity (in bits/second). By increasing the bandwidth used in transmission or increasing the SNR, the channel capacity can be improved. Further, the use of Multi-antenna systems is a very innovative approach to improve wireless communications systems overall capacity with use of more channels [4]. Increasing of the individual transmitting channels is still limited by the equation above. However, the total system capacity is therefore the sum of the individual channel capacities.

Diversity and Space-Time Coding techniques one can use replicas of the data symbol to receive independently at the receivers end for accurate reception. Besides, MIMO is also an important feature for wireless networking like IEEE 802.11n, IEEE 802.11ac, HSPA+ (3 G), WiMAX (4 G). Over recent years, MIMO has been used as

part of the International Telecommunication Union (ITU) G.hn protocol and the Home Plug AV2 standards for 3-wire power-line communication [5-6].

4.1.6. Channel coding LDPC Min-Sum Algorithm

The algorithm encodes an input binary message k to the n-bit codeword. At the decoder side, min-sum algorithm is used as a message-passing algorithm. However, the decoder performs iterations for parity check, and Check Node (CN). While, the variable node (VN) is updated, as shown in step 2 and step 3, where LLR's (Log-Likelihood Ratio) are exchanged as messages between CN and VN. Further, the LLR of VN is performed by step 4. The options for decision types available in this algorithm are hard and soft decisions [7].

In the given equation, LDPC code of length L and dimension P where parity-check matrix M with A= L - P rows and L columns contains exactly 1's in each column (column weight) and row (row weight) M_{al} is the value of the a_{th} row and l_{th} column in M.The set of bits that participate in check is denoted Equation 4.3 [7].

$$L_a=\{ l:M_{al=1}\} \qquad (4.3)$$

The set of checks that participate in bits by equation 4.4.

$$A_l= \{ a:M_{al=1} \} \qquad (4.4)$$

Assume code word equated by equation 4.5,

$$K=[K_1,K_2,K_3,.....K_L]^T \qquad (4.5)$$

Before transmission, it is mapped to a signal constellation to obtain the vector given by equation 4.6,

$$t=[t_1,t_2,....t_l]^T \qquad (4.6)$$

where, $t_l = 2 * K_l - 1$

which is transmitted by an AWGN channel with equation 4.7 variance

$$\sigma^2 = {}^{L_0}/_2 \qquad (4.7)$$

Here, is the Additive White Gaussian Noise (AWGN) with zero means. Let hard decision vector given by equation 4.8 and 4.9,

$$Q=[Q_1,Q_2,Q_3,.....Q_L]^T \qquad (4.8)$$

$$Q_l=sgn(r_l) \qquad (4.9)$$

$$Where\ sgn(r_l) = \begin{cases} 1, & r_l \\ 0, & Otherwise \end{cases}$$

R_l: A priori information of bit node, l.

$\bar{R_l}$: A posteriori information of bit node, l.

$\Theta_{a,l}$: The check to bit message from a to l.

$\lambda_{l,a}$: The bit to check message from l to a.

Step 1: Initialization

A priori information, $R_{l=} -r_l$

The bit to check message initialization, $\lambda_{l,a} = R_l$

Step 2: Horizontal Step

Check node processing by equation 4.10:

$$\Theta_{a,l} = log \frac{1 + \Pi_{l' \epsilon L(a) \backslash l} \ tanh\left(\lambda_{l'.a}/2\right)}{1 - \Pi_{l' \epsilon L(a) \backslash l} \ tanh\left(\lambda_{l'.a}/2\right)} \qquad (4.10)$$

Step 3: Vertical Step

A posteriori information by equation 4.11:

$$\overline{R_l} = R_l + \sum_{a \epsilon A(l) \backslash a} \Theta_{a,l} \qquad (4.11)$$

Bit node processing equated by equation 4.12:

$$\lambda_{a,l} = \overline{R_l} + \sum_{a' \epsilon A(l) \backslash a} \Theta_{a,l} \qquad (4.12)$$

Step 4: Decoding Attempt

$$\overline{R_l} > 0, \overline{k_l} > 0, else \overline{k_l} = 0 \qquad (4.13)$$

If $M\overline{k_l} = 0$, the algorithm stops and $\overline{k_l}$ is taken by equation 4.13 as a true consequence of decoding.

Otherwise, it goes to the next iteration until the number of iteration reaches its maximum limit [7].

The min-Sum algorithm is the modified form of the sum-product algorithm that reduces the implementation complexity of the decoder.

This can be done by altering the Horizontal step given by equation 4.14:

$$\Theta_{a,l} = log \frac{1 + \Pi_{l' \epsilon L(a) \backslash l} \ tanh\left(\lambda_{l'.a}/2\right)}{1 - \Pi_{l' \epsilon L(a) \backslash l} \ tanh\left(\lambda_{l'.a}/2\right)} \qquad (4.14)$$

Using the relationship:

$$2 \ tanh^{-1} p = log \frac{1+p}{1-p}$$

Equation 4.14 can be rewritten as equations 4.15 and 4.16,

$$\Theta_{a,l} = 2 \ tanh^{-1} \prod_{l/\epsilon L(a \backslash l)} tanh \ (\lambda_{a,l}/2) \qquad (4.15)$$

$$\Theta_{a,l} = 2 \ tanh^{-1} \prod_{l/\epsilon L(a \backslash l)} sgn \ (\lambda_{a,l}/2) \prod_{l/\epsilon L(a \backslash l)} tanh\left(\frac{\lambda_{a,l}}{2}\right) \qquad (4.16)$$

The Min-sum algorithm simplifies the calculation of (4.16) even further by recognizing that the term corresponding to the smallest Fn, m dominates the product term and so a minimum can approximate the product is given by equation 4.17.

$$\Theta_{a,l} = \prod_{l/\epsilon L(a \backslash l)} sgn \ (\lambda_{a,l}/2) \qquad (4.17)$$

In this MIN-SUM algorithm, firstly initialized the bit to check the message. Further, update the check message in the horizontal step. Then, multiply the optimization factor with the check message and proceed to the vertical step. Later, the following information is updated with the help of the check message, and then updates the bit node. Further, multiply the optimization factor with the check message. Last step is the decision-making process. However, if the decoded code word is correct, stop and

take it as the output or repeat the whole decoding process until the iteration number reaches its maximum limit [7, 8].

The main goal in designing a communication system is to achieve reliable data transmission for the power efficiency having the lowest error probability (bit error rate). Moreover, a higher data rate with a constraint on available bandwidth is another target. At the same time, the LDPC codes can be selected as an excellent coding scheme to achieve the highest reliability transmission in noisy channels. On the other hand, in terms of efficient use of bandwidth while having a high data rate to achieve functional channel capacity should be explored. Therefore, motivated by the development of low power and enhance channel capacity over low bandwidth for various SNR is a challenge. The present work tried to address these problems using source and channel coding method over the MIMO system to improvise the channel capacity by some factor in a wireless communication system.

The work is further modeled and demonstrated by the source coding method using PCA for sensor level fusion at the transmitter and regress the same at the receiver for improvising the channel capacity of the system as shown in Figure 4.6. Here, 45 signals are fused for the ensemble of sensor data having a frequency ranging from 50Hz - 2500Hz, and sampled at 5 kHz each. The efficiency of regression was measured using the RMSE of the transmitted and receive signals. And also, demonstrated the performance of BER for the M-PSK and M-QAM (M=2, 4, 8, 16, 32) modulation technique with LDPC decoder i.e., Min-Sum, for 2x2, 3x3, 4x4 MIMO channel. The performance parameters for the source coding and channel coding are evaluated over the AWGN channel ranging from -15dB to 40dB SNR value.

These works are devoted to the study of source coding and transmission over M-PSK and M-QAM modulation using LDPC Min-Sum Error Corrections over MIMO Channel.

- To present the source coding of signals originating from 45 signals for efficient transmission and reception over the communication channel for performance parameters like RMSE.
- The system evaluated for using M-PSK, M-QAM modulation for performance parameters like BER with LDPC decoder algorithm for error-free data transmission over the noisy channels.
- The system is further evaluated for the wireless communication system for data transmission with 2x2, 3x3, 4x4 MIMO channel, respectively, in the AWGN channel.

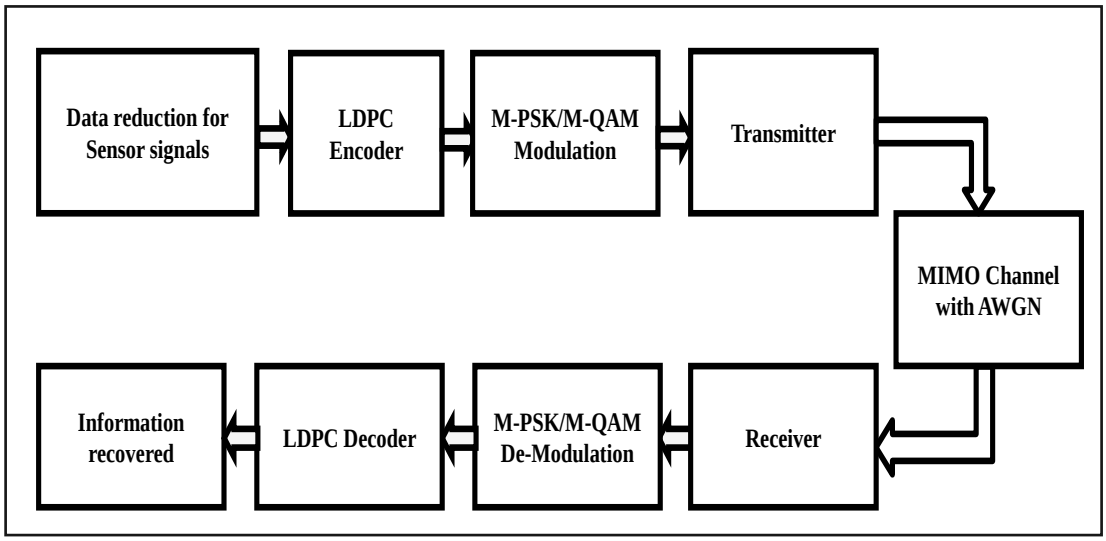


Figure. 4.6 The proposed Sensor's Data Transmission over M-PSK/M-QAM employing LDPC Error Corrections over MIMO Channels.

4.2. Experiments and Results

The present work designed an LDPC encoder and decoder, where the streams of data in a matrix form of the block length (100, 45) of 16-bitwordlength corresponding to 45 sensors signal over 100-time stamp sample. Further, it is reduced to (45, 45) of 16-bitwordlength using Eigen matrix corresponding i.e., corresponding to 45 sensors signal

data. It is alsoreduced to mapped the matrix to LDPC for the code rate of R = 1/2 to generate the block length of 64800 bits at the LDPC encoder having 32400 message bits. The encoded data from the LDPC encoder is further PSK and QAM modulator with M bits per symbol. Channel selected here is AWGN having SNR values from -15dB to 40 dB. The said system further tested for the performance over LDPC codes with 2x2, 3x3, 4x4 MIMO for wireless communication applications. The model details as specified in the experiments are illustrated in Table 4.1. The various scenarios of the experiments are given in Table 4.2. The processing order is mentioned in the flowchart as in Figure 4.7.

Table 4.1 Parameters used in data communication system

Parameters	**Value**
LDPC code rate	½ (32400 /64800)
LDPC Encoder Input Type	Bit
LDPC Decoder output Type	Information Part
LDPC Decoder Decision Input Type	Hard Decision
LDPC Decoder Number of Iterations	50
Modulator Type	M-PSK/M-QAM
Modulator Input Type	Bit
Modulator Symbol Order	Binary
Demodulator Output Type	Bit
LLR Algorithm	LLR
Channels	SISO & 2x2, 3x3, 4x4 MIMO
Channel Noise Factor	SNR

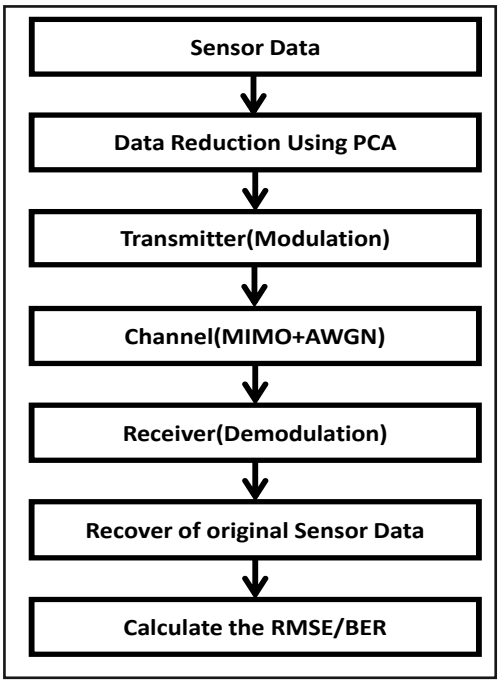


Figure. 4.7 Flowchart of the LDPC channel coding bases system.

Table 4.2 Results of different scenarios evaluation of RMSE and BER

Scenario I	RMSE and BER performance of LDPC code with code rate R = 1/2 over an AWGN channel with 2-PSK modulation
Scenario II	RMSE and BER performance of LDPC code with code rate R = 1/2 over an AWGN channel with 4-PSK modulation
Scenario III	RMSE and BER performance of LDPC code with code rate R = 1/2 over an AWGN channel with 8-PSK modulation
Scenario IV	RMSE and BER performance of LDPC code with code rate R = 1/2 over an AWGN channel with 16-PSK modulation
Scenario V	RMSE and BER performance of LDPC code with code rate R = 1/2 over an AWGN channel with32-PSK modulation
Scenario VI	RMSE and BER performance of LDPC code with code rate R = 1/2 over an AWGN channel with 4-QAM modulation
Scenario VI I	RMSE and BER performance of LDPC code with code rate R = 1/2 over an AWGN channel with8-QAM modulation
Scenario VIII	RMSE and BER performance of LDPC code with code rate R = 1/2 over an AWGN channel with 16-QAM modulation
Scenario IX	RMSE and BER performance of LDPC code with code rate R = 1/2 over an AWGN channel with 32-QAM modulation

Scenario I:

The RMSE performance is presented in Figure 4.8 for BPSK modulation with and without LDPC error correction over the AWGN channel having SNR value ranging from -15dB to +40dB. However, the performance of RMSE is better with the LDPC algorithm as compared without LDPC, for 1x1, 2x2, 3x3, and 4x4 MIMO channels. Further, the RMSE is 0.4e-4($0.4x10^{-4}$) for 1x1 LDPC MIMO at -03.0dB SNR value. And it is 0.4e-4($0.4x10^{-4}$) for 4x4 LDPC MIMO at -10 dB SNR value. Therefore, it demonstrates the error-free signal transmission up to -10dB with LDPC error correction coding.

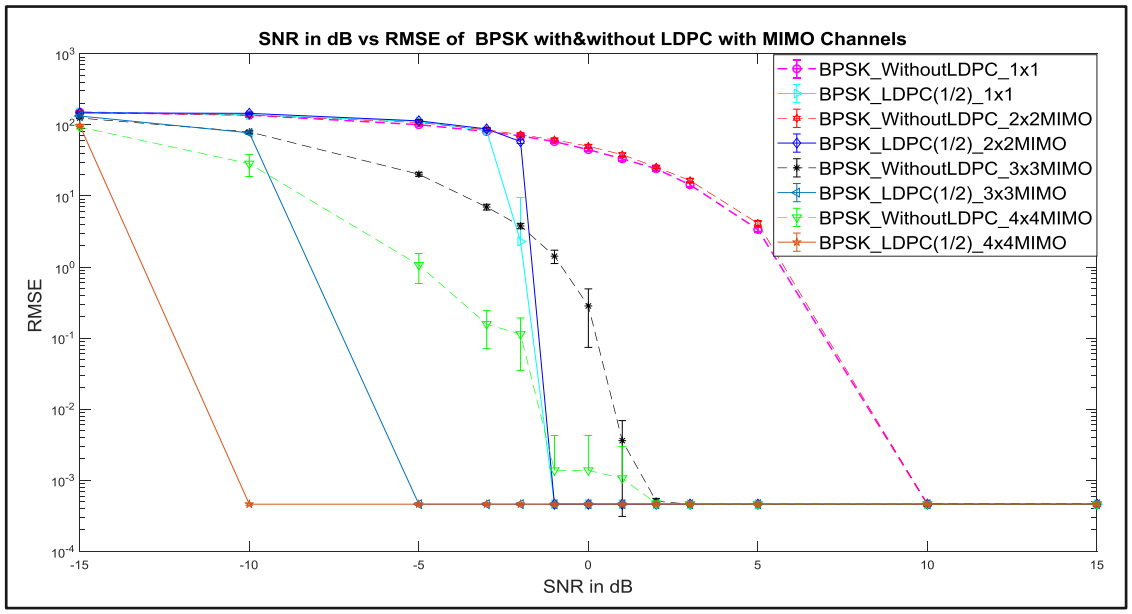


Figure. 4.8 RMSE performance of LDPC code with code rate R = 1/2 over an AWGN channel via BPSK modulation.

The BER performance is presented in Figure 4.9 for BPSK modulation with and without LDPC error correction over the AWGN channel having SNR -15dB to 40dB. Here, the performance of BER is better with the LDPC algorithm as compared to the without LDPC error corrections for 1x1, 2x2, 3x3, and 4x4 MIMO channels. Further,

the BER value is e-2(10^{-2}) at -7.5dB, -3dB, 4dB and 5dB respectively for 4x4, 3x3, 2x2, and 1x1 MIMO without LDPC. This implies that 4x4 is performing better compared to 1x1 MIMO. The BER value is 0.1e-1($0.1x10^{-2}$) at -15dB, -10dB, -3dB and -2dB respectively for 4x4, 3x3, 2x2, and 1x1 MIMO with LDPC error corrections. This indicates that 4x4 is performing better with LDPC compared to without LDPC as it can transmit the data at higher noise.

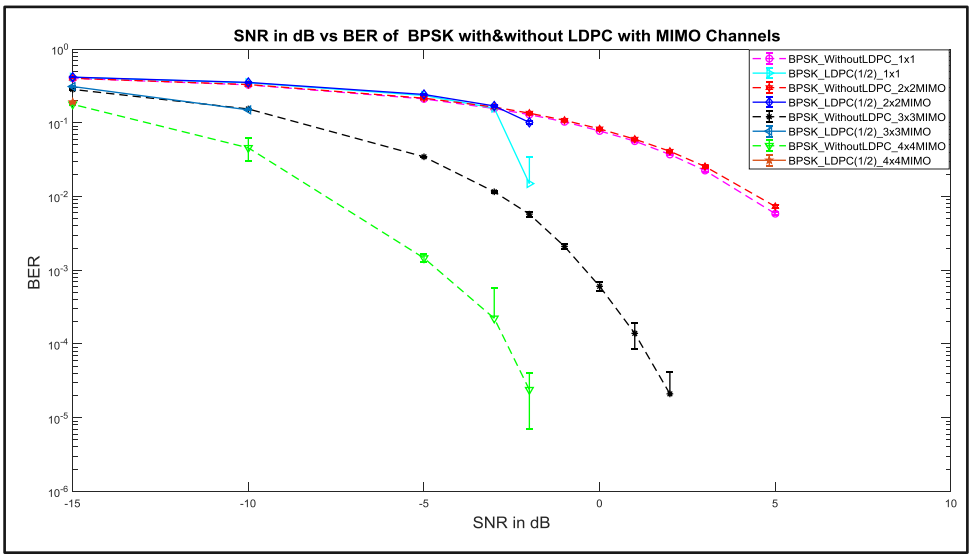


Figure. 4.9 BER performance of LDPC code with code rate R = 1/2 over an AWGN channel via BPSK modulation.

Scenario II:

The RMSE performance is presented in Figure 4.10 here the evolution of QPSK modulation performance with and without LDPC error correction over the AWGN channel having SNR -15dB to 40dB range. Here, the performance of RMSE is better with the LDPC algorithm as compared to the without LDPC error corrections for 1x1, 2x2, 3x3, and 4x4 MIMO channels. Here, the RMSE is 0.4e-4 for 1x1 LDPC MIMO at 5 dB for SNR value. And it is 0.4e-4 for 4x4 LDPC MIMO at -5 dB for SNR value. Therefore it states the error-free signals transmission up to -5dB in data communication system with LDPC error correction coding.

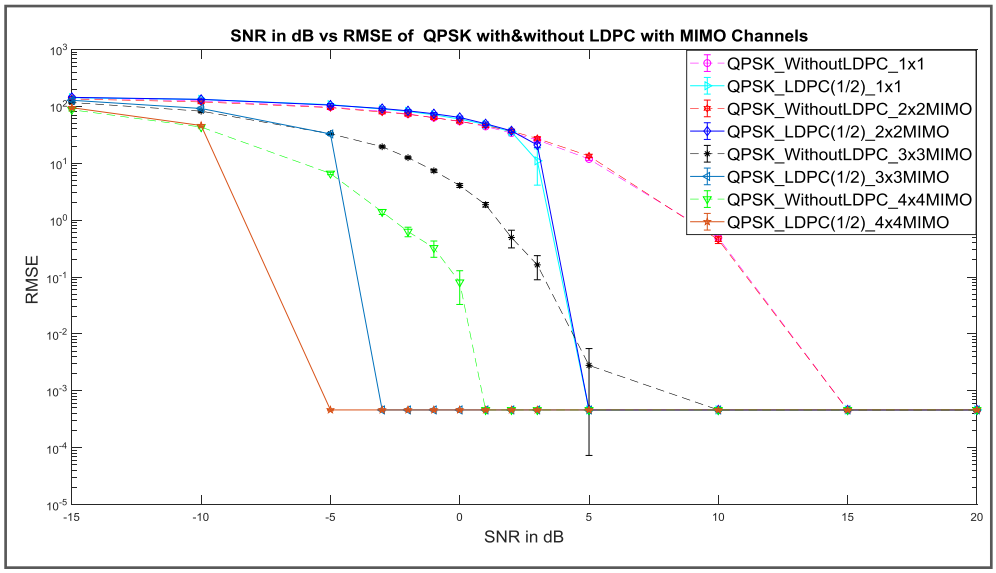


Figure. 4.10 RMSE performance of LDPC code with code rate R = 1/2 over an AWGN channel via PCA-QPSK modulation.

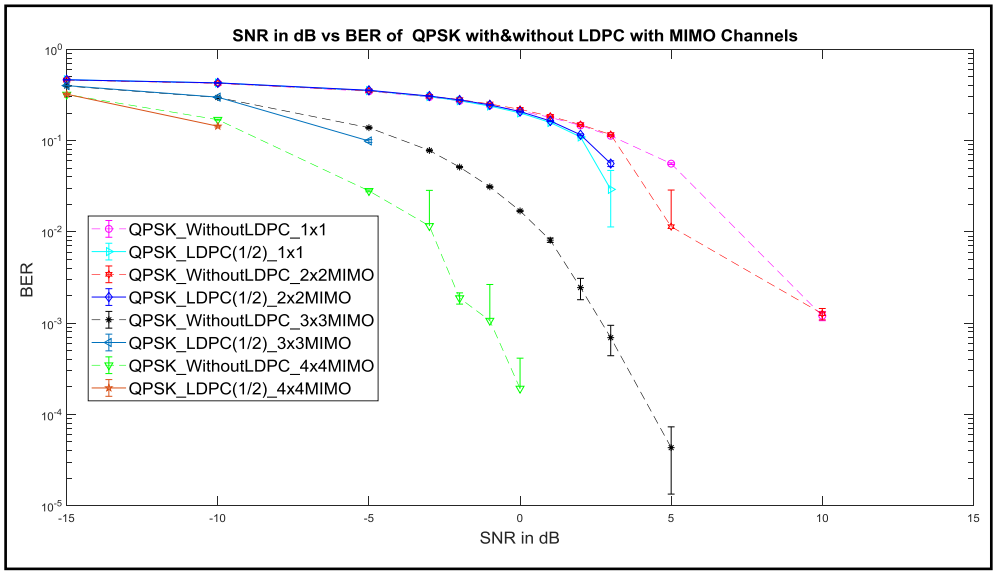


Figure. 4.11 BER performance LDPC code with code rate R = 1/2 over an AWGN channel via QPSK modulation.

