

Optical Multi-User MIMO for Multi-Mode Fiber Passive Optical Networks using Sub-Carrier Multiplexing

by

Bishal Poudel

Student ID Number: 1208006

A dissertation submitted to the
Photonic and Systems, Department of Engineering,
Graduate School of Engineering,
Kochi University of Technology,
Kochi, Japan

For the degree of
Doctor of Engineering

Assessment Committee:

Supervisor: Dr. Katsushi Iwashita

Co-Supervisor: Dr. Hirokazu Kobayashi

Co-Supervisor: Dr. Masanori Hamamura

Dr. Shuji Taue

Dr. Masahiro Fukumoto

September 2019

Abstract

Optical Multi-User MIMO for Multi-Mode Fiber Passive Optical Networks using Sub-Carrier Multiplexing

The exponential growth of global Internet Protocol (IP) traffic has triggered the demand for spectrally efficient high speed and high capacity optical network. After the maximum use of Time Division Multiplexing (TDM), Wavelength Division Multiplexing (WDM), Polarization Division Multiplexing (PDM) and multilevel modulation formats, a single strand of optical fiber have reached its transmission capacity limit. The only physical dimension of an optical fiber which is not fully utilized is the space. Mode division multiplexing (MDM) in a Multi-Mode Fiber (MMF) is a popular approach to exploit the space dimension of the fiber.

In an MMF there exist a non-linear multi-mode propagation due to random mode coupling. In a long-haul transmission, these fibers possess a challenge to compensate coupling between signals in different modes and to minimize the differential mode group delays. The crosstalk between the modes of MMFs limits the performance of the overall system, so Multiple-Input-Multiple-Output (MIMO) processing must be done to separate the received signals. The non-linearity effect is compensated using Deep learning technique. Deep learning neural networks are highly non-linear and capable of forming arbitrarily non-linear decision boundaries.

MIMO processing at the receiver-end is not suitable if all the end-users are located at a separate location. In this scenario, we need to design an optical transmission network that doesn't use MIMO processing at the receiver end. In existing Passive Optical Network (PON) each optical network unit receives not only the data that belongs to its user but also the data of other users. This makes the network vulnerable to sniffing even though encryption has been done. To solve this problem, we have proposed and demonstrated a mode forming technique in an optical network of Multi-Mode Fiber (MMF) using fused fiber coupler so that optical channels from an optical line terminal are switched directly to different user locations in such a way that the optical network unit receives only the data that belongs to its user. The use of pre-MIMO in the transmitter eliminates the need of a MIMO processor in the receiver side. We have successfully implemented a 2×2 mode forming network by transmitting two 100Mbps channels over a 1 km long Graded-

Index Multi-Mode Fiber (GI-MMF). This technique can be used to realize any $N \times N$ channel optical mode forming network to deliver them only to their destination.

Even if we assume MIMO processing is suitable to implement in the receiver side of MIMO-MDM system, none of the available MIMO detectors is an optimum detector. Optimum MIMO detector has always been a challenge in MIMO communication systems. We have designed a novel MIMO detector using a supervised Deep Learning Neural Network (DLNN) and has been implemented successfully in an MDM optical transmission system. A conventional GI-MMF is used to design an MDM optical transmission system. We have used a DLNN for MIMO detection in MDM optical transmission system and have compared its performance with Zero Forcing (ZF) detector and Semi-Definite Relaxation Row-by-Row (SDR-RBR). The results confirm that our DLNN outruns the performance of traditional MIMO detectors by compensating the non-linearity effect of multi-mode propagation.

The performance of the mode forming network depends on the feedback network. The use of pre-MIMO to eliminate the need of MIMO processor in the receiver raises questions about how accurately the channel matrix coefficients are transmitted back to the transmitter side. We have proposed and demonstrated a technique of signal extraction in the remote user location without the use of any kind of MIMO in the transmitter or receiver. We have successfully demonstrated the use of DLNN in MIMO-MDM optical transmission system for extracting the desired signal using only one composite signal. Two 1Gbps channels with sub-carrier multiplexing (SCM) has been successfully transmitted over a 1km long conventional GI-MMF and the desired signals are extracted in each remote user location using only one composite signal.

This work underlines the potential of Deep Learning technique for MIMO optical communication systems. It will inform future planned work to use Deep learning in order to enhance the BER performance of MIMO-MDM transmission system through joint system optimization.

Contents

Abstract	I
Contents	III
List of Figures	V
List of Tables	VII
List of Acronyms	VIII
List of Publications	X
Acknowledgements	XI
Chapter 1. Introduction	1
1.1 Background and problem statement.....	1
1.2 Research contribution.....	4
1.3 Outline of the dissertation.....	5
1.4 References.....	6
Chapter 2. Literature Review	10
2.1 MIMO-MDM.....	10
2.2 Non-linear propagation in multi-mode fibers.....	12
2.3 Design of a fused fiber coupler for mode multiplexing/demultiplexing.....	13
2.4 Deep learning in optical communication.....	14
2.5 Sub-carrier multiplexing in MDM transmission system.....	15
2.6 References.....	17
Chapter 3. Passive Optical Delivering Network using Conventional Graded-Index Multi-Mode Fiber with MDM and SCM	24
3.1 Mode forming network.....	24
3.2 Experimental setup.....	31
3.3 Results.....	33
3.4 Protection against sniffing.....	35
3.5 Discussion.....	36
3.6 Conclusion.....	37
3.7 References.....	38

Chapter 4. MIMO Detection using a Deep Learning Neural Network in a MDM Optical Transmission System.	41
4.1 DLNN in MDM optical transmission system	41
4.2 Design of DLNN	43
4.3 Training of DLNN.....	49
4.4 Experimental setup.....	50
4.5 Results	54
4.6 Excursion of channel condition number.....	55
4.7 Conclusion.....	56
4.8 References	56
Chapter 5. Mode Division Multiplexing MIMO Optical Transmission System that uses Deep Learning Neural Network for Signal Detection using only one Composite Signal	60
5.1 Proposed MDM optical network	60
5.2 Design of a deep learning neural network.....	61
5.3 Experimental setup.....	65
5.4 Results	67
5.5 Conclusion.....	69
5.6 References	69
Chapter 6. Summary and Future Research Direction	73
6.1 Summary	73
6.2 Further works	74

List of Figures

Fig 1.1: Different types of fibers, (a) SMF, (b) GI-MMF, and (c) MCF	2
Fig. 2.1: Schematic diagram of an MDM transmission system	10
Fig. 2.2: Multiple modes of FMF or MMF	10
Fig. 2.3: Mode-dependent fused fiber coupler for multiplexing.....	14
Fig. 2.4: SCM in MDM transmission system	16
Fig. 2.5: Illustration of SCM when 2 modes are used	17
Fig. 3.1: Proposed mode forming network	25
Fig. 3.2: Frame structure	26
Fig. 3.3: Signal receiving method.....	28
Fig. 3.4: 2×2 mode forming network	31
Fig. 3.5: Data format.....	32
Fig. 3.6: y_1 channel; (a) correlation before feedback, (b) correlation with continuous feedback, (c) eye diagram. y_2 channel; (d) correlation before feedback, (e) correlation with continuous feedback, (f) eye diagram.	34
Fig. 3.7: Output port exchange. y_1 channel; (a) correlation before feedback, (b) correlation with continuous feedback, (c) eye diagram. y_2 channel; (d) correlation before feedback, (e) correlation with continuous feedback, (f) eye diagram.....	35
Fig. 3.8: Weighting matrix. (a) magnitude. (b) phase.....	36
Fig. 3.9: Eye pattern when sniffing is done.	37
Fig.4.1: MDM optical transmission system that uses DLNN for MIMO detection.	42
Fig. 4.2: Signal processing technique on the i th channel.....	44
Fig. 4.3: Our deep learning neural network structure.	46
Fig. 4.4: A 2×2 MDM optical transmission system that uses DLNN for MIMO detection.....	50
Fig. 4.5: Confirms both the modes that are excited are received.....	51

Fig. 4.6: Performance of DLNN at a different number of hidden layers and neurons.	52
Fig. 4.7: Convergence behavior of DLNN: (a) accuracy versus epoch number, (b) mean square error versus epoch number.....	52
Fig. 4.8: Bit error rate versus channel condition number for our DLNN, ZF, and SDR-RBR.....	54
Fig. 4.9: The excursion of the channel condition number	55
Fig. 4.10: Performance of DLNN at different channel condition.	56
Fig. 5.1: Proposed MDM optical network	60
Fig. 5.2: Signal processing on i th channel	62
Fig. 5.3: Our deep learning neural network structure	64
Fig. 5.4: MDM optical transmission system for the desired signal extraction from one composite signal	65
Fig. 5.5: Both the excited modes are received	66
Fig. 5.6: Convergence behavior of DLNN: (a) accuracy versus epoch number, (b) mean square error versus epoch number.....	68
Fig. 5.7: Bit error rate versus channel condition number for our DLNN, and ZF	68