The BER performance is presented in Figure 4.11 while the evolution of QPSK modulation performance with and without LDPC error correction over the AWGN

channel having SNR -15dB to 40dB. Here, the performance of BER is better with the LDPC algorithm as compared to without LDPC error corrections for 1x1, 2x2, 3x3, and 4x4 MIMO channels. The BER value at the e-2 is -3dB, 2dB, 7dB and 10dB respectively for 4x4, 3x3, 2x2, and 1x1 MIMO in data communication system without LDPC coding. This implies that 4x4 is performing better compared to 1x1 MIMO for error correction. Further, the BER value at the 0.1e-1 is -10dB, -3dB, 5dB and 6dB respectively for 4x4, 3x3, 2x2, and 1x1 MIMO with LDPC coding. This implies that 4x4 is performs better with LDPC compared to in data communication system without LDPC.

Scenario III:

The RMSE performance is shown in Figure 4.12 in which the evolution of 8-PSK modulation performance with and without LDPC error correction over the AWGN channels having SNR -15dB to 40dB range. Here, the performance of RMSE is better with the LDPC algorithm as compared without LDPC error corrections for 1x1, 2x2, 3x3, and 4x4 MIMO channels. The RMSE is 0.4e-4 for 1x1 LDPC MIMO at 20 dB for SNR value, and the RMSE is 0.4e-4 for 4x4 LDPC MIMO at 10 dB for SNR value. Therefore it states the error-free signal transmission up to 10dB with LDPC error correction coding.

The BER performance is shown in Figure 4.13 in which the evolution of 8-PSK modulation performance with and without LDPC error correction over the AWGN channels having SNR -15dB to 40dB. However, the performance of BER is better with the LDPC algorithm as compared to the without LDPC error corrections for 1x1, 2x2, 3x3, and 4x4 MIMO channels. The BER value of the e-2 is 3dB, 8dB, 10dB and 15dB respectively, for 4x4, 3x3, 2x2, and 1x1 MIMO. This implies that 4x4 is performing better compared to 1x1 MIMO for error correction in data communication system

without LDPC corrections. The BER value of the 0.1e-1 is -7 dB, -1dB, 7dB and 7dB respectively, for 4x4, 3x3, 2x2, and 1x1 MIMO. This implies that 4x4 is performing better with LDPC compared to without LDPC.

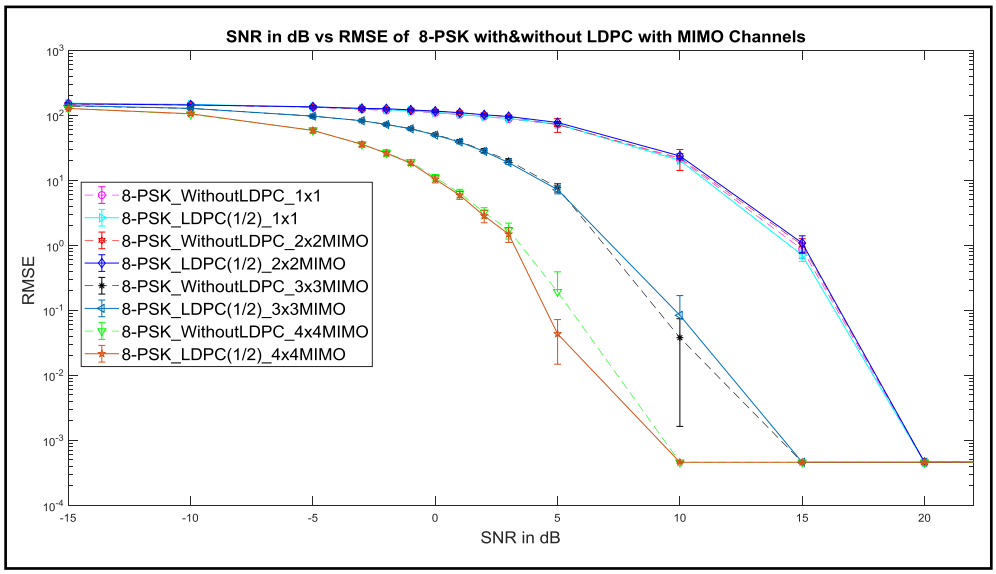


Figure. 4.12 RMSE performance of LDPC code with code rate R = 1/2 over an AWGN channel via 8-PSK modulation.

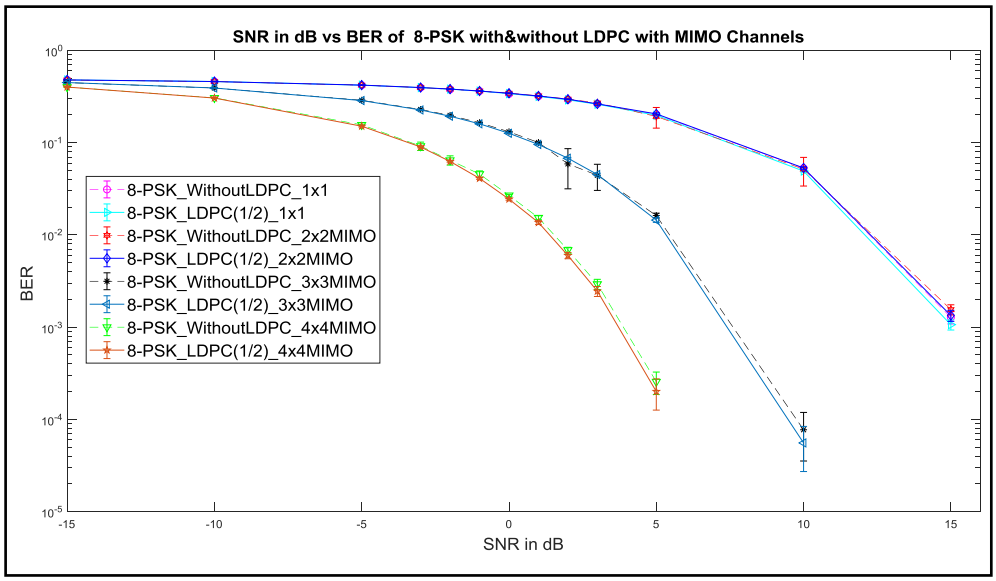


Figure. 4.13 BER performance LDPC code with code rate R = 1/2 over an AWGN channel via 8-PSK modulation.

Scenario IV:

The RMSE performance is shown in Figure 4.14 here, the performance is for 16-PSK modulation with and without LDPC error correction over the AWGN channel having SNR -15dB to 40dB. It is observed that RMSE is better with the LDPC algorithm as compared in data communication system without LDPC error corrections for 1x1, 2x2, 3x3 and 4x4 MIMO channels. The RMSE is 0.4e-4 for 1x1 LDPC MIMO at 25 dB SNR value and the RMSE is 0.4e-4 for 4x4 LDPC MIMO at 15 dB SNR value. This states that one can have error-free signal transmission up to 15dB with LDPC error correction coding.

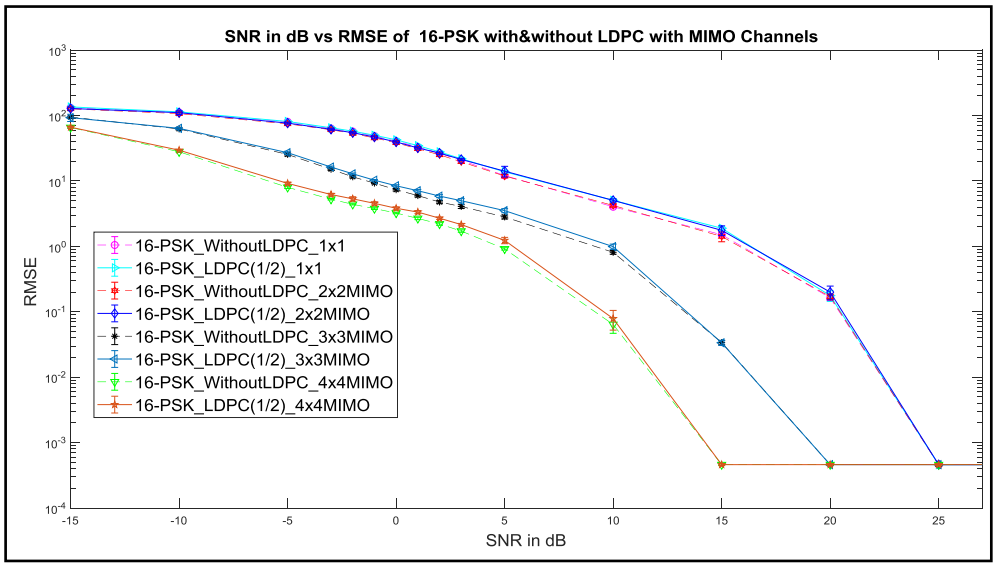


Figure. 4.14 RMSE performance of LDPC code with code rate R = 1/2 over an AWGN channel via 16-PSK modulation.

The BER performance is also presented in Figure 4.15 while the performance is for 16-PSK modulation with and without LDPC error correction over the AWGN channel having SNR -15dB to 40dB. The observation shows that the performance of BER is better with the LDPC algorithm as compared in data communication system without LDPC error corrections for 1x1, 2x2, 3x3, and 4x4 MIMO channels. The BER

value of the e-2 is 7 dB, 13dB, 18dB and 20dB respectively, for 4x4, 3x3, 2x2, and 1x1 MIMO in data communication system without LDPC. The BER value of the 0.1e-1 is -3 dB, 3dB, 13dB and 13dB respectively, for 4x4, 3x3, 2x2, and 1x1 MIMO in data communication system with LDPC. This implies that 4x4 is performing better with LDPC compared to without LDPC.

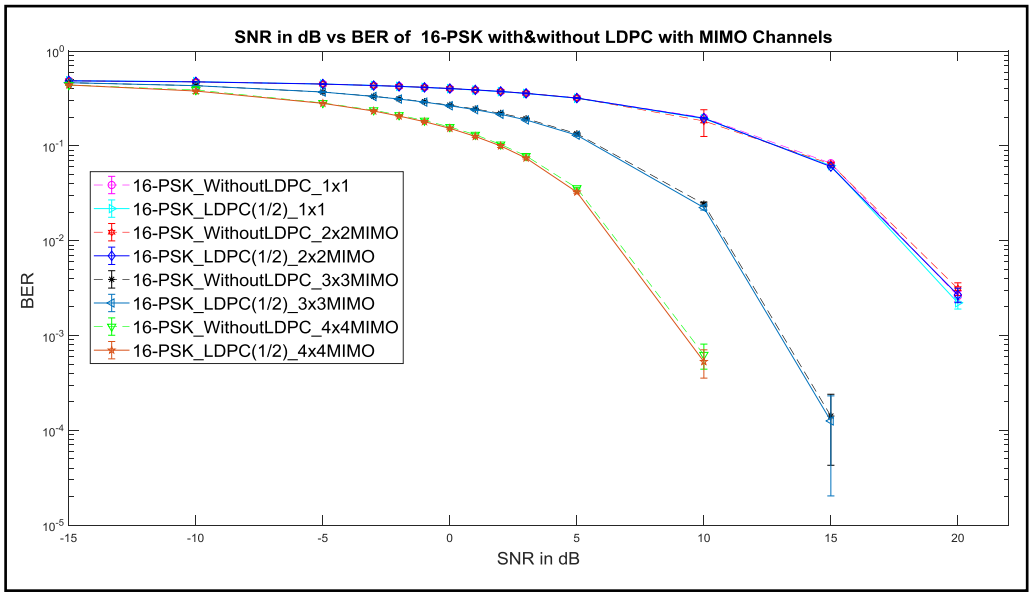


Figure. 4.15 BER performance LDPC code with code rate R = 1/2 over an AWGN channel via 16-PSK modulation.

Scenario V:

Figure 4.16 and 4.17, reveals the performance curves for RMSE and BER for various values of SNR and various configuration of MIMO over 32-PSK modulation. The RMSE performance (refer Figure 4.16) of data reduction with PCA over 32-PSK modulation with and without LDPC error correction over AWGN channel having SNR -15dB to 40dB. However, the performance of RMSE is better with LDPC channel coding as compared without LDPC error corrections for 1x1, 2x2, 3x3, and 4x4 MIMO channels. Further, the RMSE is $0.7X10^{-3}$ for 1x1 MIMO at +30 dB SNR, and similarly,

the value of RMSE is $0.7X10^{-3}$ for 4x4 MIMO at +20dB SNR in data communication system without LDPC channel coding. This indicated that the performance is better for the higher MIMO channel. So this indicates the error-free signals transmission up to +20dB without LDPC error correction coding for the proposed system. Similarly, the RMSE is $0.7X10^{-3}$ for 1x1 MIMO at +28dB SNR, and similarly, the value of RMSE is $0.7X10^{-3}$ for 4x4 MIMO at -18dB SNR with LDPC channel coding. This implied that the channel coding has better RMSE performance as compared to non-channel coding communication systems.

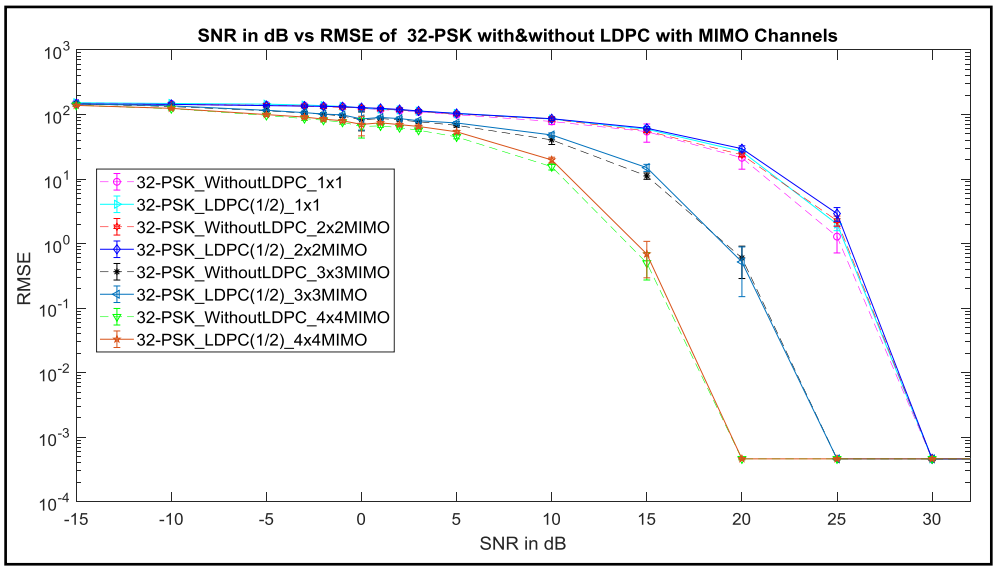


Figure. 4.16 RMSE performance of LDPC code with code rate set to R = 1/2 over the assigned AWGN channel using the technique of PCA-32PSK modulation.

The BER performance is shown in Figure 4.17. While the performance of data reduction methods using PCA over 32-PSK modulation, with and without LDPC error correction for the AWGN channel, is presented. The performance for the AWGN channel with varying SNR -15dB to 40dB is evaluated. The observations reveal that BER is better with LDPC channel coding as compared without LDPC error corrections. Further, the BER value is 0.1at +16dB,+16dB, +7dB and +5dB for 1x1, 2x2, 3x3, and

4x4 MIMO channel respectively without LDPC channel coding. This implies that the 4x4 MIMO channel is performing better compared to the 1x1 MIMO channel for error correction. Similarly the BER value is 0.6 at +5dB,+5dB, -5dB and -10dB for 1x1, 2x2, 3x3, and 4x4 MIMO channel respectively with LDPC channel coding . This implies that 4x4 MIMO channel is performing better compared to the 1x1 MIMO channel for error correction with channel coding communication systems. Therefore the system outperforms the distortion significantly when the antennas number in the channel is increased. Hence, the gain obtained with MIMO 4x4 provides the possibility to work with greater capacity in the system without a trade-off of interference and distortion. For this reason, it is very important to choose the correct modulation index. Therefore, proves that LDPC encoding works very well on a 4x4 MIMO even on the small value of SNR, wherein SNR has a larger noise value than the signal itself.

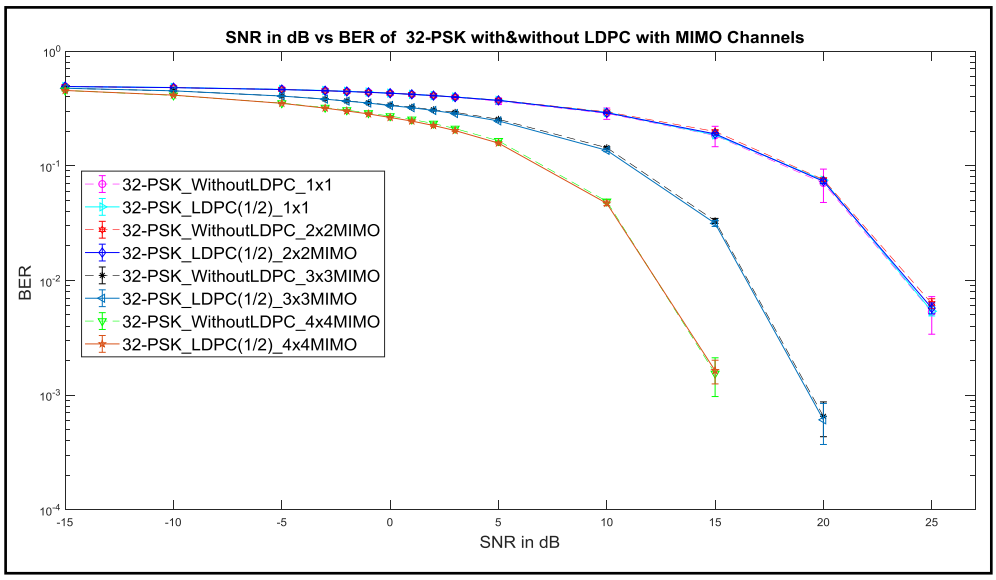


Figure. 4.17 BER performance LDPC code with code rate set to R = 1/2 over the assigned AWGN channel using the 32-PSK modulation technique.

Scenario VI:

The RMSE performance is shown in Figure 4.18. The performance is for 4-QAM modulation with and without LDPC error correction over the AWGN channel having SNR -15dB to 40dB. The performance of RMSE is better with the LDPC algorithm as compared without LDPC error corrections for 1x1, 2x2, 3x3, and 4x4 MIMO channels. The RMSE is 0.5e-4 for 1x1 LDPC MIMO at -01 dB SNR value. The RMSE is 0.5e-4 for 4x4 LDPC MIMO at -10 dB SNR value. This states that one can have error-free signals transmission up to -10dB with LDPC error correction coding.

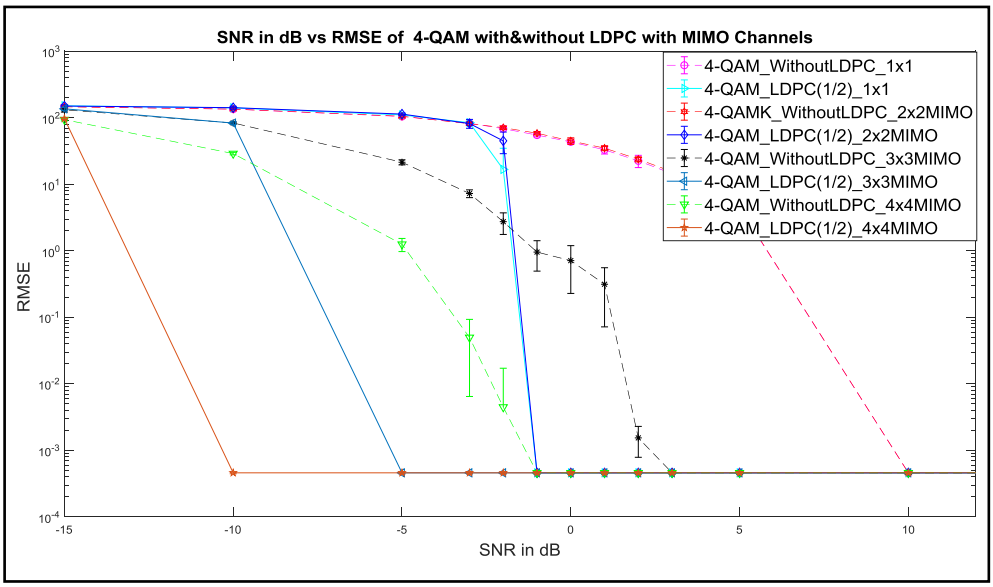


Figure. 4.18 RMSE performance of LDPC code with code rate R = 1/2 over an AWGN channel via 4-QAM modulation.

Further, the BER performance is presented in Figure 4.19. While the performance is for 4-QAM modulation with and without LDPC error correction over the AWGN channel having SNR -15dB to 40dB. One can see that the performance of BER is better with the LDPC algorithm as compared without LDPC error corrections for 1x1, 2x2, 3x3, and 4x4 MIMO channels. The BER value at the e-2 is -5.5 dB, -2dB, 4dB and 5dB respectively for 4x4, 3x3, 2x2, and 1x1 MIMO in data communication system without

LDPC channel coding. Therefore, implies that 4x4 is performing better compared to 1x1 MIMO for error correction. The BER value at the 0.1e-1 is -15 dB, -9dB, -2dB and -1dB respectively for 4x4, 3x3, 2x2, and 1x1 MIMO in data communication system with LDPC channel coding. This implies that 4x4 is performing better with LDPC compared to without LDPC.

As the model starts simulating, for every received data at the LDPC decoder, bit error rate, channel SNR and packet error rate is continuously updated. Hence the increase of SNR is required to get an error-free output signal and less noisy plot.

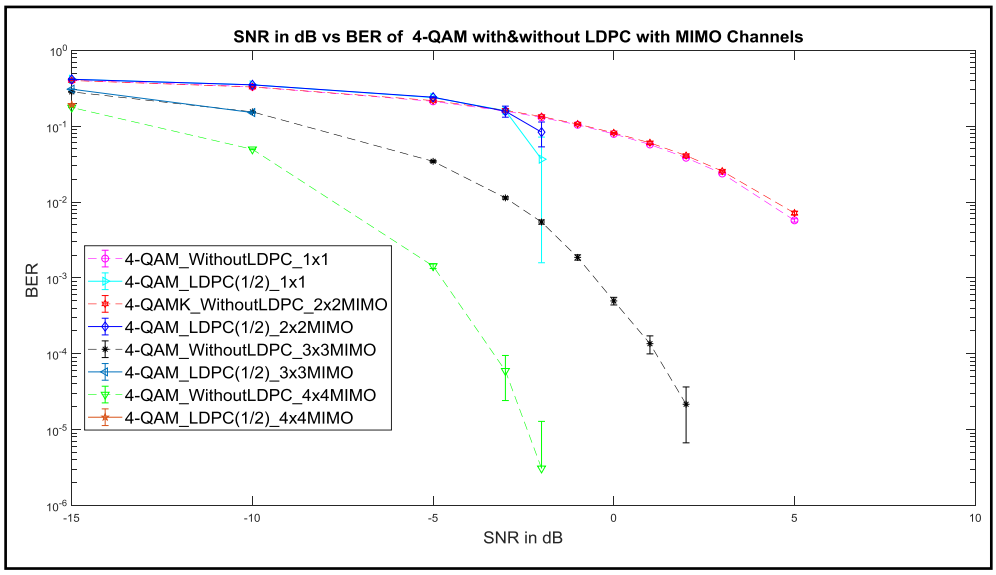


Figure. 4.19 BER performance LDPC code with code rate R = 1/2 over an AWGN channel via 4-QAM modulation.

Scenario VII:

The RMSE performance is shown in Figure 4.20. While the performance is for 8-QAM modulation with and without LDPC error correction over the AWGN channel having SNR -15dB to 40dB. The performance of RMSE is better with the LDPC algorithm as compared without LDPC error corrections for 1x1, 2x2, 3x3, and 4x4 MIMO channels. The RMSE is 0.4e-4 for 1x1 LDPC MIMO at -03 dB SNR value. The

RMSE is 0.4e-4 for 4x4 LDPC MIMO at -10 dB SNR value in data communication system with LDPC coding. This states that one can have error-free signals transmission up to -10dB with LDPC error correction coding.

The BER performance is also presented in Figure 4.21 and the performance is for 8-QAM modulation with and without LDPC error correction over the AWGN channel having SNR -15dB to 40dB. One can see that the performance of BER is better with the LDPC algorithm as compared without LDPC error corrections for 1x1, 2x2, 3x3, and 4x4 MIMO channels. The BER value at the e-2 is -7.5 dB, -3dB, 4dB and 5dB respectively for 4x4, 3x3, 2x2, and 1x1 MIMO in data communication system without LDPC Coding. This implies that 4x4 is performing better compared to 1x1 MIMO for error correction. The BER value at the 0.1e-1 is -15 dB, -10dB, -3dB and -2dB respectively for 4x4, 3x3, 2x2, and 1x1 MIMO. This implies that 4x4 is performing better with LDPC compared to without LDPC.

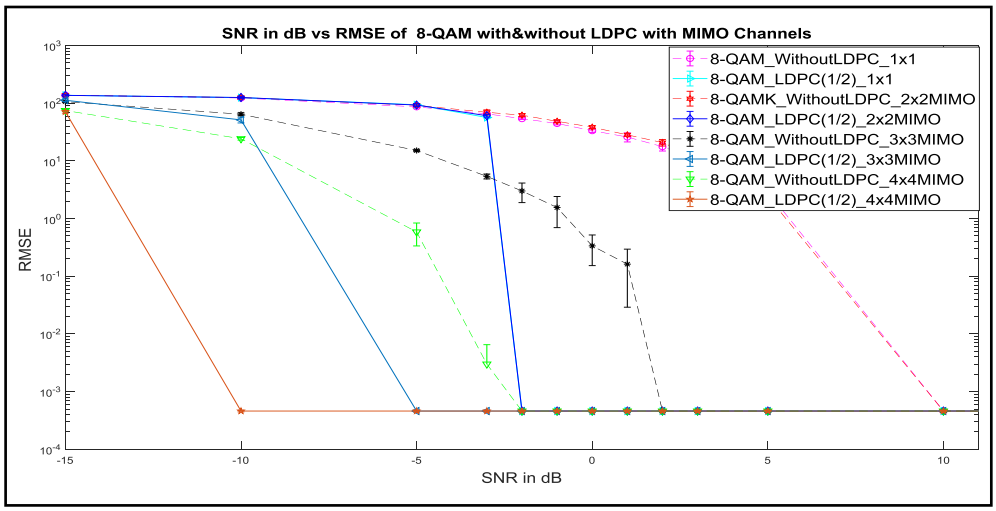


Figure. 4.20 RMSE performance of LDPC code with code rate R = 1/2 over an AWGN channel via 8-QAM modulation.

The above scenarios says PCA with QAM modulation with LDPC 4x4 MIMO shows better results compare to the BPSK modulation. According to this study says

higher-level QAM modulation with higher configuration MIMO using LDPC better results gives for the communication system.

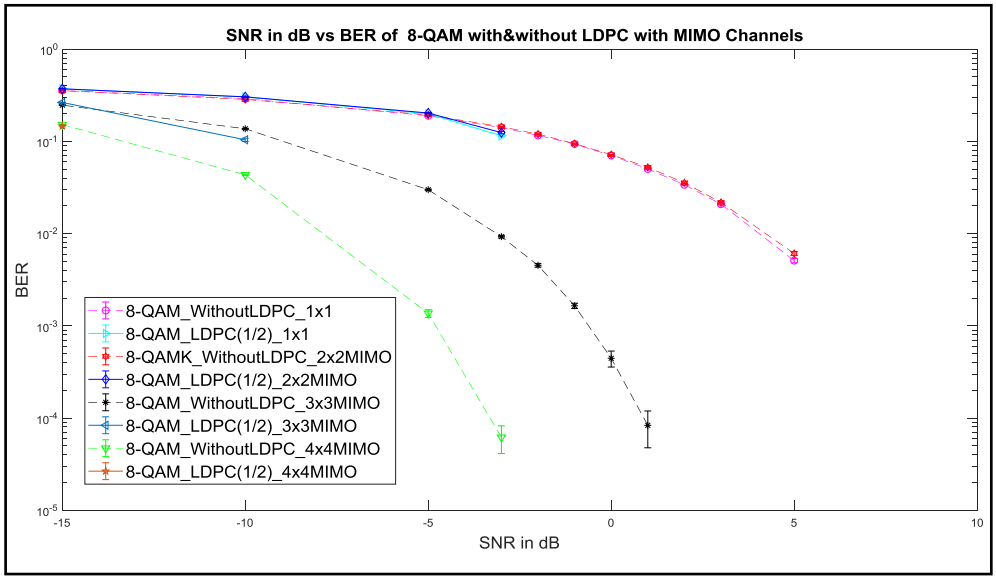


Figure. 4.21 BER performance LDPC code with code rate R = 1/2 over an AWGN channel via 8-QAM modulation.

Scenario VIII:

In Figure 4.22 and 4.23, we have presented the performance curves for RMSE and BER for various values of SNR and various configuration of MIMO over 16-QAM modulation. This Figure shows that 4x4 MIMO with LDPC good performance with efficiency up to -10 dB compare with 1x1 MIMO.

The RMSE performance (refer Figure 4.22) of data reveals the reduction with PCA over 16-QAM modulation with and without LDPC error correction over AWGN channel having SNR -15dB to 40dB. The performance of RMSE is better with LDPC channel coding as compared without LDPC error corrections for 1x1, 2x2, 3x3, and 4x4 MIMO channels. The RMSE is 1.0 for 1x1 MIMO at +5 dB SNR, and similarly, the value of RMSE is 1.0 for 4x4 MIMO at -5dB SNR without LDPC channel coding. This indicated that the performance is better for the higher MIMO channel. So this indicates

the error-free signals transmission up to -5dB without LDPC error correction coding for the proposed system. Similarly, the RMSE is 1.0 for 1x1 MIMO at +2.5dB SNR, and similarly, the value of RMSE is 1.0 for 4x4 MIMO at -13.5dB SNR with LDPC channel coding. Hence, the channel coding has better RMSE performance as compared to a non-channel coding communication system.

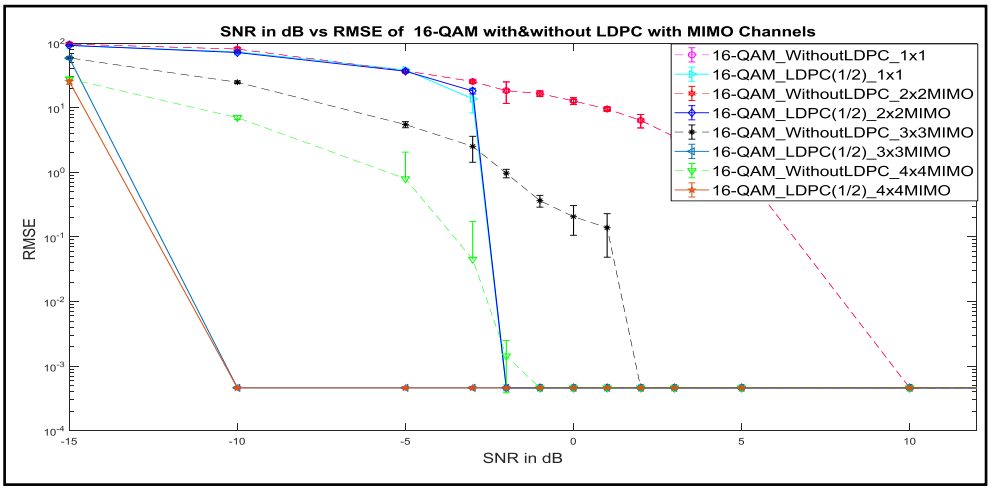


Figure. 4.22 RMSE performance of LDPC code with code rate set to R = 1/2 over the assigned AWGN channel using the technique of PCA-16QAM modulation.

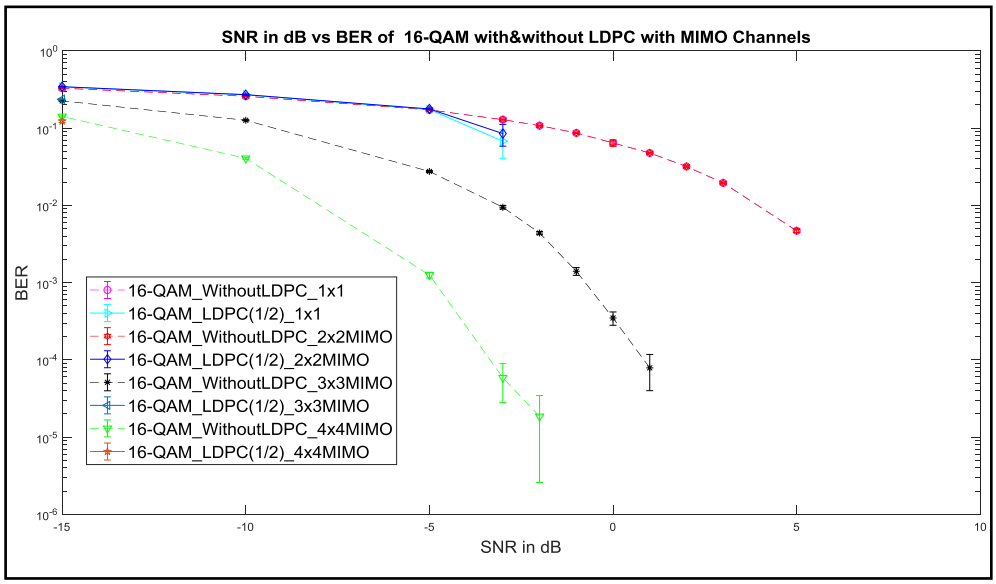


Figure. 4.23 BER performance LDPC code with code rate set to R = 1/2 over the assigned AWGN channel using the 16-QAM modulation technique.

The BER performance is also presented in Figurer 4.23 and the performance of data reduction methods using PCA over 16-QAM modulation, with and without LDPC error correction for the AWGN channel, is presented. However, the performance for the AWGN channel with varying SNR -15dB to 40dB is evaluated. The BER is better with LDPC channel coding as compared without LDPC error corrections. The BER value is 0.01at +4dB,+4dB, -3dB and -8dB for 1x1, 2x2, 3x3, and 4x4 MIMO channel respectively in data communication system, without LDPC channel coding. This implies that the 4x4 MIMO channel is performing better compared to the 1x1 MIMO channel for error correction. Similarly the BER value is 0.2 at -5dB,-5dB, -15dB and -15dB for 1x1, 2x2, 3x3, and 4x4 MIMO channel respectively in data communication system with LDPC channel coding. This implies that the 4x4 MIMO channel is performing better compared to the 1x1 MIMO channel for error correction with channel coding communication systems. Therefore the system outperforms the distortion significantly when the antennas number in the channel is increased. The gain obtained with MIMO 4x4 provides the possibility to work with greater capacity in the system without a trade-off of interference and distortion. Hence it is important to choose the correct modulation index. This proves that LDPC encoding to work well on a 4x4 MIMO even on the small value of SNR, wherein SNR has a larger noise value than the signal itself.

Scenario IX:

In Figure 4.24 and 4.25, reveals the performance curves for RMSE and BER for various values of SNR and various configuration of MIMO over 32-QAM modulation. The Figure shows that 4x4 MIMO with LDPC has a good performance with efficiency up to -15 dB compare with 1x1 MIMO.

The RMSE performance (refer Figure 4.24) of data reduction with PCA over 32-QAM modulation with and without LDPC error correction over AWGN channel having

SNR -15dB to 40dB. It is observed that RMSE is better with LDPC channel coding as compared without LDPC error corrections for 1x1, 2x2, 3x3, and 4x4 MIMO channels. The RMSE is 1.0 for 1x1 MIMO at 5dB SNR, and similarly, the value of RMSE is 1.0 for 4x4 MIMO at -12dB SNR without the LDPC channel coding. This indicates that the performance is better for the higher MIMO channel. Hence, provide error-free signals transmission up to -10dB without LDPC error correction coding for the proposed system. Similarly, the RMSE is 1.0 for 1x1 MIMO at -3dB SNR, and similarly, the value of RMSE is 1.0 for 4x4 MIMO at -15dB SNR with LDPC channel coding. This implied that the channel coding has better RMSE performance as compared to a non-channel coding communication system.

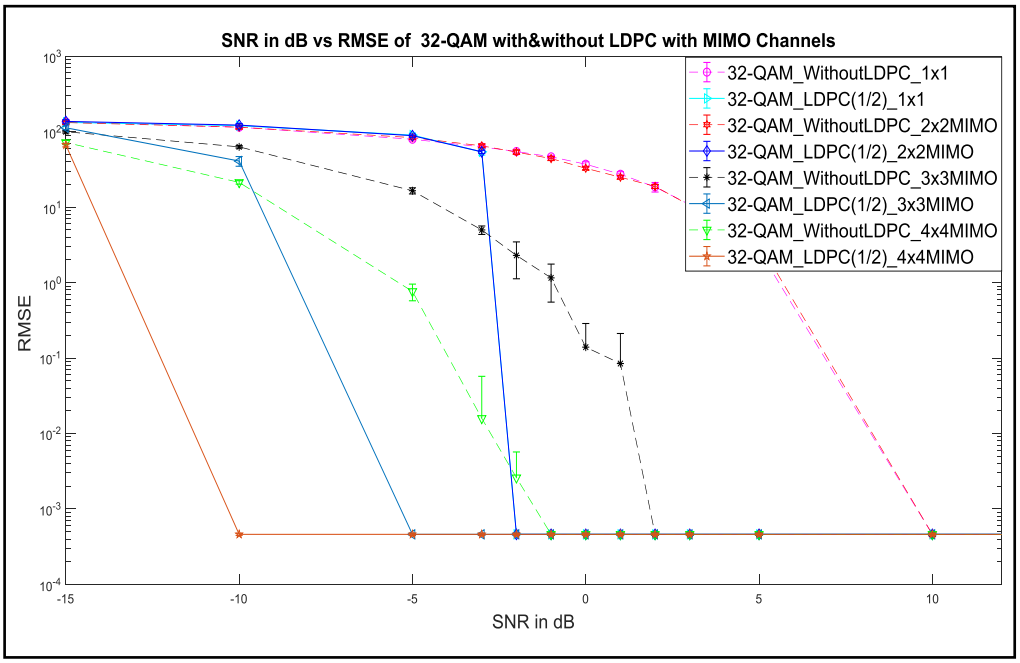


Figure. 4.24 RMSE performance of LDPC code with code rate set to R = 1/2 over the assigned AWGN channel using the technique of PCA-32QAM modulation.

The BER performance is also presented in Figurer 4.25 and the performance of data reduction methods using PCA over 32-QAM modulation, with and without LDPC error correction for the AWGN channel, is presented. Performance for the AWGN

channel with varying SNR -15dB to 40dB is evaluated. It is observed that BER is better with LDPC channel coding as compared without LDPC error corrections. The BER value is 0.01at +1.5dB,+1dB, -4dB and -9dB for 1x1, 2x2, 3x3, and 4x4 MIMO channel respectively in data communication system without LDPC channel coding. This implies that the 4x4 MIMO channel is performing better compared to the 1x1 MIMO channel for error correction. Similarly the BER value is 0.2 at -7dB,-7dB, -10dB and -15dB for 1x1, 2x2, 3x3, and 4x4 MIMO channel respectively in data communication system with LDPC channel coding. Hence, the 4x4 MIMO channel is performing better compared to the 1x1 MIMO channel for error correction with channel coding communication systems. Therefore the system outperforms the distortion significantly when the antennas number in the channel is increased.

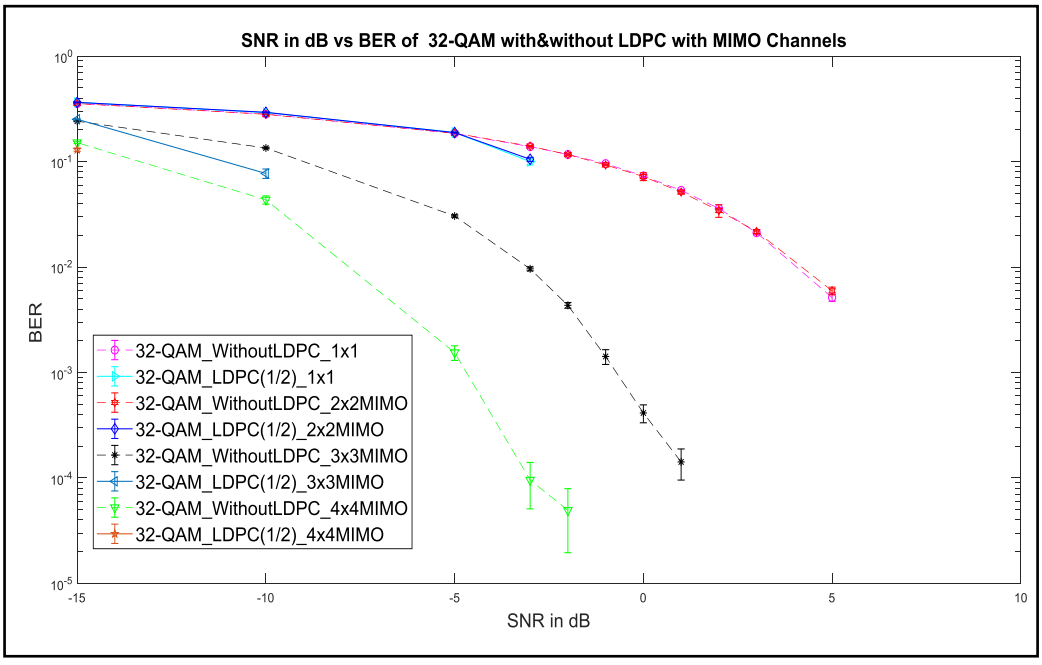


Figure. 4.25 BER performance LDPC code with code rate set to R = 1/2 over the assigned AWGN channel using the 32-QAM modulation technique.

The gain obtained with MIMO 4x4 provides the possibility to work with greater capacity in the system without a trade-off of interference and distortion. Therefore it is

important to choose the correct modulation index. This proves that LDPC encoding works very well on a 4x4 MIMO even on the small value of SNR, wherein SNR has a larger noise value than the signal itself.

4.3. Synthesis of source coding data communication model for BPSK modulation on Xilinx FPGA

Field Programmable Gate Array (FPGA) technology enables reconfigurable device systems has become a feasible for applying data processing algorithms for applications. This section addresses the FPGA synthesis and implementation of PCA data reduction using BPSK modulation. The model-based architecture provides the ability to conduct quick FPGA based prototyping of data reduction communication. The model design is developed using Xilinx ISE 13.1 IDE with Xilinx System Generator (XSG) configured with MATLAB R2013a which combines the Xilinx Block set and Mat lab Simulink environment. The design is synthesized for the Virtex-6 FPGA. The MATLAB framework is a high-level technical language for the creation of algorithms, data processing, data interpretation, and numerical computation. System generator automatically defines the FPGA configuration with the aid of Simulink Xilinx DSP generator, generating the bit file which can be programmed on FPGA.

The source coding PCA technique is integrated using BPSK with MATLAB Simulink and Xilinx system generator (XSG) as shown in Figure 4.26. Using Simulink library from the communication system, composite data for the ensemble is generated in a matrix form of the block size (100, 10) for 16-bitword length corresponding to 10 sensors signal data having 100 time-stamped data samples. The demonstration is made to ensemble 10 sensor signal set within the frequencies of range between 10-100Hz. Further, the signal data is reduced to (10, 10) using PCA data transformation, which is a subsequent reduction in the form of features matrics or eigenmatrix having 16-bitword

length. However, the compressed features matrix data is later modulated using BPSK modulator. While the channel selected here is AWGN with SNR values ranging from 0 dB to 9dB. Later the data is demodulated and regressed for original. The system is further synthesized and implemented on the Virtex-6 FPGA (refer Table 4.3) untether communication.

Table. 4.3 Xilinx synthesis model parameters used in the data communication system.

Parameter	**Value**
Modulator Method	BPSK
Modulator Input Type	Bit
Modulator Symbol Order	Binary
Demodulator Output Type	Bit
Channels	AWGN
Channel Noise Factor	SNR

Implementation parameters of the data reduction with BPSK modulation is shown in Table 4.3. The flow steps for the design are defined as follows. In the first step, MATLAB Simulink is used to implement reduction or compression algorithm with the PCA method and later converted using the Xilinx block set library. Xilinx models are given the input data in the form of a vector in fixed point as per Xilinx block set format. This model is simulated with the proper simulation time in MATLAB Simulink settings. The system generator has been configured for the FPGA board virtex-6.

The Xilinx device generator itself has the ability to create user limit register, test bench and check vectors for design validation. Compilation of the bit stream is performed to construct an FPGA bit file that is suitable for implementation on target device Virtex-6 LX240T device. The hardware realization of the proposed PCA based data reduction using BPSK based on FPGA technology provides a fast, compact and low power solution for the communication system. Figure 4.27 shows the Resource Utilization in FPGA.

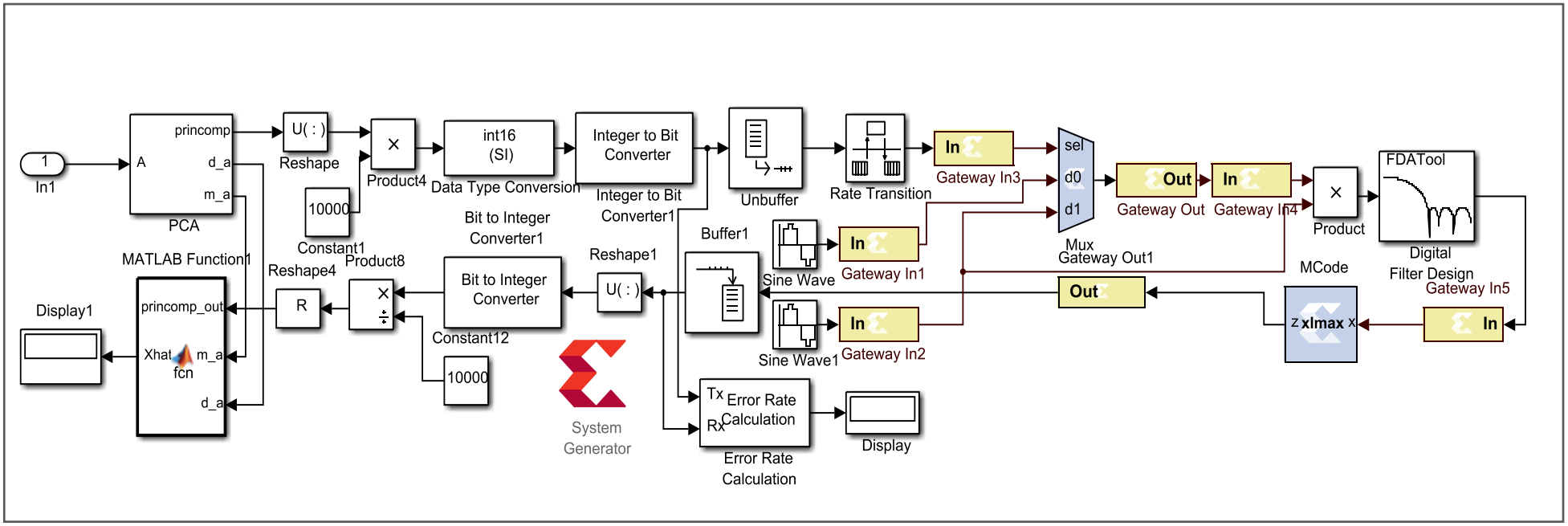


Figure. 4.26 Xilinx based data sourcing model synthesis using BPSK modulation.