List of Tables

Table 3.1: weighting matrix coefficients	36
--	----

List of Acronyms

AMP	Approximate Message Passing
BER	Bit Error Rate
BPSK	Binary Phase Shift Keying
CN	Condition Number
CPU	Central Processing Unit
DFB-LD	Distributed Feedback Laser Diode
DL	Deep Learning
DLNN	Deep Learning Neural Network
DMGD	Differential-mode Group Delay
EVM	Error Vector Magnitude
FMF	Few-Mode Fiber
FM-MCF	Few-Mode Multi-Core Fiber
GI-MMF	Graded- Index Multi-Mode Fiber
IMDD	Intensity Modulation with a Direct-Detection
IP	Internet Protocol
LP	Linearly Polarized
MCF	Multi-Core Fiber
MD	modal dispersion
MDL	Mode Dependent Loss
MDM	Mode division multiplexing
ME	Manakov Equations
MIMO	Multiple-Input-Multiple-Output
ML	Maximum Likelihood
MMF	Multi-Mode Fiber
MMSE	Minimum Mean Square Error
MSE	Mean Square Error
MUX/DEMUX	mode multiplexer/de-multiplexer
MZM	Mach-Zehnder Modulator
NTT	Nippon Telegraph and Telephone
OCDM	Optical Code Division Multiplexing
OFDM	Orthogonal Frequency Division Multiplexing
ONU	Optical Network Unit
OOK	On-Off Keying
PDM	Polarization Division Multiplexing
PON	Passive Optical Network
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase Shift Keying
SCM	Sub-Carrier Multiplexing
SDM	Space Division Multiplexing
SDR	Semi-Definite Relaxation
SDR-RBR	Semi-Definite Relaxation Row-by-Row
SGD	Stochastic gradient descent
SMF	Single Mode Fiber
SNR	Signal to Noise Ratio

TDM	Time Division Multiplexing
TDMA	Time Division Multiple Access
WDM	Wavelength Division Multiplexing
ZF	Zero Forcing

List of Publications

The following publications and presentations are the parts of this work.

1. Bishal Poudel, Joji Oshima, Hirokazu Kobayashi, and Katsushi Iwashita, "MIMO detection using a deep learning neural network in a mode division multiplexing optical transmission system," *Optics Communications*, vol. 440, pp. 41-48 (2019).
2. Bishal Poudel, Joji Oshima, Hirokazu Kobayashi, and Katsushi Iwashita, "Passive Optical Delivering Network Using Conventional Graded-Index Multi-Mode Fiber With Mode Division Multiplexing and Sub-Carrier Multiplexing," *Journal of Optical Communications and Networking*, Vol. 10, Issue 3, pp. 252-259 (2018).
3. Bishal Poudel, Joji Oshima, Hirokazu Kobayashi, and Katsushi Iwashita, "Deep-learning neural network for MIMO detection in a mode-division multiplexed optical transmission system," presented at SPIE Photonics West 2019 Conference, San Francisco, California, United States, February 2019.
4. Bishal Poudel, Joji Oshima, Hirokazu Kobayashi, and Katsushi Iwashita, "Stability of an Optical Access Network that uses Mode Forming Technique over Conventional Graded-Index Multimode Fiber with Mode Division Multiplexing and Sub-carrier Multiplexing," presented at an ISFT Conference, Kochi University of Technology, Japan, November 2017.
5. Bishal Poudel, Joji Oshima, Yuki Morizumi, Hirokazu Kobayashi, and Katsushi Iwashita, "Passive Optical Delivering Network using Conventional Graded-Index Multimode Fiber with Mode Division Multiplexing and Sub-carrier Multiplexing," presented at the OECC Conference, Marina Bay, Singapore, August 2017.
6. Joji Oshima, Bishal Poudel, Hirokazu Kobayashi, Katsushi Iwashita, "Continuous Control in Multi-Mode Fiber Mode-Forming Networks," *Asia Communications and Photonics Conference*, Guangzhou, China, paper Su2A.37, November, 2017

Acknowledgements

First of all, I would like to express my sincere gratitude and appreciation to my supervisor Katsushi Iwashita for giving me the opportunity to work on this topic. Without his guidance and support, this would not have been possible. I feel immensely fortunate to work with him.