The representation of the signal data is converted into binary format, i.e., ‘0’ and ‘1’, in which the present model of PCA with the BPSK system model synthesis process, sometimes requires synchronization issues. This gives rise to in Xilinx synthesis, shift in the original data word into one or two bits forward when recovered at the receiving end over buffer, hence changing the original recovered data entirely. In the present work, 16-bit data transmission in the communication model is observed. Hence due to this bit shift error, the study shows high RMSE value. Further, the 16-bits word length, single and double bits shift is shown in Table 4.4.

bpsk_cw Project Status			
Project File:	bpsk_cw.xise	**Parser Errors:**	No Errors
Module Name:	bpsk_cw	**Implementation State:**	Synthesized
Target Device:	xc6vlx240t-3ff1759	**•Errors:**	No Errors
Product Version:	ISE 14.7	**•Warnings:**	5414 Warnings (16 new)
Design Goal:	Balanced	**•Routing Results:**	
Design Strategy:	Xilinx Default (unlocked)	**•Timing Constraints:**	
Environment:	System Settings	**•Final Timing Score:**	

Device Utilization Summary (estimated values)			[-]
Logic Utilization	**Used**	**Available**	**Utilization**
Number of Slice Registers	9708	301440	3%
Number of Slice LUTs	5714	150720	3%
Number of fully used LUT-FF pairs	3904	11518	33%
Number of bonded IOBs	540	720	75%
Number of BUFG/BUFGCTRLs	1	32	3%
Number of DSP48E1s	4	768	0%

Detailed Reports					[-]
Report Name	**Status**	**Generated**	**Errors**	**Warnings**	**Infos**
Synthesis Report	Current	Mon Oct 12 10:52:01 2020	0	5414 Warnings (16 new)	1240 Infos (238 new)
Translation Report					
Map Report					
Place and Route Report					
Power Report					
Post-PAR Static Timing Report					
Bitgen Report					

Secondary Reports	[+]

Date Generated: 10/12/2020 - 10:52:17

Figure. 4.27 The Resource Utilization in FPGA.

Table 4.4 One and two-bit right shit for 16-bit word length over Xilinx system generator for BPSK

The bit stream representation in buffer consist of 16-bit word length generated in order	**The bit stream representation in received buffer with Single Bit shift**	**The bit stream representation in received buffer with Double Bit Shift**
0000 0000 0000 0000	0000 0000 0000 0000	0000 0000 0000 0000
0000 0000 0000 0001	0000 0000 0000 0000	0000 0000 0000 0000
0000 0000 0000 0010	1000 0000 0000 0001	0100 0000 0000 0000
0000 0000 0000 0011	0000 0000 0000 0001	1000 0000 0000 0000
0000 0000 0000 0100	1000 0000 0000 0010	1100 0000 0000 0001
0000 0000 0000 0101	0000 0000 0000 0010	0000 0000 0000 0001
0000 0000 0000 0110	1000 0000 0000 0011	0100 0000 0000 0001
0000 0000 0000 0111	0000 0000 0000 0011	1000 0000 0000 0001
0000 0000 0000 1000	1000 0000 0000 0100	1100 0000 0000 0010
0000 0000 0000 1001	0000 0000 0000 0100	0000 0000 0000 0010
0000 0000 0000 1010	1000 0000 0000 0101	0100 0000 0000 0010
0000 0000 0000 1011	0000 0000 0000 0101	1000 0000 0000 0010
0000 0000 0000 1100	1000 0000 0000 0110	1100 0000 0000 0011
0000 0000 0000 1101	0000 0000 0000 0110	0000 0000 0000 0011
0000 0000 0000 1110	1000 0000 0000 0111	0100 0000 0000 0011
0000 0000 0000 1111	0000 0000 0000 0111	1000 0000 0000 0011

To provide a comparison of the detailed evaluation results for each one of the ensembles, a tenfold experiment is computer to generate the RMSE and BER over SNR values with AWGN channel as given in Figure 4.28 & Figure 4.29. The demonstration of RMSE as shown in Figure 4.28, where extreme of each box is 25^{th} and 75^{th} percentiles and their whiskers stretch to the most extreme values found, minus those of outliers identified by "+" labels. As the lowest RMSE median values is at -6dB and -3dB the highest data dispersion respectively. The RMSE ‘minimum’ and the

‘maximum’ observed values are at 0dB and -6dB value respectively. The consistency is not seen in the RMSE due to the single bit and double bit shift error as discussed above, while the RMSE is acceptable range.

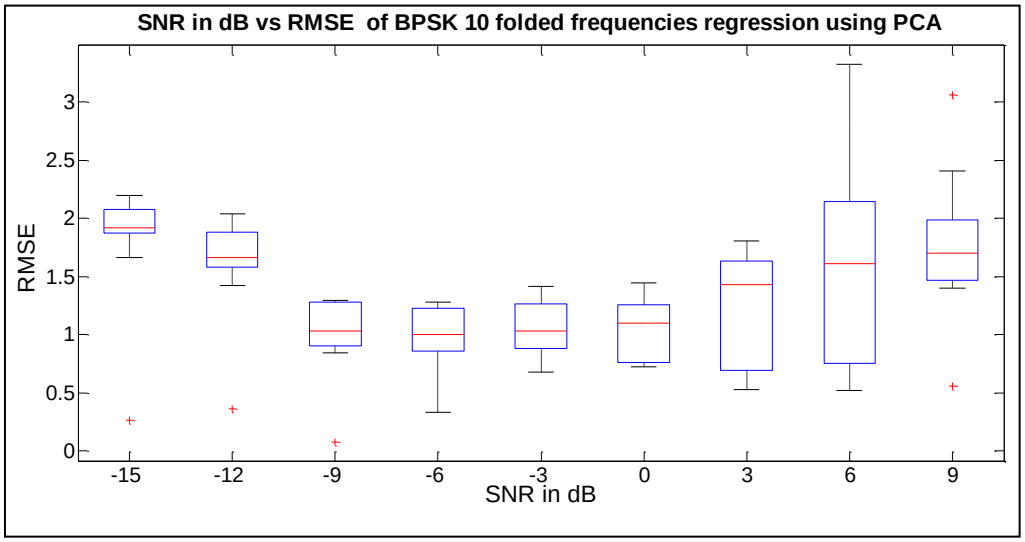


Figure. 4.28 RMSE performance over the AWGN channel synthesis on FPGA for BPSK modulation technique for 10 folds experiment.

As seen in Figure 4.29, the lowest BER median values were presented at 9dB representing lowest distortion. The generated plot has consistency and is in line with the theory.

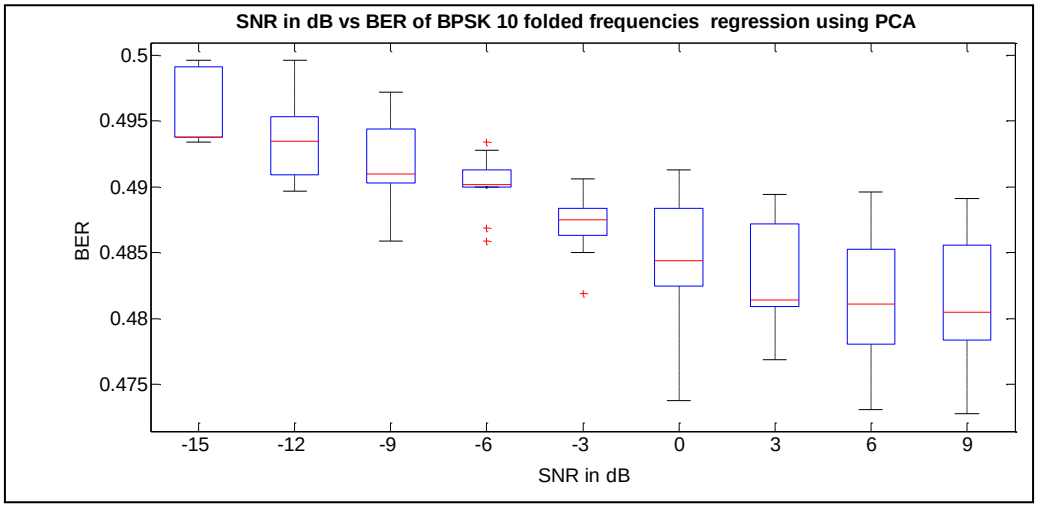


Figure. 4.29 BER performance over the AWGN channel synthesized on FPGA for BPSK modulation technique for 10 folds experiments.

4.4. Discussion

Vinaya et al. [8] specified that the results of BER performance improve with an increase in block length, which can be optimized for error correction by setting the number of decoding iterations 5 to 50 for SNR close to 2dB. Since more errors are introduced for larger frame lengths, number of iterations needs to be increased for better Error Correction performance.

Similar works are found like that of Raghuwansi et al. in which the BER varying with the changes in modulation techniques as well as FFT points, which says that with 256 FFT points BPSK modulation the wireless LDPC based MIMO-OFDM system outperform [9]. Also, Indira Bestari et al. presented the LDPC encoding process, LDPC encoding works very well on a single RF-based MIMO OFDM system, even on the negative value of SNR dB. However, the coding rate also needs to be noted, which states that the lowest rate will require a larger parity matrix to produce a higher level of accuracy in the error correction process. The process of channel estimation also works very well by generating a minimal difference between theoretical and estimation values [10].

Furthermore, Diouf M et al. investigated the complexity and BER performance analysis of Maximum likelihood zero-forcing and minimum mean square error MIMO receivers by using a small length binary polar code. At the same time, the performance gain of 3dB at $10e^{-2}$ of the proposed is obtained compared to the optimal maximum likelihood detector. Zero Force is always the worst performance, as in the literature [11]. The author, Lamia et al. demonstrate that the channel codes provide high channel efficiency with small code rate, low decoding complexity, and low BER, describing the concatenated BCH-CC-LDPC for the MIMO system. Also, the concatenated scheme improves the performance of a low receive diversity MIMO system. Therefore, the concatenated schema enhances the system decoding latency because an LDPC decoder

is replaced by a CC decoder and a short BCH decoder, which has lower decoding latencies than LDPC.

Further, the proposed system could also be enhanced in the case of LDPC convolutional codes [12]. The author, A.Z.M. Touhidul Islam et al. presented the BER performance of the 1/2-rated LDPC channel encoded STTC MIMO-OFDM system under different digital modulations on a Raleigh fading channel, which degrades due to fading channel effect. The system outperforms at BPSK modulation and shows the poor performance at QPSK modulation, while the Doppler frequency shifts more influences the system with QPSK modulation,and its performance degrades. At the SNR value of 0.75 dB, the system performance is improved by 2.7195 dB in the case of BPSK modulation as compared with QPSK [13].

Moreover, Nordam R et al. proposed the soft decision decoding and signal combining algorithms of MRC and the Alamouti code to the orthogonal MIMO codes for three and four T_x antennas. The complexities of the novel combiners do not substantially exceed that of the Alamouti combiner, making them very attractive in terms of implementation. With the soft decision MIMO combiners, the FEC decoders attain the same error rates with 2 to 3 dB smaller data bit energy than with the hard decision MIMO. The novel combiners require a close to constant channel over four instants, implying that they are better suited for low mobility environments. However, the soft decision MIMO combiners provide the largest gain when linked to a bitwise soft decision demodulator. Extracting the bitwise LLR values from the combined output symbols is shown to be a very simple task not only in QPSK but also in Gray-mapped 16-QAM. Therefore, it is notable, that in order to get correct noise variance to the bitwise LLR equations, multiplication by the pertaining channel tap energies has to be carried out [14].

Pramanik A et al. show low-density parity-check codes and space-time block codes on far-end data reconstruction offer significant visual quality improvement at reduced data transmission. Further, the proposed soft-decision decoding scheme offers a coding gain of 3 dB at −2 dB channel signal-to-noise ratio for an 8 antenna system. However, the simulations show a coding gain increases with the increase in signal-to-noise ratio and the number of transmit antennas. Decoding performances of the block codes with zero and no-zero entries highlight the fact that the former offers marginally better bit error rate performance (around 0.5 dB for number of antennas less than 8) compared to the latter one. However, zero-entry codes give rise to a peak-to-average power ratio problem [15].

Further, Ibrahim et al. revealed the BER performance of LDPC codes in Weibull fading channels, were analyzed for different decoding rules using comparative computer simulations. However, an LDPC coded BPSK communication system was designed as a communication infrastructure to perform the analysis. Four different decoding rules and a regular LDPC code with (1008, 504) block length and ½ code rate were used in simulations [16].

The Min-Sum algorithm based LDPC decoder is presented in the present study. Among various message passing algorithms, one of the vital algorithms is Min-Sum, which provides suitability for implementation of the decoder with less hardware implementation area. The data reduction methodology for enhancing the channel capacity of the MIMO channel has helped in designing a high-performance communication system, especially in the low bandwidth wireless networks. For future work, the study addresses, different modulation techniques and their performance in terms of BER and RMSE efficiency versus SNR of received data.

The present investigation has successfully developed PCA based data reduction using BPSK modulation for data communication application with hardware-software

co-simulation implementation on FPGA. Based on quality metrics such as RMSE and BER, it is found that the proposed method is a better option for enhancing data is the reconstruction process in the communication system. The proper utilization of Simulink/ Xilinx system generator DSP blocks for FPGA greatly shortens the development cycle from a software algorithm to hardware. It leads to fast time to market of the design.

4.5. References

[1] Andrews, J.G., Choi, W. and Heath Jr, R.W. MIMO in 3G cellular systems: Challenges and future directions. *Submitted to IEEE Communications Magazine*, 2005.

[2] FPGA Implementation of MIMO Module, at www.rfwireless-world.com.

[3] Siavash M. Alamouti. A Simple Transmit Diversity Technique for Wireless Communications. *IEEE Journal on Select Areas in Communcations*, 16(8):1451–1458, 1998.

[4] Vahid Tarokh, Hamid Jafarkhani, and A. Robert Calder bank. Space-time block coding for wireless communications: Performance results. *IEEE Journal on Selected Areas in Communications*, 17(3):451–460, 1999.

[5] Herrero, A.F., Jimenez-Pacheco, A., Caffarena, G. and Quiros, J.C. Design and implementation of a hardware module for equalisation in a 4G MIMO receiver. In *2006 International Conference on Field Programmable Logic and Applications*, Pages 1-4, 2006.

[6] Berger, L.T., Schwager, A., Pagani, P. and Schneider, D. eds. *MIMO power line communications: narrow and broadband standards, EMC, and advanced processing*. CRC Press, 2014.

[7] Johnson SJ. *Iterative Error Correction: Turbo, Low-Density Parity-Check and Repeat–Accumulate Codes*. Cambridge Ubiversity Press, Pages 64-69, 2010.

[8] Vinaya R. Gad, Udaysingh V. Rane, Rajendra S. Gad and Gourish M. Naik. Performance Evaluation of Low Density Parity Check (LDPC) Codes over Gigabit Ethernet Protocol. *Transactions on Networks and Communications, Society for Science and Education United Kingdom*, 4(5):18-24, 2016.

[9] Raghuwansi, M. and Sharma, A. Designing of efficient LDPC based MIMO-OFDM using 4-PSK scheme. In *2015 International Conference on Computer, Communication and Control (IC4)*, Pages 1-4, 2015.

[10] Bestari, I. and Astawa, I.G.P., 2017, September. MIMO-OFDM 2× 2 based on single RF using LDPC code and Kalman filter detection. In *2017 International Electronics Symposium on Engineering Technology and Applications (IES-ETA)*, Pages 69-74, 2017.

[11] Diouf, M., Diop, I., Dioum, I., Farssi, S.M., Diouf, B. and Tall, K., 2016. Soft output detection for MIMO systems using binary polar codes. In *2016 5th International Conference on Multimedia Computing and Systems (ICMCS)*, Pages 400-404, 2016.

[12] Berriche, L. and Al Qahtani, A. Low Error Floor Concatenated LDPC for MIMO Systems. *SYSTEM*, 8(4):571-574, 2017.

[13] Islam, A.T., Islam, M.S., Alam, M.M. and Ullah, S.E., 2011. Performance of iterative LDPC-based space-time trellis coded MIMO-OFDM system over AWGN and fading channels. *International Journal of Computer Science, Engineering and Information Technology*, 1(4):1-13, 2011.

[14] Nordman, R. Soft decision decoding of the orthogonal complex MIMO codes for three and four transmit antennas. *Physical Communication*, *5*(1):33-46, 2012.

[15] Pramanik, A., Maity, S.P. and Sarkar, S. Compressed sensing image reconstruction by low density parity check codes and soft decoding of space time block codes. *Computers & Electrical Engineering*, *72*:553-565, 2018.

[16] Ibrahim Develi, Yasin KabalciA. Comparative simulation study on the performance of LDPC coded communication systems over Weibull fading channels. *Journal of Applied Research and Technology*, 14(2):101-107, 2016.

Chapter 5

Data Communication

application in NoC

5.1. Introduction

The data transfer and processing via the router system is feasible in wide area networks. The router is used to transfer the packets from two or more separate topologies of the network. While Microprocessors or digital signal processors are built into a router system that handles data transmission and reception control functions. Whereas, the NoC On-Chip communication should happen through packet switching. The Routing Algorithm decides the mode of the packet be transmitted from source to destination. Therefore these routers are used between nodes in the network, where this routers can guide the packets based on the routing algorithm implemented in the NoC architecture. However, the developed NoC routing algorithm results in machine efficiency. Also, the distribution of latency, efficiency and load are very critical parameters that should be regarded in architecture.

As the number of processors on a single chip and the computational difficulty grows day by day, the processor-to-processor interconnection and communication mechanism become critical factors influencing chip-multiprocessor efficiency. Therefore to improve the performance, it requires more efficient interconnection and communication between processors, rather than depending on the processing speed. Hence, the system needs a full range of communication requirements and characteristics of all types of processors, which can provide improved efficiency in data transmission under restricted conditions such as chip location, power usage, network bandwidth etc. Thus, higher on-chip communication demands such as high speed, high bandwidth, and high throughput while limited area and low power consumption are needed.

Over the last decade, chip-based computing has developed while the system-on-chip (SoC) has been used for a long time, which reduces costs and creates further power usage. Therefore, it helps the researchers to discover Network-on-Chip (NoC), which is low in size and with decreased costs, has less power usage. However, as the complexity

of the design grows, the overall length of the interconnection wires decreases, resulting in lengthy transmission delays and increased power usage. Furthermore, the gap between the wires shrinks with the technology, increasing the capacitance of the connection and the height of the wire material resulting in greater capacitance of the fringe.

Besides, the wireless NoC (WNoC) architectures are based on omnidirectional antennas. However, in such WNoC systems, most situations follow a token-based medium access system for transferring data through the shared wireless channel. Since all antennas share a common channel, only a single communication is possible at any instant of time. The performance benefits of utilizing wireless links are thus limited. While having multiple non-overlapping wireless channels is one potential solution for achieving simultaneous communication. But building several transceivers tuned to non-overlapping channels, due to interference effects and heavy overheads, is an incredibly difficult task. Therefore, interference decreases the usable bandwidth and degrades the bit error rate (BER). Such cumulative impacts lead to low quality-of-service (QoS).

Network on chip (NoC) can be made more effective by building faster routers with larger buffers for communications, a greater number of ports and channels, and adaptive routing, both of which require substantial hardware costs overhead. These Networks on Chip (NoC) are suggested as a flexible framework for communication, which can also provide assured efficiency. NoCs typically focus on physical and spatial scalability. Sometimes, the practical scalability, i.e., the potential to satisfy increasingly challenging demands with a continuous output, cost ratio, is neglected.

The architecture model of the chip-multiprocessor (CMP) is expected in the future to replace the current single-core approach to embedded Nano scale circuits and systems [1]. A network-on-chip (NoC) interconnect fabric can connect the CMP cores. Further, the RF/wireless interconnections have the potential to provide multi-access

communication features, which are useful in a NoC system, particularly for cores located far away in the CMP environment. The use of multi-band RF interconnection [2] is the signal which propagates at light speed to provide shortcuts to a multi-core network-on-chip (mesh) topology. A recent work [3] proposed a method called wireless NoC (WiNoC) as the backbone of interconnection for the estimation, characterization, and simulation of interconnected networks. All the techniques employed in RF/wireless interconnect, however, are single-input single-output based systems. As RF/wireless channel is used for data transmission, calculating power holds significance in the overall efficiency of the network. Because the radio spectrum is small, the SISO device can not satisfy the RF/wireless interconnect channel capacity needs without significantly increasing the communication spectral efficiency.

The signal transmission over through-silicon via (TSV) connections inherently creates coupling interference on signal TSVs in 3D integrated circuits (ICs). These interventions might be a crucial factor in the design of 3D ICs and thus needs to be addressed. However, the system of multi-TSV channel equalization deployed in the digital world, which provides virtually perfect crosstalk compensation by inverting the effect of the individual coupling channels. Therefore, Crosstalk mitigation techniques are mainly based on traditional signal shielding, have been researched by using ground TSVs [4] or ground guard rings and coaxial TSVs [5] [6]. The loss of passive contacts is used to compensate for the capacitive effects of TSV interconnections and to partially equalize the response to the TSV frequency [7] [8].

Further, the approach to crosstalk compensation is provided by using optical signal processing to equalize the TSV voltage signals. This approach is influenced explicitly by the world of mobile communications, where the signals are spatially multiplexed, distributed over a time-dispersive multiple-input multiple-output (MIMO) radio channel, and eventually equalized at the receiver [9]. Whereas in comparison to

wireless communications, is considered as the multi-TSV channel as a MIMO channel and can be specifically applied to both equalization and channel estimation techniques. The present study also analyzed the robustness of this approach against analog-to-digital converter (ADC) impairments, namely ADC quantization noise and ADC clock jitter.

Network on Chip (NoC) is an up-coming world view that adapts to the expanding many-sided quality and correspondence requirement of future System on Chip (SoC). Whereas numerous topologies with various capacities have been proposed for NoCs where various topologies and parameters are chosen based on different NoC applications. However, the present works have been modeled to the Mesh topology for 4X4, 8X8, 16x16, 32x32, 64x64 nodes for varying packets size (0.1 and 16000Kbytes), queue size (5-200), link bandwidth (10-200Mbps), link propagation delay (10-200ms), over CBR (5 and 10Mbps) and FTP applications. Further, the performance of throughput and propagation delay of packets from the given source node to the destination node is studied for low (0.512Kbytes) and high load (64Kbytes) applications. Point by point, similar investigation of the reproduction brings about terms of latency and throughput is displayed. The outcomes can be utilized as a rule for NoC architects to settle on fitting decisions keeping in mind the end goal to accomplish ideal execution for respective applications of future wireless communications systems is to provide sensor data transmission high-data-rate, quality of service (QoS), low cost, speed of wireless access.

5.2. Methodology

The transmission rate R also is known as Channel Capacity C expressed as C=R=B*SE, here B is bandwidth (Hz) and SE=$\log_2$ (1+SNR) is spectral efficiency (bps/Hz) of modulation over the data link. The R is measured in terms of bits per second (bps) it can be improvised by increasing B at the expense of an area and

power cost. This relation is approximately linear up to a specific limit imposed by the technology. An alternative approach is using better modulation methods with higher SE, which usually non-linear in area and power due to complex digital modulator circuits. Further Channel capacity of the advance configuration of the communication systems can be enhanced by NR (number of antennas at the receiver) and NT (number of antennas at the transmitter) like SISO ($C= B*\log_2(1 + SNR)$), SIMO ($C = NR*B*\log_2(1 +SNR)$), MISO ($C = NT*B*\log_2(1 +SNR)$), MIMO ($C= \min(NT,NR)*B*\log_2(1 +SNR)$) having a transmitter power of P with an average signal to noise ratio of SNR. This concludes that there is a linear increase in capacity with the number of transmitting antennas, which can be low powered channels than using a single high power channel.

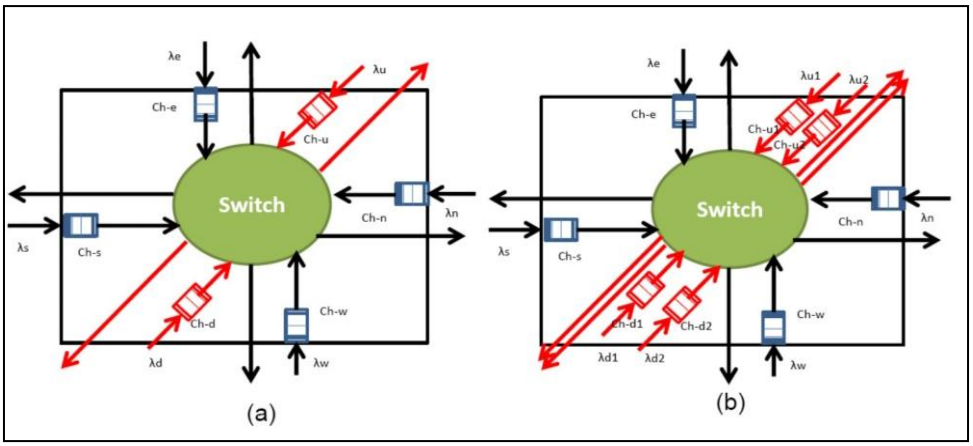


Figure. 5.1. (a) Router for 3D NoC with six bidirectional channels. (b) Router for 3D NoC with six bidirectional channels having more than one virtual channel for Up and Down ports.

Further, the proposed FIFO buffer based router with north, south, east, west, up and down direction and the next end the router with up and down channels with more than one virtual channel to enhance the channel capacity. These channels were proposed to establish connections with the up-down layer in the 3D NoC

architectures (Figure 5.1 (a) & 5.1 (b)). Further, to develop this model, the work of U. Y. Orgas et al. [10] were referred to extend the model developed and to understand the virtual channel capacity throughput and latency at equilibrium.

Furthermore, The average number of input buffer packets, at each input channel east (e), west (w), north (n), south (s), up (u), down (d) is $N = [Ne, Nw,..., Nd]^T$. The arrival is Poisson distribution was considered because the equilibrium equation for the input buffer is valid at any given channel j is $\lambda_j = N_j/\tau_j$, where τ denotes an incoming packet spends an average time in queue j. The average time is for packets waiting until the incoming packet in the other buffer of the same router, packets waiting for response time in the same buffer, and the residual service time provided from the incoming packet. And hence τ_j can be written as follows equation 5.1 [10].

$$\tau_j = TN_j + T\sum_{K=1,K\neq j}^{P} c_{jk}N_k + R \qquad (5.1)$$

Where the coefficients c_{jk} represent the probability of conflict as channels j and k contend for the same performance, for example, The East and West channels compete for the North Channel. For packets served before the incoming packet, the first component, i.e., the average waiting time. Moreover, in addition, c_{ij}' shows the probability of contention between arriving packets at ports i and j in compared with incoming packets. Here the packet service time is T and Residual packet waiting time is R. Now let C_e be the row vector C_e= [C_{ee}, C_{ew}, C_{en}, C_{es}, C_{eu}, C_{ed}] of the contention probabilities matrix *C*, where $C_{jj}=1$. Then Equation 5.1 can be written as equation 5.2 [10].

$$\lambda_j = N_j/\left(TC_jN + R\right)$$

$$where\ C_jN = N_j + \sum_{K=1,K\neq j}^{P} c_{jk}N_k$$

and $$N_j = \lambda_j T C_j N + \lambda_j R \qquad (5.2)$$

Equation 5.2, describe the buffer's equilibrium condition in one direction, say east. So for the entire router, was to denote the arrival rate to the router, as shown in Figure5.1(a) that describe the diagonal matrix α, N describes the average number of packets on each input channel, the content on matrix C, and the dual time R equated by equation 5.3.

$$\alpha = \begin{bmatrix} \lambda e & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda w & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda n & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda s & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda u & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda d \end{bmatrix}_{6x6} \qquad (5.3)$$

where $N=[N_e,N_w,N_n,N_s,N_u,N_d]^T$, $C=[C_e,,C_w,C_n,C_s,C_u,C_d]^T$

$$and \quad \bar{R} = R[1,1,1,1,1,1]^T$$

Then the equilibrium condition for the number of packets overall buffer in various directions for the router can be written as equations 5.4, 5.5 and 5.6 [10].

$$N = T\alpha CN + \alpha\bar{R} \qquad (5.4)$$

$$\alpha\bar{R} = (1 - T\alpha C)N \qquad (5.5)$$

$$N = (1 - T\alpha C)^{-1}\alpha\bar{R} \qquad (5.6)$$

5.2.1. Router with physical channels with multiple virtual channels

The router model as shown in Figure 5.1, in which the total number of packets at the buffer of each port can be expressed in the form as in equation 5.7 [11] [12]

Packet contention probability C with appropriate router design specifications and applications such as CBR and FTP for traffic arrival rates. Further, let equation 5.6 provides the probability of a packet coming to channel *i* and leaving the router via channel *j* to be f_{ij}. Then, the forwarding probability matrix can be expressed as in equation 5.7, where λ_{ij} is the input channel *i* the arrival rate of traffic is routed to the output channel *j* [10].

$$F = \begin{bmatrix} 0 & f_{ew} & f_{en} & .... & f_{ed} \\ f_{we} & 0 & f_{wn} & .... & f_{2d} \\ .... & .... & .... & .... & .... \\ f_{de} & f_{dw} & f_{dn} & .... & 0 \end{bmatrix} \quad (5.7)$$

$$\text{Where } f_{ij} = \lambda_{ij}/(\sum_{k=1}^{P} \lambda_{ik}) \; ; 1 \leq i, j \leq P \quad (5.8)$$

If the forwarding probabilities are independent, then the probability of contention can be expressed in terms of forwarding probabilities is given by equation 5.9 [10],

$$1 \leq i, j \leq p, i \neq j, c_{i,j} = \sum_{k=1}^{P} f_{ik} f_{jk}, i = j, c_{ii} = 1 \quad (5.9)$$

The physical channels can be improvised by having multiple virtual channels, which can use a multiplexing method to reply the input received. In this situation, the average number of packets for each virtual channel can signify the router's condition (Figure 5.2 (d)). Therefore the physical channel number of the router layout element is (P) times the physical layer number of virtual channels (V). So, it can further be useful equation 5.2 to compute the equilibrium condition of the number of packets in the buffers connected by virtual channels say up the channel as follows as described in equation 5.10.

$$N_u = \lambda_u T_{virtual} C_u N + \lambda_u \overline{R} \qquad (5.10)$$

Where '$T_{virtual}$' is the service time of the virtual channels, and C_u is the contention vector for the upper channel. However, it is possible to make a qualitative comparison among the equilibrium conditions of the buffer in the physical channel with and without virtual channels. In general, to demonstrate that the NoC architecture of the virtual channel is better than the NoC physical channel architecture, it is a need to prove $N_{Noc\text{-}virtual} \leq N_{NoC\text{-}physical}$ as given by equation 5.11 with the help of equation 5.7.

$$\left(1 - \alpha T_{virtual} C_{virtual}\right)^{-1} \alpha \overline{R} \leq \left(1 - \alpha T_{physical} x C_{physical}\right)^{-1} \alpha \overline{R} \qquad (5.11)$$

The service times $T_{virtual}$, and $T_{physical}$ depends on the micro-architecture of the front-end of the router input port. Typically it consists of how many arbitration clock cycles are needed for buffer read/write, packet routing, machine cycle, and clock frequency for each operation. One should note that Mutual bus and P2P implementations benefit from control logic flexibility and clock frequency, and the clock frequency is small because of the large potential of the interconnections. In the case of the router with and without virtual channel logic, $T_{vertical}$ and $T_{Physical}$ are almost the same. Therefore, by considering the same rate of generation in both configurations, the router Capacity is enhanced in the virtual channel as explained earlier, and hence $N_{Noc\text{-}virtual} \leq N_{NoC\text{-}physical}$ is ascertained.

5.2.2. NoC with the combinations of channels in the sphere

The probability of contention can be expressed in terms of probability of routing, as seen in equation 5.7 estimating the probability of the routing. The maximum number of packets in the channel where the router configuration

dimension is the total number of input channels is the number of physical channels (P). Hence, making a qualitative comparison among equilibrium conditions of the buffer in the physical channel at various spatial spheres, as shown in Figure 5.2 (a-c) i.e., the channel at the centre, mid-centre, and at the periphery of 2D mesh. In general, the equilibrium condition for the summation of the number of packets over the vertical buffers in up and down directions for the router as in equation 5.12.

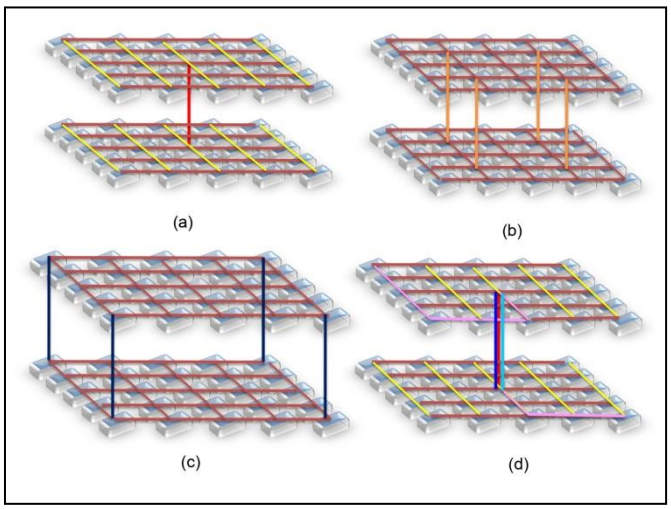


Figure. 5.2. (a) 2 layer 3D NoC model having one vertical channel at centre sphere link; (b) having 4 nos. the vertical channel at mid sphere links. (c) Having 4 nos. vertical channel at periphery sphere links (d) 2 layer 3D NoC model having 4 number virtual channel at centre sphere link.

$$N_l = \sum_{\beta}^{m}\left((1 - \alpha T_m C_m)\alpha \bar{R}\right);\ Where, \beta = 1,2,3,4 \qquad (5.12)$$

Where β is the number of vertical channels at various spheres over NoC, in this case, β= *[4],[4],[4];* in the centre, mid-centre and periphery positions respectively, wherein l is the number of spheres, here l = *1,2,3* $\equiv\equiv \beta$ = *[4],[4],[4]* vertical channels, as there are only three-sphere in spatial dimensions of NoC mesh. Further, the total equilibrium condition overall sphere, i.e., periphery, mid-centre, and centre, can be expressed as in equation 5.13.

$$N_T = \sum_{\gamma}^{P} N_l \qquad (5.13)$$

Here N_T is a summation of the total number of packets at equilibrium over the various spatial sphere $\gamma = 1,2,..,p$ (where in p= 3 in Figure 5.2(a-c)) of buffers in up and down directions for the router over the 2 layers.

5.2.3. Latency over the Packets

Ideal packet latency, T_{ideal}, over the average wire length between the source and destination D (the Manhattan distance) with the channel bandwidth B for the packet size L with the propagation velocity v is given by equation 5.14 [13].

$$L_{Ideal} = (D/v + L/B) \qquad (5.14)$$

Considering the contention probabilities C in case of packet-switched network, the latency over the first to last flit of the packet [13] being injected from source to destination node is given by equation 5.15.

$$L = \left(D/v + L/B + H \cdot T_{router} + T_c\right) \qquad (5.15)$$

There, the average hop-count is H, the one-router delay is T_{router}, and the contention delay is Tc. The $H\ T_{router}$ is time spent in the router coordinating the multiplexing of packets, while T_c is the contention delay spent waiting for the same port resources. The latency expression can be reduced to the objective function, as described in equation 5.14. For that, we have to assume that the network is contention-free wherein $Tc = 0$. Further, since links are connected in horizontal and vertical directions only, we use two-dimensional variables $X_{i,j,k,}$ and $Y_{i,j,k}$ to indicate the wire router in x or y-direction. L and T_{router} are constants. The packet latency for source to a destination over various routers can be summed up in the bi-directional variables. Let $T_{i,j}^{m}$ and $T_{i,j}^{l}$ be the packet latency in the mesh network and long-

link (or vertical link connecting 2D NoC) network, respectively. Then the objective function [14] can be expressed as equation 5.16.

$$MAX\left(\sum_{k=0}^{P-1}\sum_{i=0,i\neq 0}^{N-1}\sum_{j=0}^{N-1}\left(X_{i,j,k}+Y_{i,j,k}\right)\left(T_{i,j}^{m}-T_{i,j}^{l}\right)\right) \qquad (5.16)$$

This function described the various routes from the source to a destination over layers k of 3D NoC. The packets are routed over various 2D NoC layer k with the help of two-dimensional variables $X_{i,j,k,}$ and $Y_{i,j,k,}$ and then delayed by the latency of $T_{i,j}^{l}$ to connect to up/down layer and continues to search the destination. As said earlier, if every 2-D layer is interconnected with every other 2-D layer requires huge numbers of global interconnect that has a detrimental effect on the overall chip dimension, and thus D. Hence, the reduction in the global connect over the bisection plain of the network is proposed as given in Figure 5.2 (a-c) over various spheres of the 2D NoC. So reducing the global interconnect (especially between the layers of 2D NoC), one can reduce the routes from source to destination in the objective function [14].

5.3. Experiments and Results

Case study 1

This chapter considers some variants of 3D Mesh topologies in this paper for study. Here in these topologies, each switch has four neighboring switches, and a computing resource, hence the switch number is equivalent to the resource number. The resources are connected through the communication channel, which is two unidirectional links between a switch and a resource. Here, apply the deterministic XY routing algorithm. Here measure throughput as the total number of received packets by the destinations per unit time. Table 5.1 showed the various scenarios of 4x4 and 8x8, 16x16, 32x32, 64x64 mesh topologies over FTP and CBR traffic applications.

Table 5.1 Various scenarios of 4x4 and 8x8, 16x16, 32x32, 64x64 mesh topologies over FTP and CBR traffic applications.

NoC Model Parameter		Parameter Constraint applied in NS2	
Topology		Mesh	
Connections		Resource-Router, Router-Router	
Transmission Protocols		Transmission Control Protocol (TCP)	
Routing Scheme		Static	
Routing Protocol		Shortest Path	
Queue Mechanism		Drop Tail (FIFO)	
Simulation time		20 seconds	
Number of Nodes		4X4 (16 nodes) and 8X8 (64 nodes), 16x16 (256 nodes), 32x32 (1024), 64x64 (4096)	
Scenario 1	Scenario 2	Scenario 3	Scenario 4
CBR :10Mbps; Link delay: 10ms; Link BW: 10Mbps Queue Size: 100; Varying packet size: 0.1 to 16000 Kbytes	CBR :10Mb/s; Link delay: 10ms; Link BW:10Mbps Queue Size: 5 to 200; Fixed packet size: 0.512 and 64 Kbytes	CBR : 10Mb/s; Link delay: 10ms; Link BW varying: 10 to 200; Queue Size: 100; Fixed packet size: 0.512 and 64 Kbytes	CBR : 10Mb/s; Link delay: 10-200ms; Link BW: 10Mbps; Queue Size: 100; Fixed packet size: 0.512 and 64 Kbytes

Packet Loss happens when one or more packets do not reach their destination due to the error introduced by the network, the contention for network link or lack of buffer space etc. NS-2 is an exceptionally normal and broadly utilized used tool to simulate and expansive territory networks. Because of likenesses amongst NoCs and networks, NS-2 has been a decision of numerous NoC researchers to simulate and watch the conduct of a NoC at a higher reflection level of design. Based on this fact, we have successfully simulated a 16x16 nodes 2D Mesh, 2D Torus, 3D Mesh-based NoC using our reliable protocol for safe delivery of packets.

Scenario I: Throughput and delay calculation with varying packet size

The link bandwidth and delay of the link were kept constant at 10Mbps and 10ms, and the bulk data size and constant data were generated using FTP and CBR (10Mbps) rate application with varying packet size generated.

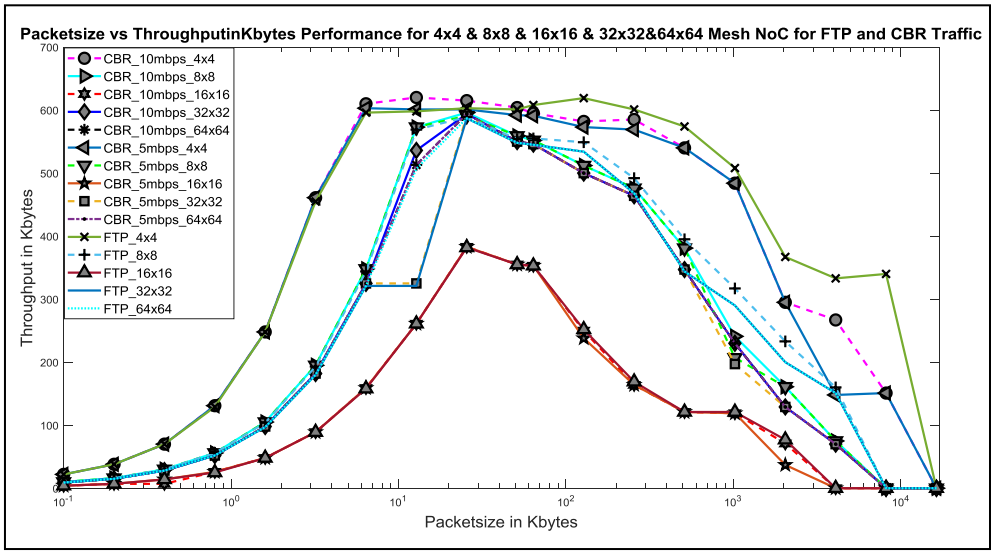


Figure. 5.3 (a) Packet size v/s Throughput for 4X4 and 8X8, 16x16, 32x32, 64x64 Mesh topology for CBR and FTP application.

It is observed (Figure 5.3 (a)) that for FTP and CBR application as the packet size increases throughput increases linearly initially and saturates for packet size in the range of 6 to 512 Kbytes per second having a maximum value of 619 Kbytes/s which is corresponding to 4.952Mbps (close to 50% of the bandwidth since it is the simplex link) and later degrades. The throughput for the 4X4, 8x8 is higher than 16X16, 32x32, and 64x64 mesh topology, which is due to the higher length of the path.

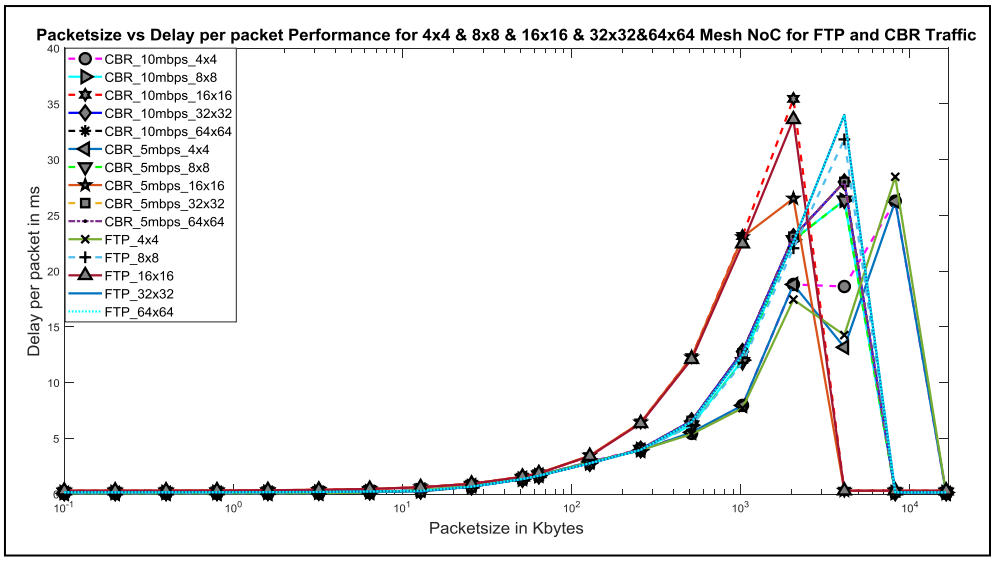


Figure. 5.3 (b) Delay per packet v/s Packet size for 4X4 and 8X8, 16x16, 32x32, 64x64 Mesh topology for CBR and FTP application.

Further, the study on the performance of delay (Figure.5.3 (b)) using FTP and CBR (10Mbps) rate applications with varying packet sizes generated. It is observed (Figure 5.2 (a)) that for FTP and CBR application as the packet size increases, delay increases slowly and is highest from the 1000 to 10,000 packet size and later drop down drastically. This decrease is due to a drop in the throughput from 4.952Mbps to 320 Bits per second. The delay for the 64x64, 32x32, 16X16 is higher than the 8x8, 4X4 mesh topology, which is due to higher path length.

Scenario II: Throughput and delay calculation with varying queue size with low and high load packets

In this scenario, each link of bandwidth was kept at 10Mbps, the propagation delay of each link at 10ms, the bulk data size, and constant data was generated using FTP and CBR (10Mbps) rate application with varying queue size from 5 to 200. It is observed (Figure. 5.4 (a)) that the performance of low load i.e., 0.512 Kbytes is very low as compared to a high load of 64 Kbytes having a factor of 4-5 times. Probably because at a low load of 0.512Kbytes, the network resource i.e., queues size, is not utilized efficiently. Also, it may be noted that within the low load, the 4X4 topology has higher throughput as compared to 8X8, 16x16, 32x32, 64x64 topology is due to short path length. Therefore holds the same for high load packets wherein the 4X4 topology has higher delay throughput as compared to 8X8, 16x16, 32x32, 64x64 topology is due to short path length.

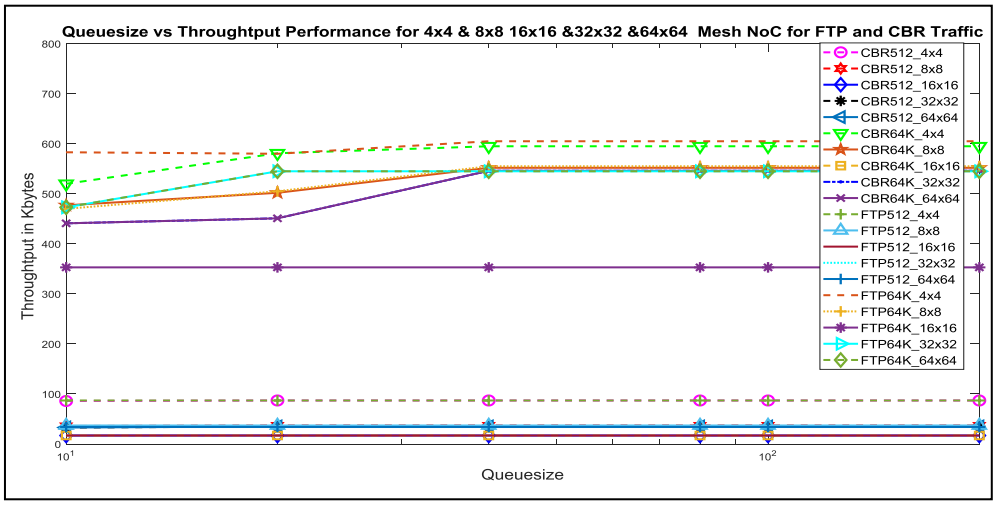


Figure. 5.4 (a) Throughput v/s Queue size for 4X4 and 8X8, 16x16, 32x32, 64x64 Mesh topology for CBR and FTP application.

Similar performance was studied for transmission delay from N (0) to N (15) and N (0) to N (64), N (0) to N (256) source, and the destination node, respectively. While it is observed (Figure. 5.4 (b)) that the delay performance of low load, i.e., 0.512Kbytes is

very low as compared to a high load of 64 Kbytes having a factor of 4-5 times on an average. Due to the low load of 0.512Kbytes, the network resource, i.e., queues size, is not utilized efficiently. Further, it may be noted that within the low load, the 4X4 topology has better performance over transmission delay as compared to 8X8, 16x16, 32x32, 64x64 topology. It may be noted that the delay performance over higher load is saturating at the queue size of 40. Therefore, it could be the optimized size of the queue at that packet size of 64Kbytes. Further, it may be noted that from the queue size of 10 to 40, the delay performance is gradually increasing linearly, and here it is again better performance for transmission delay over 4X4 topology as compared to 8X8, 16x16 topology, which is due to higher path length.