I would like to thank my co-supervisors Hirokazu Kobayashi and Masanori Hamamura for all the interesting discussions and advices. I am highly motivated by their wisdom, knowledge and enthusiasm for research.

I would like to appreciate all other members of my dissertation assessment committee: Shuji Taue, and Masahiro Fukumoto for their valuable suggestions and discussions.

A heartfelt thanks goes to all my current and previous members of Light Control and Optical Network Labs. Joji Oshima, I appreciate your kindness, support, and generosity.

To my family

Chapter 1. Introduction

1.1 Background and problem statement

The exponential growth of global Internet Protocol (IP) traffic has triggered the demand for spectrally efficient high speed and high capacity optical transmission systems [1]. After the maximum use of Time Division Multiplexing (TDM), Wavelength Division Multiplexing (WDM), Polarization Division Multiplexing (PDM) and multilevel modulation formats, a single strand of optical fiber have reached its transmission capacity limit [2]. The space dimension of an optical fiber is the only physical dimension which is not fully utilized [3]. Researchers have concentrated on Space Division Multiplexing (SDM) to further increase the optical fiber capacity and spectral efficiency. One approach of implementing SDM is by using Multi-Core Fiber (MCF) [4-6]. And another approach to implementing SDM is Mode Division Multiplexing (MDM) [7-9].

In MCF technique, multiple single-mode cores are used [10]. Although it is easy to migrate from single mode fiber systems to uncoupled single-mode MCF systems, the performance of the MCF optical transmission system is affected by the maximum achievable cladding diameters and inter-core crosstalk. This crosstalk depends on fiber perturbations and is stochastic [11-13]. Crosstalk limits how closely cores can be placed. In papers [14-15], various techniques to suppress inter-core crosstalk have been studied. MCFs with low crosstalk and/or large effective area have also been demonstrated [15-16]. In 2011, the first transmission exceeding 100Tbps was demonstrated using the first homogeneous, trench-assisted 7- core fiber [17]. In 2015, the capacity reaches 2.15Pbps using 22-core homogeneous single mode MCF and wideband optical comb [18].

Whereas in MDM technique, different modes of Few-Mode Fiber (FMF) or Multi-Mode Fiber (MMF) are used as unique data channels at the same wavelength to aid the fiber capacity [19]. FMF is a special class of MMF. They differ in the number of modes and available data channel. FMF supports up to a few tens of guided modes. In contrast, an MMF supports a few hundred modes. FMF was first developed by two NTT laboratories independently [20]. Some of the challenges of FMF system are; channel crosstalk due to mode coupling, pulse distortion due to large Differential-mode Group Delay (DMGD), and a MIMO processor to compensate the effect of crosstalk and DMGD. 283Tbps data transmission has been experimentally demonstrated over a 30Km span of FMF and the capacity of each mode of FMF has also been reported to be within

90% of the record data-rate reported for single-mode fibers [21]. By simultaneously using all the modes of FMF/MMF to transmit data, we can achieve the maximum capacity of the MDM system.

A combination of MCF and MDM approach is also investigated [22-23]. In a Few-Mode Multi-Core Fiber (FM-MCF) technology, the number of spatial multiplicities can be increased over 100 which helps to increase the transmission capacity per fiber [24]. Single Mode Fiber (SMF), Graded-Index Multi-Mode Fiber (GI-MMF) and MCF fiber structure is shown in Fig. 1.1. A record high transmission capacity of 10.16Pbps with an aggregate spectral efficiency of 1099.9b/s/Hz has been achieved using an FM-MCF with 6 modes in each of 19 cores [23].