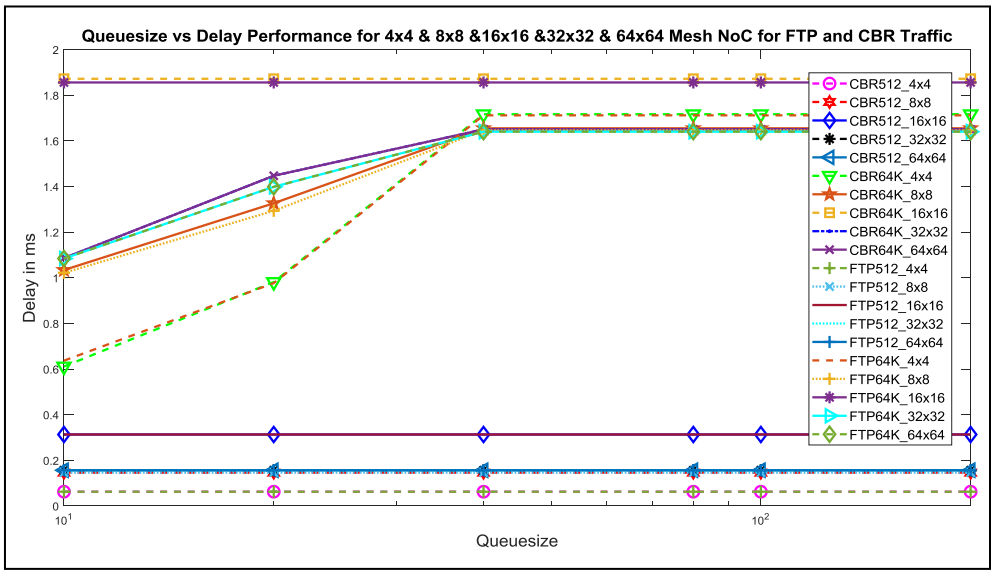


Figure. 5.4 (b) Delay per packet v/s Queue size for 4X4 and 8X8, 16x16, 32x32, 64x64 Mesh topology for CBR and FTP application.

Scenario III: Throughput and delay calculation with varying link Bandwidth for low & high load packets

In this scenario, the queue size was kept 100, and the propagation delay of each link is at 10ms, where the bulk data size, and the constant data was generated using FTP

and CBR (10Mbps) rate application with varying link bandwidth from 10 to 200. Further, it is observed (Figure 5.5 (a)) that the throughput performance for low load, i.e., 0.512Kbytes is very low below 100 Kbytes compared to a high load of 64 Kbytes having an average factor of 20-25 times. Probably due to the low load of 0.512Kbytes, the bandwidth is not exploited. Hence the throughput remains very low. Whereas at the high load of 64Kbytes, FTP performs better as compared to CBR, as the CBR rate is low to exploit the given bandwidth, and hence it remains saturated even after increasing bandwidth. While the FTP application, the bandwidth is linearly increasing as the bandwidth increases, and it is found to be best for 4X4 topology as compared to 8X8, 16x16, 32x32, 64x64 topology due to path length.

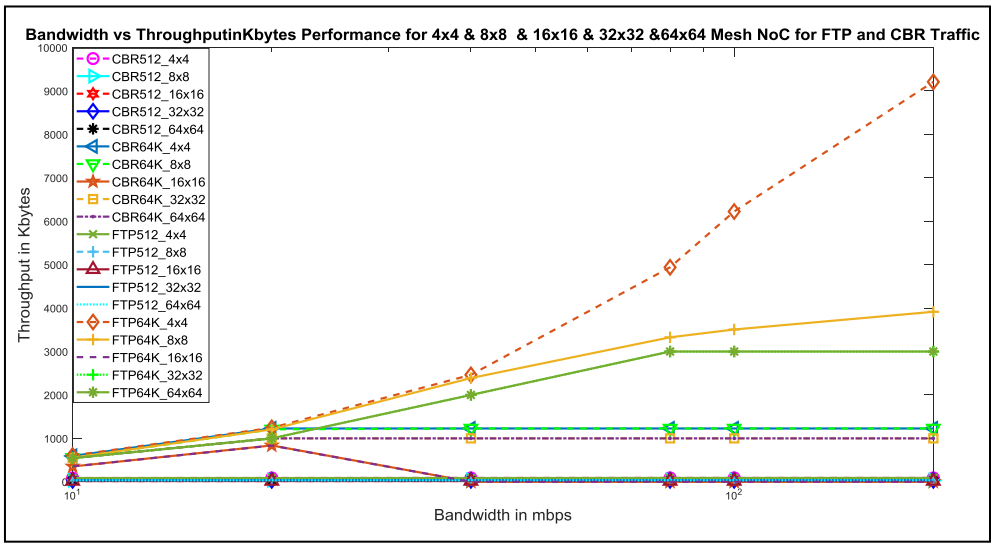


Figure. 5.5 (a) Throughput v/s Bandwidth for 4X4 and 8X8, 16x16, 32x32, 64x64 Mesh topology for CBR and FTP application.

Further, similar performance was studied for transmission delay from N (0) to N (15) and N (0) to N (64), N (0) to N (256) source, and the destination node, respectively. It is observed (Figure 5.5 (b)) that the delay performance of low load, i.e., 0.512Kbytes is better as compared to a high load of 64 Kbytes having factored of 4-5 times on an

average between the bandwidth of 10-40 Mbps. The high performance of the transmission delay is due to the size of the packets 128 times smaller than the high load packet. Thus the initial delay is high for the low link bandwidth within 10-40 Mbps as for the higher load, and the bandwidth is not enough for transmission, which satisfies after at 40Mbps for 64Kbytes load.

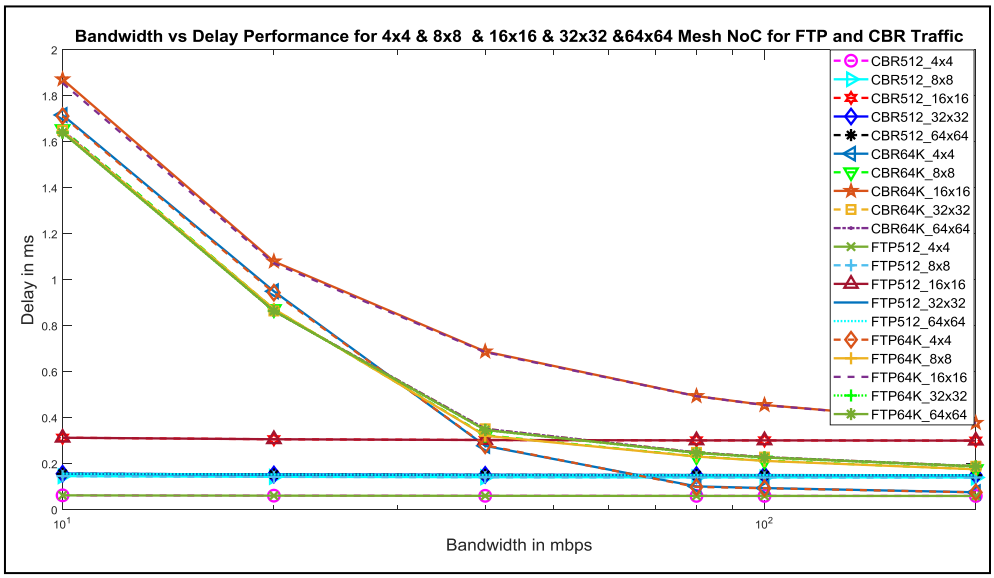


Figure. 5.5 (b) Delay per packet v/s Bandwidth for 4X4 and 8X8, 16x16, 32x32, 64x64 Mesh topology for CBR and FTP application.

Scenario IV: Throughput and delay calculation with a varying propagation delay of link with low & high load packets.

In this scenario, the queue size was kept as 100, link bandwidth from 10Mbps and the bulk data size and the constant data was generated using FTP and CBR (10Mbps) rate application with varying link propagation delay each link within 10-200ms. It is observed (Figure 5.6 (a)) that the throughput performance for low load, i.e., 0.512Kbytes, is very low below 100 Kbytes compared to a high load of 64 Kbytes factor of 4-5 times, having an average value of having an average of 400 Kbytes. However, the throughput of the high load decreases as the link propagation delay

increases. Hence, the 4X4 topology performs better compared to 8X8, 16x16, 32x32, 64x64 due to low path length. And the decrease in performance at the higher propagation delay is natural and implicit.

Similar performance was studied for transmission delay from N (0) to N (15) and N (0) to N (64), and N (0) to N (256) source and the destination node, respectively. It is observed (Figure 5.6 (b)) that the delay performance is better over topologies as compared to scenarios 2 and 3, which was for high and low loads. It is found that the 8X8 topology the delay performance is poor at 1.5-factor times. The transmission delay remains within the average value of 1.5ms for high load packets, i.e., 64Kbyts, using FTP and CBR (10Mbps) rate application. Probably due to the propagation delay increases, the packets transverse time matches with that of propagation delay and the system seem to be in unison within specific propagation link days limits. While in other applications low load 8X8, 16x16 topology and 4X4 load topology, the packets delay performance degrades.

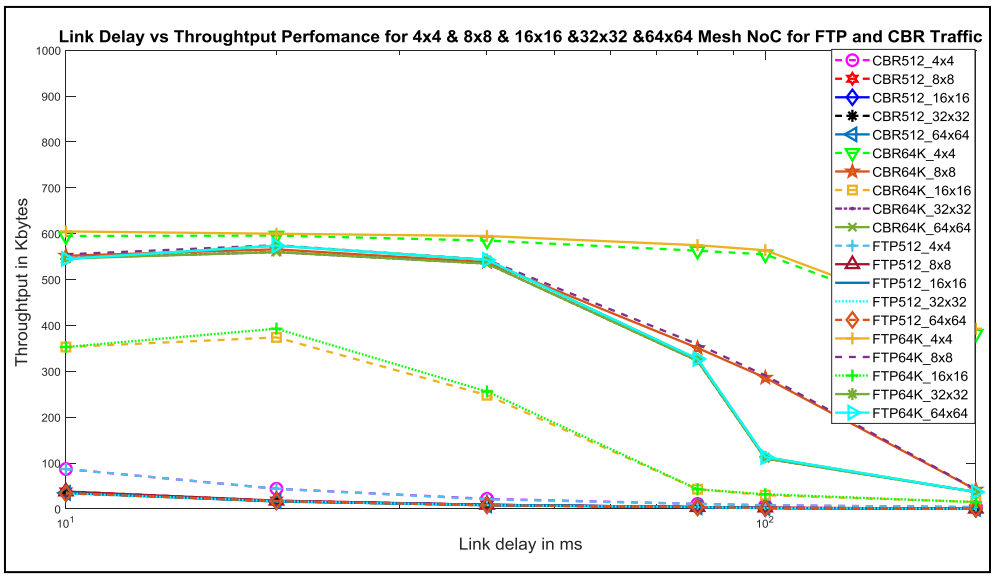


Figure. 5.6 (a) Throughput v/s Link delay for 4X4 and 8X8, 16x16, 32x32, 64x64 Mesh topology for CBR and FTP application.

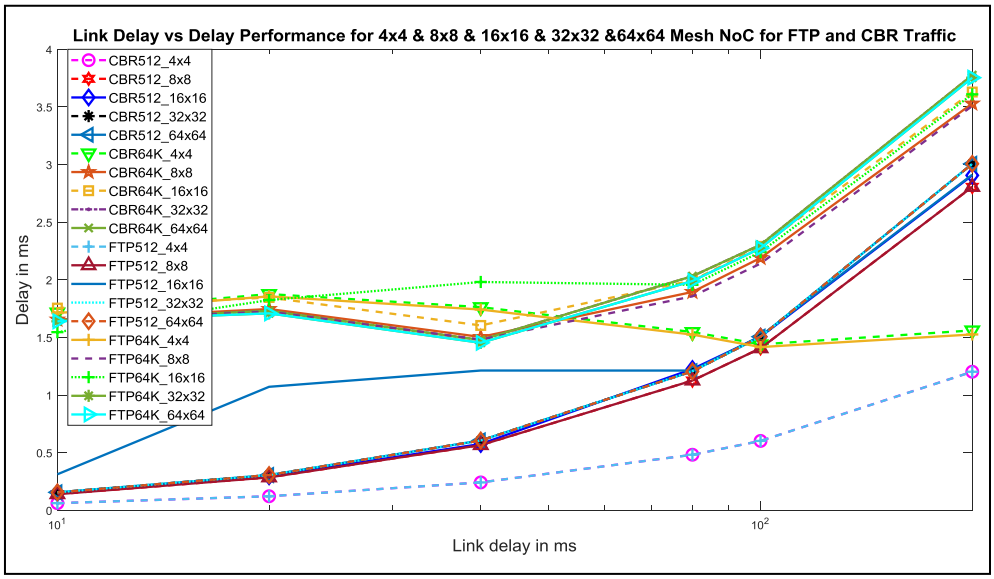


Figure. 5.6 (b) Delay per packet v/s Link delay for 4X4 and 8X8, 16x16, 32x32, 64x64 Mesh topology for CBR and FTP application.

Case study 2

The study was considered for some variants of 3D mesh topologies, i.e., 2, 4, 8 layers the simulation study. While the multilayer variants are considered to demonstrate the effectiveness of the TSV's like links, between the layers. Here in the 2-D dimensional mesh topology, each router is routed to four nearby routers and a computational resource. The routers and computational resources are also related through communication channels. Through channel is linked unidirectional between routers and computational resource. Hence, the deterministic XY routing algorithm for routing over this topology was used. The vertical channels, i.e., up and down (refer Figure. 5.1 (a) and (b) for red colour links) establish a connection between the 2-dimensional layers. The performance parameter, i.e., throughput, is the total number of received packets by the destinations per unit time. When one or more packets do not reach their destination due to the contention over a network link or lack of buffer space, etc. the error introduced over the network, which results in the packet loss. A standard tool was used to simulate small and large area networks is NS-2. Since there are

similarities between NoCs and networks, NS-2 is preferred by researchers to understand behaviours at a higher abstraction level of design. There are various topologies and protocols support, and even customized protocols can be incorporated. Therefore, the parameters for routers and links can easily be scaled to simulate the real situation a chip. With this understanding, stimulation of 4x4 and 18x18 nodes 3D mesh-based multilayer NoC for performance analysis for packets delivery.

Scenario I: Evaluation of throughput and delay for 3D 18x18 Mesh NoC topology over varying packet size with 10-50% break-in overall links number for CBR and FTP.

The link bandwidth and delay of the link were kept constant at 10Mbps and 10ms, and the bulk data size and constant data were generated using FTP and CBR (10Mbps) rate application with varying packet size. This case study explores the performance analysis throughput and delay of 2layer 3D 18x18 mesh NoC for faults over 10, 20, 30, 40, 50% of breaking links. As more percentage of links are disturbed, the magnitude of the shortest path over the network increase, decreasing the throughput and increasing the latency (refer Figure 5.7 & Figure 5.8) over FTP and CBR applications.

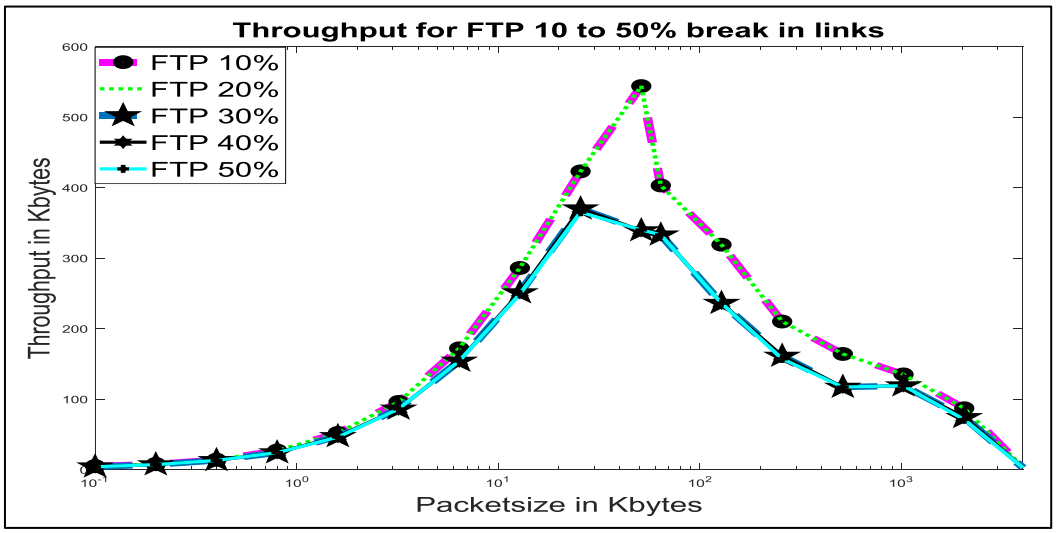


Figure. 5.7 Throughput v/s packet size for 2layer 18x18 3D Mesh NoC with 10, 20, 30, 40, 50 % of breaking overall links nodes for FTP application.

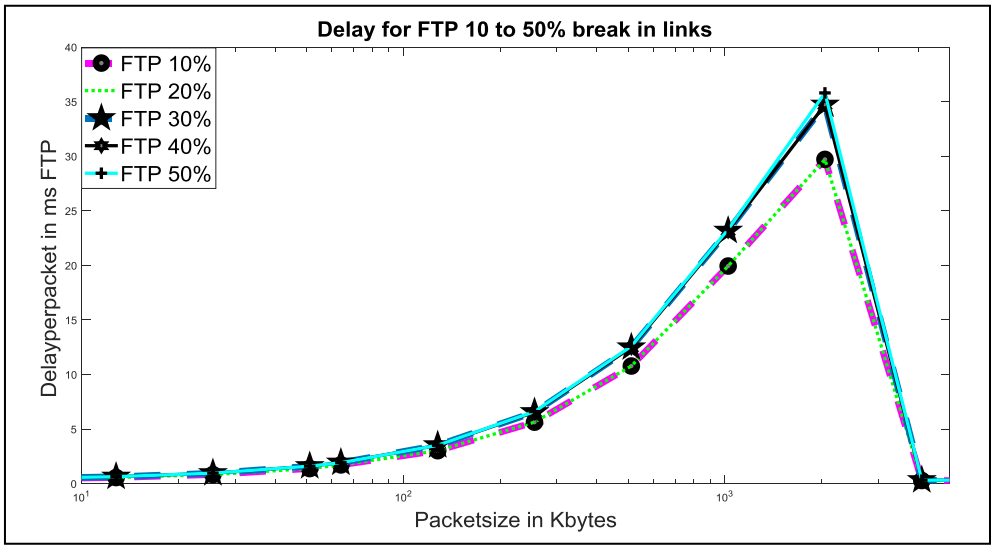
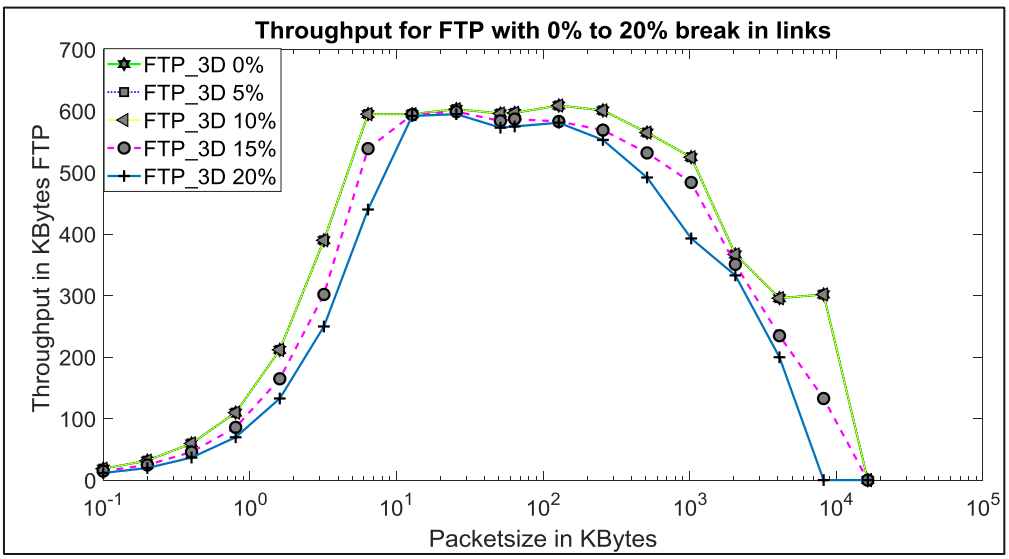
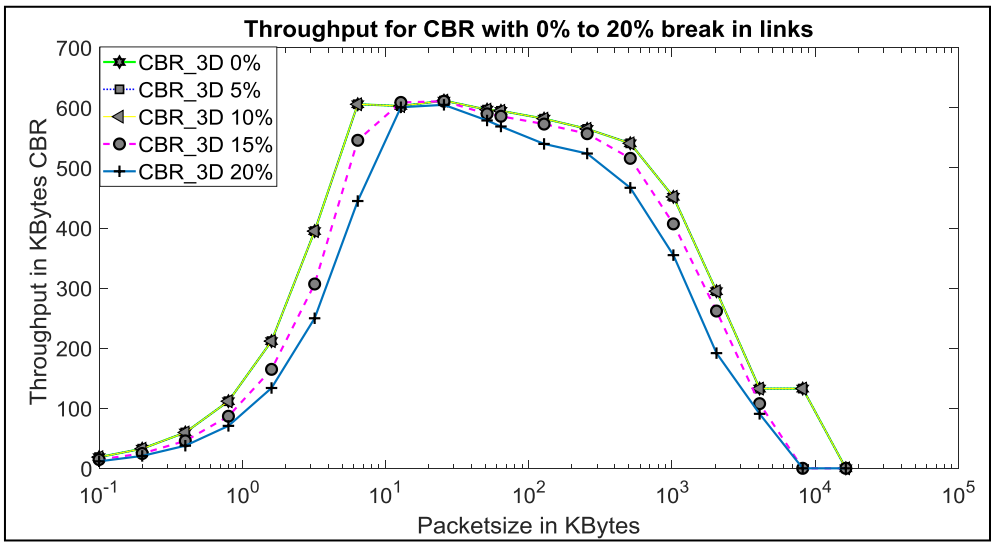


Figure. 5.8 Delay v/s packet size for 2layer 18x18 3D Mesh NoC with 10, 20, 30, 40, 50 % of breaking overall links over for FTP application.

Scenario II: The performance matrix for throughput and delay plot over 2 layers 3-D 4x4 mesh topology as a function of both varying packets size and random breakdown in horizontal links (0-20%). Here robustness of network is demonstrated for CBR and FTP applications.



(a)



(b)

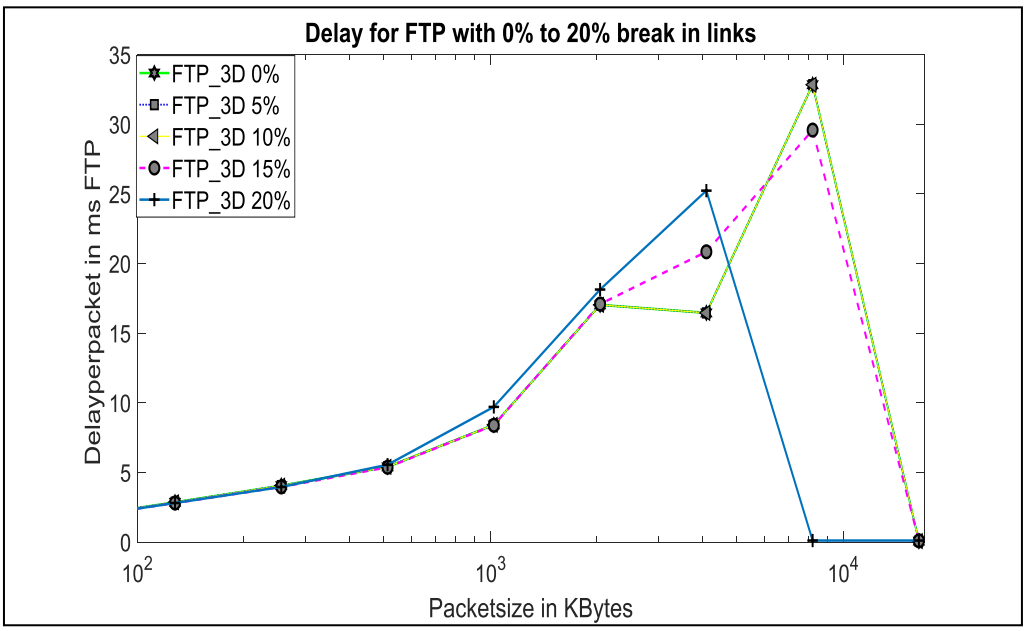
Figure. 5.9 Throughput v/s Packet size for 3D 4x4 2-layer NoC Mesh topology for random break-in links (0, 5, 10, 15, 20 %) (a) FTP traffic. (b) CBR traffic.

The bandwidth and delay of the horizontal link are kept constant at 10Mbps and 10ms, and the bulk data size and constant data are generated using FTP and CBR (10Mbps) with varying packet size (10016384kbps). However, the single vertical link having BW of 10MBps connecting (refer Figure 5.2 (a)) between the layers of the 3D NoC has been established for connectivity, which demonstrates the scenario of single TSV for the performance study.

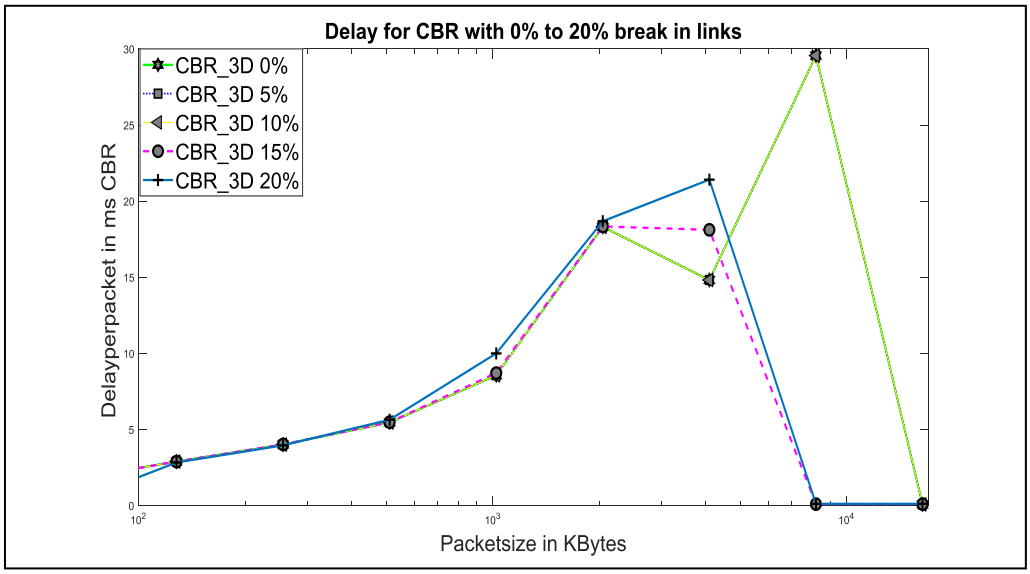
Figure 5.9 (a) & (b) and 5.10 (a) & (b) shows the total NoC throughput and delay performance of 2-layer 3 D mesh topology network. Figure 5.9(a) it is apparent that the network reaches a peak of 5MBps (50% of 10MBps as it is bidirectional) at a throughput of 600KBps at significant packet size of 6KB-9KB. It conveys that 3D NoC network throughputs perform linearly as a function of the size of the packet until the channel capacity is reached. The peak throughput is maintained from 6KB-1000KB for a 10% break-in links due to bottleneck in buffer size assigned to the routers. To be specific, the peak range for 20% break-in links is lower (10KB to 300KB) than the peak

range of 10% links. Once the channel capacity is exceeded by data rate and buffer size, throughput begins to decrease to the point of ruin from 1000KB onwards. As the number of horizontal links is disconnected in a random manner, the shortest path from source to destination increases due to more numbers of hops, and the throughput decreases further below 600KBps. Likewise, the same takeaway can be applied to the CBR application in Figure 5.9(b).

Considering specifically the 20 % break-in links case in Figure 5.9(a), the average peak throughput value of about 550 KBps is reached at a packet size of 6KB-100KB, resulting in lowering throughput value as compared to 0-15 % break-in links. While focusing on the CBR application corresponding plot, the 20% break-in links case in Figure 5.9(b), peak throughput is reached at about packet size of 6KB-9KB. Therefore, it is a reasonable conclusion that FTP and CBR are capable of exploiting the bandwidth at a reasonable packet size in 2 layer 3-D NoC.



(a)



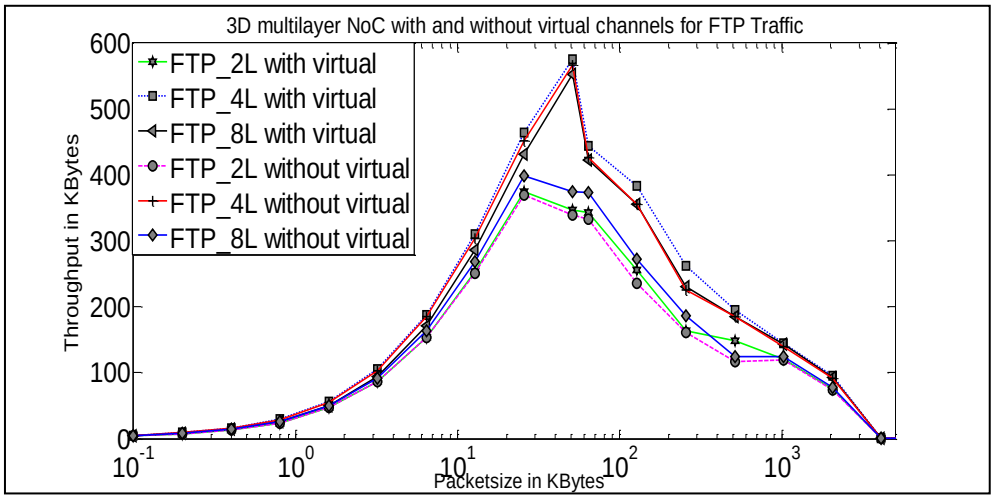
(b)

Figure. 5.10 Delay v/s Packet size for 3D 4x4 2-layer NoC Mesh topology for random break-in links (0, 5, 10, 15, 20 %) (a) FTP traffic. (b) CBR traffic.

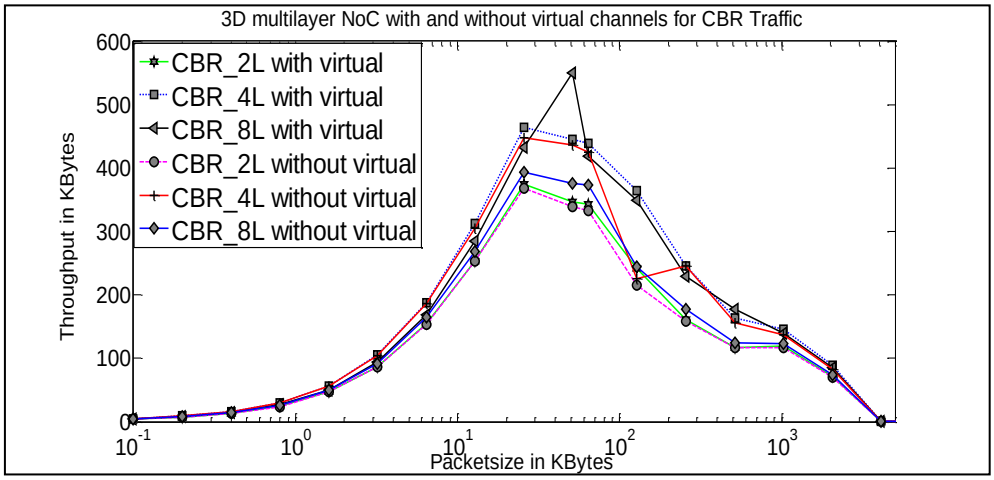
Packet delay is below the link delay of 10 ms) is seen up to 1000KB packets size over a random breakdown in horizontal links (0-20%). However, the network is not congested up to 1000KB packets size. It also conveys that 3-D NoC network latency performs linearly as the function of packet size and delay. In this view, since the throughput is maintained at the peak to 1000KB (Figure 5.9(a) & (b)), a seamless uncongested network is reached to the equivalent value of 1000KB in latency plot. Consider specifically the 20% break-in links case in Figure 5.10(a) for FTP traffic 25ms is the highest delay value obtained for 4000KB while 21ms and 17ms is the value of delay for 15% and 10% break-in links respectively. Above 4000KB packets size, the network is uncertain. Likewise, the same takeaway can be applied to the CBR application in Figure 5.10(b). Therefore, due to the bottleneck of the resources (like buffer size, link bandwidth, etc.) assigned to the routers [15].

Scenario III: The performance matrix for throughput and delay over 2, 4, 6 layer 3-D 18x18 mesh topology (648, 1296, 2592 nodes respectively) as a function of varying packets size is performed with the other parameters as in Scenario II in case study 2.

The single vertical link (bandwidth of 10MBps) and multiple virtual links, which can be implemented using OFDM like channel multiplexing technique (five channels with a bandwidth of 5, 10, 20, 30, 40 MBps) connecting (refer Figure. 5.2(d)) between the layers of the 3D NoC have been used for connectivity, which demonstrates the scenario of the router with and without virtual channels as discussed in sections 5.2.1 and 5.2.2, respectively. In 2-8 layer 18x18 NoC (four-time as in scenario II in case study. 2) the peak throughput is maintained up to an average of 300KBps from packet size of 10KB to 100KB as per the resources in place with the router as discussed in scenario II. With a higher number of layers and more nodes, the shortest distance between the source and destination node increases, increasing the latency for the packet and decreasing the throughput of the network. Specifically for the 2 layers, 18x18 NoC with zero % break-in links, the peak performance throughput range is 300KBps for the packet size of 11KB to 90KB, which is quite below the range of 2 layers 4x4 NoC. Also, the multilayer NoC throughput with-out virtual channels is always marginally lower than with virtual channels, as seen in Figure 5.11(a). This case study demonstrates the effectiveness of virtual channels over non-virtual channels, i.e., a physical channel. Hence, the effectiveness of additional resource, i.e., the virtual channels to enhance the bandwidth of the physical channel, thereby increases the performance. Likewise, the same takeaway can be applied to the CBR application in Figure 5.11(b).



(a)

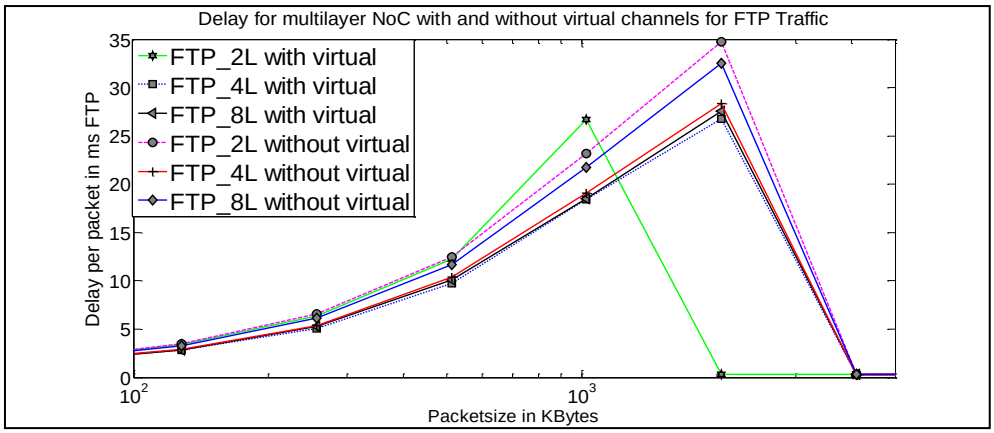


(b)

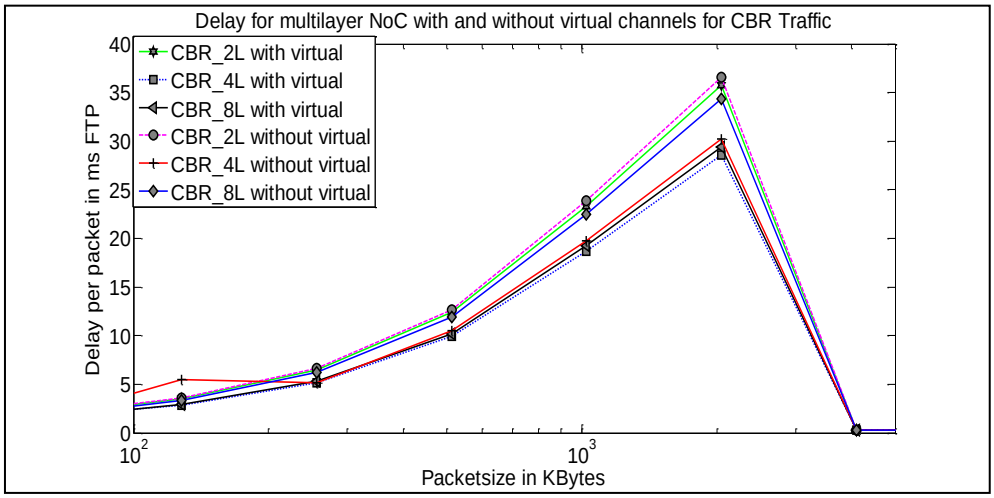
Figure. 5.11 Throughput v/s Packet size for 3D 18x18 multilayer NoC Mesh topology without a virtual channel and with five virtual channels (5, 10, 20, 30, 40 MBps) (a) FTP traffic. (b) CBR traffic.

From Figure 5.12 (a) & (b) the smooth traffic flow (since the packet delay is below the link delay of 10ms) is seen up to 400KB as compared to 1000KB as in 4x4 NoC network for both the FTP and CBR application. Probably is due to an increase in latency time of the packet being delivered from source to destination, which comprises the average time an incoming packet spends over the shortest path. As discussed, it consists of packets waiting in the other router buffer (participating in the shortest path)

were served at the front of the incoming packet, packets waiting for connection time in all buffer, and the residual service time seen by an incoming packet. Since more than one router is participating in the routing, it is observed to increase in latency from 10ms to 25ms for a packet size of 1000KB as compared to Figure 5.10(a) & (b). It also conveys that 3-D NoC network latency performs linearly as the function of packet size delay.



(a)



(b)

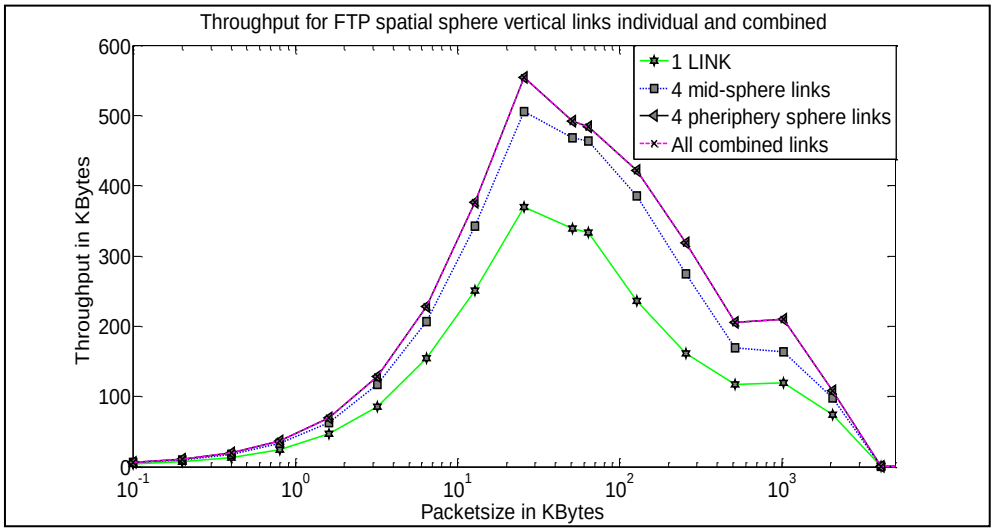
Figure. 5.12 Delay v/s Packet size for 3D 18x18 multilayer NoC Mesh topology without a virtual channel and with five virtual channels (5, 10, 20, 30, 40 MBps) (a) FTP traffic. (b) CBR traffic.

Also, in the multilayer NoC, the latency factor with virtual channels is always marginally higher than without virtual channels due to additional resources of virtual channels. Hence, the result for throughput and latency is in agreement with [17] effective utilization of frequency resources.

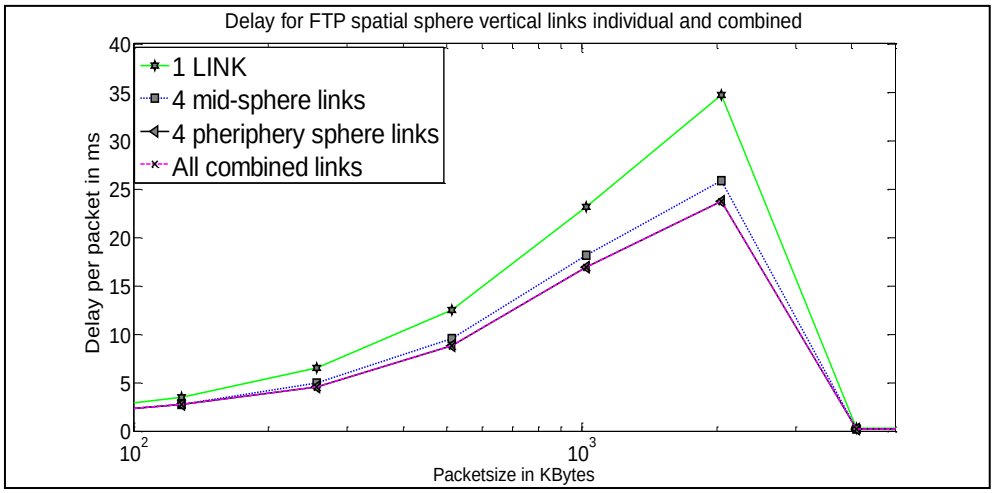
Scenario IV: Evaluation of 2 layers 3-D 18x18 nodes mesh NoC for variation and placement of channels (i.e., vertical channels connecting layers) from various positions (i.e., centre, mid-sphere, and periphery over the sphere) is studied. The study is performed independent and combination channel modes for FTP traffic.

Four physical channels connecting two layers placed equidistance over the spatial sphere, i.e., at the centre, mid sphere and periphery as seen in Figure 5.2(a), (b) & (c), with the bandwidth of 10MBps and propagation delay 10ms are assigned to said NoC network. The source and destination nodes are selected on the periphery sphere. As discussed, in scenario II the case study.2 from Figure 5.9 (a) network reach the peak of 5MBps at the throughput of 600KBps at a significant packet size of 6-9KB. Therefore, the 2 layers 3-D 18x18 Nodes NoC, the throughput is around 500KBps at packet size 30KB. This value is reached when the resources like buffers are fully exploited over participating routers in the shortest path traversal. After 30KB packet size, the performance degrades, as seen in Figure 5.13(a), as the bottleneck is reached. However, compared to scenario II in a case study. 2, the average throughput value of 250 KBps is observed over the packet size of 6KBps-1000KBps as compared to 600KBps. Also, across the 3-D routers placed over spheres from the centre, the performance is lowest for the central sphere router as compared to the peripheral sphere router. Specifically, as seen in Figure 5.13(a) the central 3-D, the router has a peak throughput of 370KBps while peripheral has 550KBps. This is probably due to the source, and destination nodes are on the peripheral sphere and are close to 3-router placed on the peripheral sphere

compared to central 3-router. Hence, with the entire 3-D router enabled, the throughput matches with the peripheral 3-D router value as the source is in close proximity to it.



(a)



(b)

Figure. 5.13 (a) Throughput v/s Packet size for 2 layers 18x18 NoC Mesh Topology without and with virtual channel for FTP traffic (b) Delay v/s Packet size for 2 layers 18x18 NoC Mesh Topology without and with the virtual channel for FTP traffic.

Similarly, the latency of the optimum packet value below the 10ms is obtained up to packets size of 400KB, as observed in Figure 5.13(b). The maximum latency of the average value of 26ms is obtained for the packet size of 2000KBps. Generally, for the central 3-D router, the value is 35ms, while for combination, the value is 23ms. Probably, due to the source and destinations are on the peripheral edge, and hence the shortest path XY routing algorithm on selecting one of the peripheral vertical channels for Z-traversal. Therefore, it concludes that in the case of packets switch network, the traversal over the vertical channel will be based on minimum neighborhood distance of the sphere placed from the source node. Such combination of vertical channel network can help in reducing the latency, as the directional variables $X_{i,j,k}$ and $Y_{i,j,k}$ are minimum for the immediate vertical channels at the outer edge. Similar results are reported [16] for center architecture having much lower average packet latency compared to other TSV reduced architectures, and also the performance of boundary improves as the increase in the number of 3D routers from 25 %, 50%, and 75%. Also, for lower elevator percentage, the network becomes saturated in a lower ordered load such that the elevator percentage of 20% results in an impractical network [17].

Scenario V: Evaluation of throughput and delay for 3D & 4D 18x18 Mesh NoC topology over 50% faults in links over FTP.

The link bandwidth and propagation delay of the link were kept constant at 10mbps and 10ms. The bulk data size and consistent data were generated using FTP and CBR (10Mbps) application over varying packet size. While the source and the destination nodes were over the peripheral sphere. Hence, it was introduced to 2 layers of 18x18 nodes 3D & 4D Mesh NoC at 50% links faults.

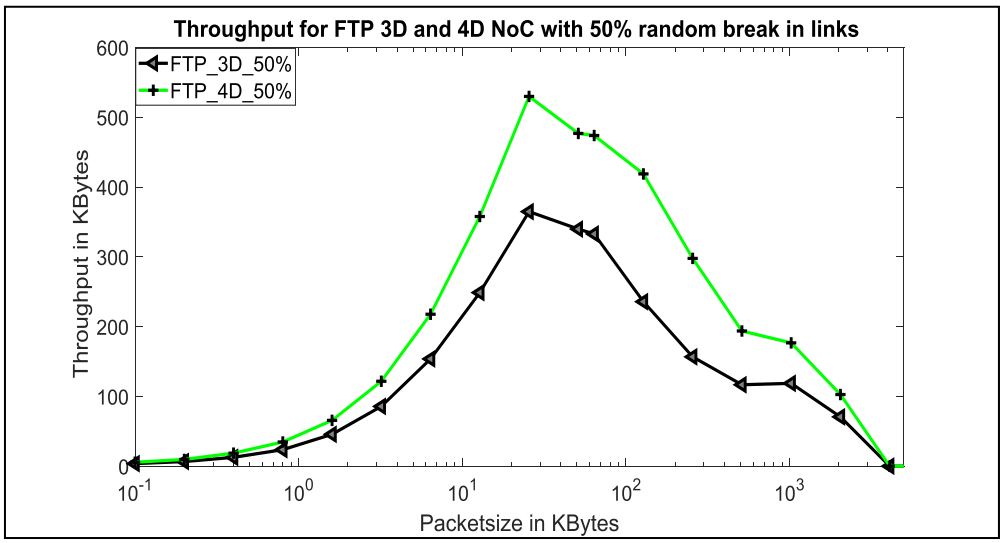


Figure. 5.14 Throughput v/s Packet size for 2 layers 18x18 3D & 4D Mesh NoC with 50% of breaking overall links over topology for periphery sphere vertical channels for FTP application.

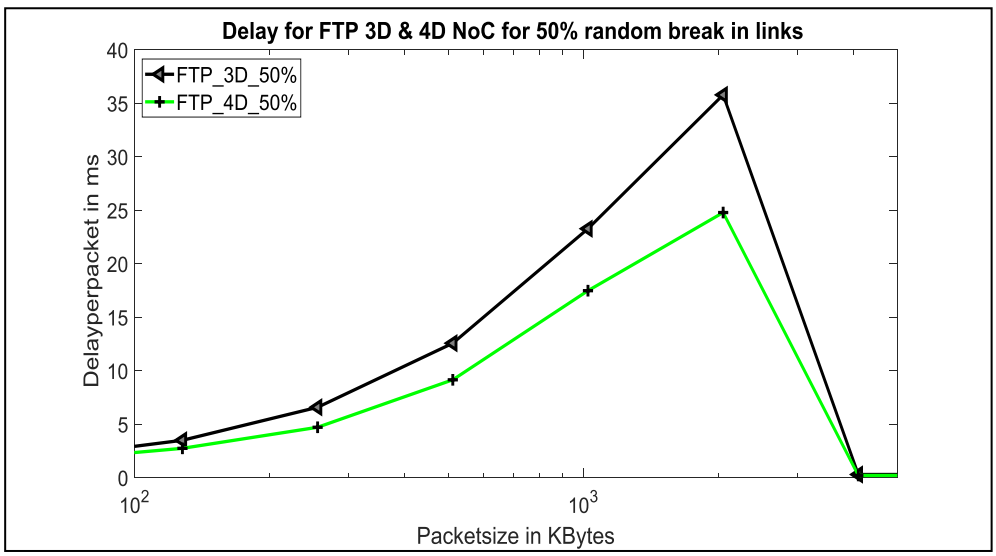


Figure. 5.15 Delay v/s Packet size for 2 layers 18x18 3D & 4D Mesh NoC with 50% of breaking overall links over topology for periphery sphere vertical channels for FTP application.

The 4D NoC shows an increase in the throughput (refer Figure 5.14) as compared to 3D NoC model and a decrease in the latency compared to the 3D NoC model.

Probably due to the 4D NoC model selects the shortest path directly over the layer corner edges between two 3D NoC, and hence the latency also improved (refer Figure.5.14). Similar trends are seen over the CBR application.

Map the source node *S (i,j,k)* in the neighborhood of available sphere *(S_mi.e.S_1,S_2,S_3)* while XY traversal over layer k. Here, first, *for loop,* identify the spatial sphere TS_m for z-traversal for the given source node overall sphere channel *(S_m)*. The second *for loop* identifies the vertical channel *(TV_m)* over selected z-traversal sphere channel *(TS_m)* for Z-axis.

It is considered as the S_m multilayer is 6 layers NoC model, as seen in Figure 5.16. Let's consider that the mesh is 19x19 node having the S_2, S_6, S_{10} at 2, 6, and 10 spatial spheres over 19x19 NoC. Where in the central sphere is at spatial sphere number 2, mid-centre is at 6 and peripheral at 10. Let's assume the source node be at spatial sphere number 5, i.e., *n (5,y,1)* and the destination node beat the sphere 4 *n(4,y,6)* (as considered the spatial sphere are numbered on x-scale one can name it also on y-scale). Also, let s as sum the vertical channels assigned at spatial sphere S_n =3 be at V_1 = 1, V_2 = 3, V_3 = 5, V_4 = 7 in clockwise directions (by considering the reference of frame analog clock, let us assume V_1=1 as 12'o clock position).

Following the above algorithm one can derive that length *(S (i, j, k) − S_m)*=3; and length *(S (i, j, k) - S_{m-1})*=1; i.e., 3 $\leq$ 1, where S_m=2 and S_{m-1}=6 and hence TS_m = 6 i.e., Sphere number to be traverse for the vertical link. Similarly, it can locate as the vertical channel V_m for traversal and can reach the destination.

Algorithm for the adaptive XYZ traversal in the 3D mesh topology

---

1: **Procedure INPUT**:–Read source S(i,j,k) and destination D(i,j,k) nodes over the layers OUTPUT:–Shortest path over the sphere of 3D mesh NoC between sources to destination

2:

3: **For** All Sm in the neighborhood of $S\ (i,j,k)$

4: ***Select the closest spatial sphere TS_m for traversal***

5: If $S\ (i,j,k)\text{-}S_m \leq S\ (i,j,k)\text{-}\ S_{m-1}$

6: Traversal sphere channel as $TS_m=S_m$

7: **Else**

8: Traversal sphere channel as $TS_m=S_{m-1}$

9: **End If**

10: **End For**

11:

12: **For** All V_m **Do**

13: **Select the closest vertical channel** TV_m **for traversal**

14: If $S\ (i,j,k)\ \text{-}V_{mleft} - S \leq (i,j,k)\text{-}V_{mright}$;

15: $TV_m=V_{mle\ ft}$

16: **Else**

17: $TV_m=V_{mleft}$

18: **end Do**

19: **End For**

20:

21: **End Procedure**

5.4. Discussion

The need to expedite the growth of interconnect and packaging technologies is required, as there is a bottleneck over system performance gains, which further, demands the innovation in silicon ancillary technologies. Also, to encourage more than Moor one has to integrate antennas with RFIC to enable phased arrays for millimeter-wave communication [18] or even wireless router ancillary. The heterogeneous integration of high-density bandwidth, low-power communication for upcoming applications is the drivers for innovative technologies. Presently, the state of the Art, 2.5-D integration is an assembly of multiple ICs (in a 2-D plane) over passive interposer substrates, such as silicon, glass, or fine-pitch organic substrates [19] [20] [21].

The optical orthogonal frequency division multiplexing (OOFDM) multiplexing scheme is promising technology for the future high-capacity optical networks. Here, a high bit rate data stream is divided over low bit rate streams modulated over orthogonal subcarriers. These orthogonal subcarriers can be seen as virtual channels integrated into physical channels, as discussed earlier in section 5.2.2. Generally, the subcarriers are generated in the digital domain and hence, they consist of many subcarriers [22]. For example, the subcarriers are generated using coherent Wavelength-Division Multiplexing (WDM) systems in the fiber-optic communication system. Coherent WDM systems do not use training symbols but rely on blind channel estimation [23]. Therefore, information theory states that MIMO systems enhance the channel capacity when a rich scattering environment is adequately exploited. Hence, the combination of the MIMO techniques with OFDM for communication systems is seen as a promising basis for next-generation high data rate communication systems.

Further, the transition of data needs to be between the host and the external memory element. As an appropriate solution for managing data transmission through NoC is identified. The findings obtained show that it is possible to address data faster

and allow for a competent method to share information by means of a specific way between nodes in the network and the arbiter helps to avoid conflict if it happens at all. As the network size grows, latency and throughput values often increase as more nodes appear in the picture.

The TSV's issues in 3D integration and demonstrated the performance of the multilayer (2, 4, 8 layers) 3D NoC for 18x18 mesh topology over varying packet sizes for FTP and CBR traffic. Further, demonstrating the robustness of the NoC networks for random break-in horizontal links of the two-dimensional layer. Also, the performance with and without the virtual channel is demonstrated for the throughput and delay over the fixed set of sources and destinations over the spatial sphere over mesh topology. However, the proposed spatial sphere-based vertical link for optimizing the TSV's in 3D integration for higher dimensional integrated circuits and subsequently higher dimensional NoCs'.

This work reveals a spatially placed sphere-based vertical channel scheme for 3D NoCs to support TSV's density optimization to enhance reliability and yield. One can use the XYZ traversal algorithm for the traversal over the multilayer NoC from source and destination node over the spatial sphere placed to connect the layers (refer Figure. 5.16 layers L_m and spatial sphere S_m) over vertical channels. The listing of the algorithms is described above section.

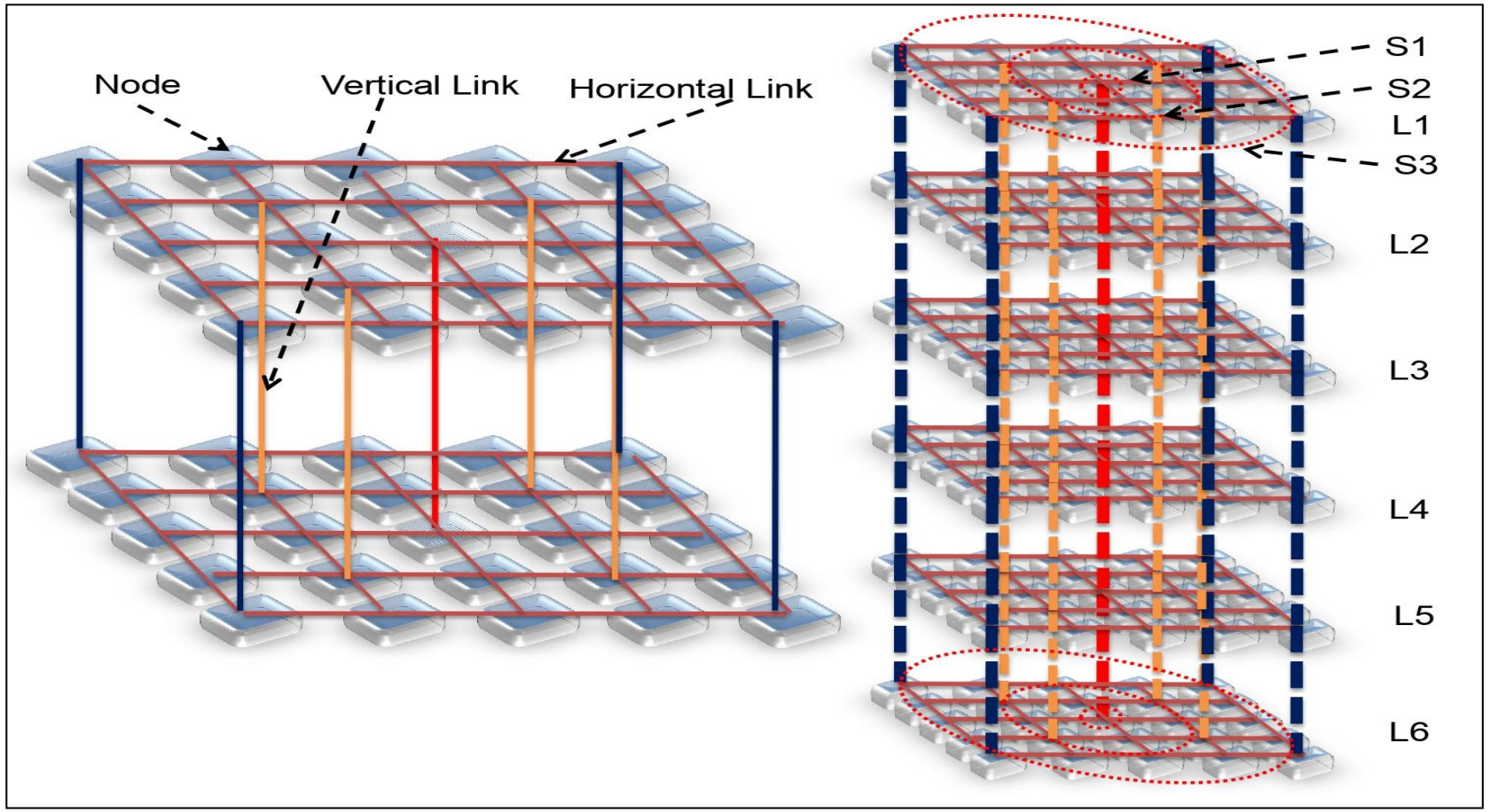


Figure. 5.16 Global NoC model based on the proposed spatial sphere-based Vertical channels.

This scheme has the potential to give optimum performance matrix after adaptively choosing the vertical channels over various spheres for adaptive routing depends on the spatial position of the source and destination nodes. Where the source and destinations are peripheral, then peripheral sphere vertical links have to be chosen for the routing the packets from one to other layers in the 3D & 4D mesh NoC and so on. Link utilization under uniform traffic is higher towards the centre of the NoC than its corners inhomogeneous 3D NoC's. Hence, the decision over the neighborhood sphere, towards the central position has to be given priority over the outer sphere. Probably due to the nature of often used XYZ routing algorithm; central routers relay most of the traffic from different corners to their destinations. Hence, one can conclude that the said scheme can be a good toll to customized the positions of the TSV's for 3D integration to reduce the density of the TSV's and to identify the keep-out zone parameter.

5.5. References

[1] Ullah, M.S., Alharbi, A.G. and Chowdhury, M.H. Capacity of gaussian mimo channel for network intra-chip rf interconnect. In *2017 9th IEEE-GCC Conference and Exhibition (GCCCE)*, Pages 1-9, 2017.

[2] Chang, M.F., Cong, J., Kaplan, A., Naik, M., Reinman, G., Socher, E. and Tam, S.W. CMP network-on-chip overlaid with multi-band RF-interconnect. In *2008 IEEE 14th International Symposium on High-Performance Computer Architecture*, Pages 191-202, 2008.

[3] Deb, S., Ganguly, A., Pande, P.P., Belzer, B. and Heo, D. Wireless NoC as interconnection backbone for multicore chips: Promises and challenges. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, *2*(2):228-239, 2012.

[4] Chang, Y.J., Chuang, H.H., Lu, Y.C., Chiou, Y.P., Wu, T.L., Chen, P.S., Wu, S.H., Kuo, T.Y., Zhan, C.J. and Lo, W.C. Novel crosstalk modeling for multiple through-silicon-vias (TSV) on 3-D IC: Experimental validation and application to Faraday cage design. In *2012 IEEE 21st Conference on Electrical Performance of Electronic Packaging and Systems*, Pages 232-235, 2012.

[5] Khan, N.H., Alam, S.M. and Hassoun, S. Through-silicon via (TSV)-induced noise characterization and noise mitigation using coaxial TSVs. In *2009 IEEE International Conference on 3D System Integration*, Pages 1-7, 2009.

[6] Cho, J., Song, E., Yoon, K., Pak, J.S., Kim, J., Lee, W., Song, T., Kim, K., Lee, J., Lee, H. and Park, K. Modeling and analysis of through-silicon via (TSV) noise coupling and suppression using a guard ring. *IEEE Transactions on Components, Packaging and Manufacturing Technology*, *1*(2):220-233, 2011.

[7] Kim, J., Song, E., Cho, J., Pak, J.S., Lee, J., Lee, H., Park, K. and Kim, J. Through silicon via (TSV) equalizer. In *2009 IEEE 18th Conference on Electrical Performance of Electronic Packaging and Systems*, Pages 13-16, 2009.

[8] Sun, R.B., Wen, C.Y. and Wu, R.B., 2011. Passive equalizer design for through silicon vias with perfect compensation. *IEEE Transactions on Components, Packaging and Manufacturing Technology*, *1*(11):1815-1822, 2011.

[9] J. Proakis and D. Manolakis.*Digital Signal Processing (4th Edition)*. Prentice Hall, 2006.

[10] Ogras, U.Y. and Marculescu, R. Analytical router modeling for networks-on-chip performance analysis. In *2007 Design, Automation & Test in Europe Conference &Exhibition*, Pages 1-6, 2007.

[11] Liu, X., Chen, Q., Dixit, P., Chatterjee, R., Tummala, R.R. and Sitaraman, S.K. Failure mechanisms and optimum design for electroplated copper through-silicon vias (TSV). In *2009 59th Electronic components and technology conference*, Pages 624-629, 2009.

[12] T. C. Huang, U. Y. Ogras, R. Marculescu, Virtual channels planning for networks-on-chip, in: 8th International Symposium on Quality Electronic Design (ISQED'07), Pages 879–884, 2007.

[13] A. Kumar, L.-S. Peh, P. Kundu, and N. K. Jha. Express Virtual Channels: Towards the Ideal Interconnection Fabric. Proc. of the 34th International System on Computer Architecture. pp. 150-161, 2007.

[14] Xu, Y., Du, Y., Zhao, B., Zhou, X., Zhang, Y. and Yang, J.. A low-radix and low-diameter 3D interconnection network design. In *2009 IEEE 15th International Symposium on High Performance Computer Architecture*, Pages 30-42, 2009.

[15] Loi, S. Mitra, T. H. Lee, S. Fujita, L. Benini. A low- overhead fault tolerance scheme for tsv-based 3d network on chip link. *In Proceedings of the 2008 IEEE/ACM International Conference on Computer-Aided Design, ICCAD '08, IEEE Press, Piscataway*, Pages 598–602.NJ, USA, 2008.

[16] M. O. Agyeman, A. Ahmadinia, N. Bagherzadeh. Performance and energy aware inhomogeneous 3d networks-on-chip architecture generation. *IEEE Transactions on Parallel and Distributed Systems,* 27 (6):1756–1769, 2016.

[17] S. Foroutan, A. Sheibanyrad, F. Ptrot. Assignment of vertical-links to routers in vertically-partially-connected 3- d-nocs. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems,* 33 (8):1208–1218, 2014.

[18] X. Gu, A. Valdes-Garcia, A. Natarajan, B. Sadhu, D.Liu,K. Reynolds. W-band scalable phased arrays for imaging and communications. *IEEE Communications Magazine*, 53(4): 196–204, 2015.

[19] B. Banijamali, S. Ramalingam, H. Liu, M. Kim. Outstanding and innovative reliability study of 3d TSV interposer and fine pitch solder micro-bumps. *In 2012 IEEE 62nd Electronic Components and Technology Conference*, Pages309–314, 2012.

[20] B. Sawyer, H. Lu, Y. Suzuki, Y. Takagi, M.Kobayashi,V. Smet, T. Sakai, V. Sundaram, R. Tummala. Modeling, design, fabrication and characterization of first large 2.5d glass interposer as a superior alternative to silicon and organic interposers at 50-micron bump pitch. In *2014 IEEE 64th Electronic Components and Technology Conference (ECTC)*, Pages 742–747, 2014.

[21] Oi, S. Otake, N. Shimizu, S. Watanabe, Y. Kunimoto, Kurihara, T. Koyama, M. Tanaka, L. Aryasomayajula, Z. Kutlu. Development of new 2.5d package with

novel integrated organic interposer substrate with ultra-fine wiring and high-density bumps. In *2014 IEEE 64th Electronic Components and Technology Conference (ECTC)*, Pages 348– 353, 2014.

[22] Shieh, W., Bao, H. and Tang, Y., 2008. Coherent optical OFDM: theory and design. *Optics express*, *16*(2):841-859, 2008.

[23] Agrawal, G.P. Nonlinear fiber optics. In *Nonlinear Science at the Dawn of the 21st Century*, Pages 195-211, Springer, Berlin, Heidelberg, 2000.

Chapter 6

Conclusion

6.1 Conclusion

The data traffic can rise by approximately 30 percent by 2020, which is enabled by current mobile technologies. There is also a need for next-generation wireless networking (5G) aimed at providing a maximum bandwidth and data rate at 1Tbps [1] level. Hence the required flexibility and data-rate can be accomplished by using a better spectral efficiency technology. However, the successor of 3G Technology is 4G, where Orthogonal Frequency Division Multiplexing (OFDM) is used as a modulation system, which offers Ultra Large Band connectivity for a mobile device with more power and data rate and improved coverage relative of 3 G Mobile [2].

Cisco estimated that by 2020, a total of 50 billion new connections will be a part of the IoT [3]. While increasing the billions of connected devices can produce a significant amount of data to be stored and processed. A few papers, therefore, discussed the possibilities of incorporating a compression algorithm into IoT devices or sensor nodes. Thus, compression is the optimal option for rising both storage requirements and input/output processes [4], hence, it is beneficial when transferring data across a network. The data leakage from numerous sensors and other processing devices surpass the capacities of all computing technologies, as the data growth trend is continuing, which is very tough to handle.

Besides, data compression is one of the digital world's major signals processing applications, where the bandwidth is often a powerful resource. However, a challenging task is a compression in wireless sensor networks (WSN), which have a limited amount of energy. Hence the sensor nodes are densely deployed in WSN to have highly clustered data which contributes to the reliability of data. Due to the data from various sensor nodes, the traditional compression methodology is challenged in WSN.

The energy conservation in WSN is one of the major QoS factors towards increasing data transfer is the WSN's, and largest energy-consuming operation, tremendous energy savings can be accomplished by reducing data communication.

Further, MIMO systems can improve the transmission efficiency and strength of the wireless networks without increasing the bandwidth and transmitter power of the system. While there are advanced wireless applications such as WLAN and WiMAX, have incorporated MIMO technology in their communication systems, and the MIMO platform is now widely included in cell 3G and 4G communications. The recent MIMO development researchers have focused mainly on elements of signal processing, channel modeling, and coding, rather than antenna design issues.

Furthermore, Forward Error Correction (FEC) coding has now progressed from 2G convolution/block codes to more effective 3G Turbo codes, as key wireless communication expertise. The latest designers looked for help with the more complex 4G systems somewhere else. However, the Low-Density, Parity-Check (LDPC) encoding format is an attractive intention for these schemes, due to its enormous error correction performance and highly parallel decoding scheme that is capable of operating super-fast 4G Technology, so an LDPC encoder and decoder design are built for 4G technology [5].

In the proposed work, the M-PSK (2, 4, 8, 16, 32), M-QAM (4, 8, 16, 32) modulation performance was evaluated with the Bit Error Rate (BER) and the REMSE. Hence, resulting in the rate of BER to decrease as increase in antennas on both transmitter and receiver sides. Whereas, in the case of even power, the bit error rate in the case of QAM is small when compared to the odd capacity, resulting in QAM to have a rather low BER. Further, the system output depends also on the number of transmitters and receivers. It is also concluded that the bit error rate of 4 transmitters and 4 MIMO receivers is smaller than

that of the first case when using 2 transmitters and 2 receivers. Hence, concluding that the BER and RMSE are very low while large numbers of transmitters and receivers are used.

The present thesis work is carried out based on the application for network on chip design (NoC). Hence, the work conducted in this thesis, along with the experimental results, draws the following specific conclusion, listed as follows.

The first contribution of the present work is the transmission of signals using PCA in the communication system. The study is also carried out the data reduction using PCA for an ensemble of 10, 20, 30, 40, 50 sets of signals for QPSK and QAM (2, 4, 8, 16, 32 constellations) modulation. The RMSE and BER studies were performed for various SNR in the AWGN channel. Further, the experiments were carried out for 10 folds to compute the average RMSE and BER performance. Moreover, the performance of the AWGN channel with Rayleigh fading for the QPSK and QAM modulation scheme was done. Later the work was extended for the various configuration of MIMO for 1x1, 2x2, 3x3, 4x4 and demonstrated that the data transmission could be performed at the low power. Therefore, concluding that the PCA technique for data reduction can ensemble the sets of signals for efficient data communication over MIMO configurations.

The second and third contribution is, the data communication using PCA as source coding over an ensemble of 10, 20, 30, 40, 50 signal sets are demonstrated for M-PSK (2, 4, 8, 16, 32 constellations) and M-QAM (4, 8, 16, 32 constellations) modulation with LDPC channel coding. The same system is extended for the various configuration of MIMO, i.e., 2x2, 3x3, and 4x4 channels. However, the assumed model is synthesized on Vertex-6 FPGA.

The fourth contribution is, based on the measure of NoC performance in terms of network delay and throughput. The NoC applications for future wireless communications

systems towards high BW, quality of service (QoS), low cost, speed of wireless access is studied and analysed.

The technological advancement in recent days has made digital communication society to explore the possibility of utilizing higher data transmission. Mostly, by using better source coding methods along with Error Corrections Codes (LDPC) and MIMO channel have indicated significant potential in the data communication system, due to the increase in channel capacity and quality of service. The current research work presents the source coding using PCA for data reduction over the MIMO channel for the data communication system. The evaluation of the system for efficiency and reliability shows the significance of data communication for low power communication.

Then outcomes of the work can be further utilized as a rule for NoC architects to settle on fitting decisions by considering the end goal to accomplish ideal execution for respective applications of future wireless communications systems to provide sensor data transmission high-data-rate, quality of service (QoS), low cost, speed of wireless access. MIMO wireless technology addresses these demands with spectral efficiency and improved reliability of the link.

6.2. References

[1] Prasad, N. Croitoru, V. Prasad, R. and Badoi, C. 5G Based Cognitive Radio. *Wireless Personal Communication Springer*, 57(2):441-465, 2013.

[2] Shukla, S. Khare, V. Garg, S. and Sharma, P. Comparative Study of 1G, 2G, 3G and 4G. *Journal of Engineering, Computers & Applied Sciences (JEC&AS)*, 2(4):55-63, 2013.

[3] S. E. Collier. The Emerging Internet: Convergence of the Smart Grid with the Internet of Things. *Rural Electric Power Conference (REPC), 2015 IEEE*, Asheville, NC, Pages 65-68, 2015.

[4] Sheikh, S. and Dakhore, M.H., 2015. Data compression techniques for wireless sensor network. *International Journal of Computer Science and Information Technologies (IJCSIT)*, 6(1):818-821, 2015.

[5] Mohit Lakra, Rajiv Dahiya and ShalooKikan. Performance Analysis of Low-Density Parity-Check Codes (LDPC) in Wimax OFDM System. *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, 2(7):2869-2874, 2013.