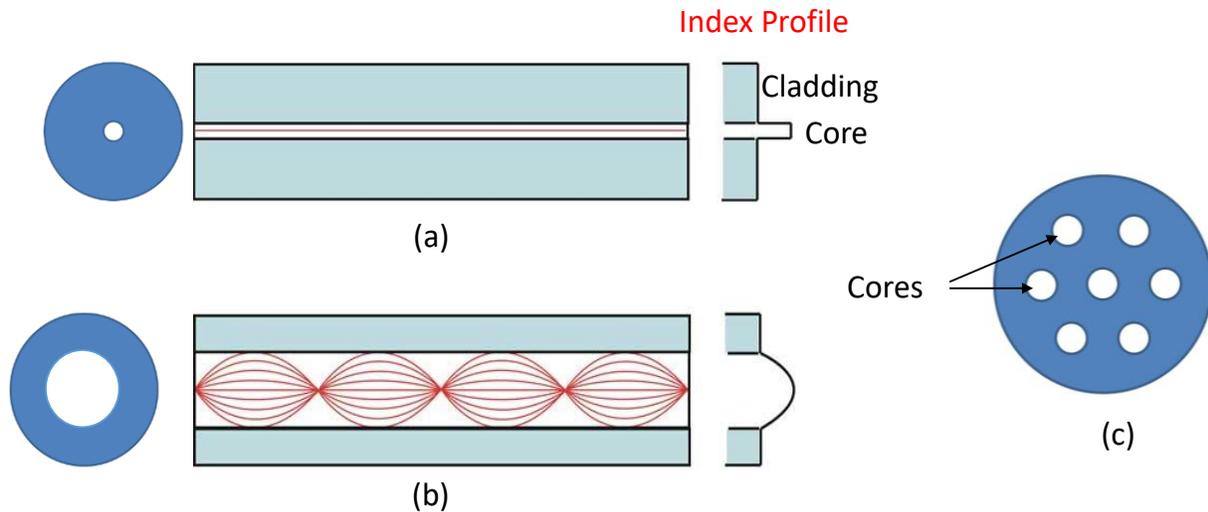


Fig 1.1: Different types of fibers, (a) SMF, (b) GI-MMF, and (c) MCF

One of the advantages of using FMF for MDM is that the fiber capacity can be increased without exponentially increasing Signal to Noise Ratio (SNR) [25-28]. FMF has better mode selectivity and easy mode impairments management compared to conventional MMF. However, all the existing optical networks use conventional MMFs and replacing all the existing networks with FMFs is not economically efficient. A conventional MMF supports many modes. There exists a non-linear multimode propagation in an MMF. In a long-haul transmission, these fibers possess a challenge to compensate coupling between signals in different modes and to minimize the DMGDs. The crosstalk between the modes of MMFs limits the performance of the overall system. So, Multiple-Input-Multiple-Output (MIMO) processing must be done to separate the received signals.

A conventional GI-MMF supports many modes, so if we use it to develop the MDM optical network, MIMO processing must be done to improve the performance. Common suboptimal MIMO detectors are Zero Forcing (ZF) detector [29] and Minimum Mean Square Error (MMSE) detector [30]. They are known for their low computational complexity because they use linear detection operation. ZF detector suffers from noise enhancement and has poor Bit Error Rate (BER) [29]. MMSE does not eliminate noise, but it minimizes. At high Signal to Noise Ratio (SNR), the MMSE detector converges to the ZF detector [29]. On the other hand, Maximum Likelihood (ML) uses a non-linear detection technique and provides an optimum MIMO detection only if all the transmitted symbol vectors are of equal probability. Also, the complexity of the ML detector increases exponentially with the size of MIMO and modulation order [31]. Some advanced detectors are Approximate Message Passing (AMP) [32], and Semi-Definite Relaxation (SDR) [33]. AMP is simpler and cheaper to implement in practice but requires proper knowledge of noise variance. An improper value of noise variance would severely degrade the performance. Furthermore, AMP detector may not converge when the MIMO channels are spatially correlated [34]. SDR is an alternative to ML detection because the complexity is polynomial rather than exponential. At high SNR, its performance is similar to the ML detector. Although SDR is more robust, it maintains a polynomial-time complexity with respect to the MIMO dimension [35].

This raises three questions:

1. If end-users are located at a separate location, then MIMO processing at the receiver-end is not suitable. To solve this problem, we need to design an optical transmission network that doesn't use MIMO processing at the receiver end.
2. In the existing optical transmission system, each Optical Network Unit (ONU) receives not only the data that belongs to its user but also the data of other users. So even if data is encrypted, the network still remains vulnerable to sniffing. It raises security questions. To solve this problem, we need to develop a new optical transmission system where each user receives only the designated optical channel.
3. Even if all the end-users are located at the same location, none of the available MIMO detectors is an optimum detector. Therefore, we need to develop a novel optimal MIMO detector using a supervised deep learning neural network (DLNN).

1.2 Research contribution

In today's Passive Optical Network (PON), each optical network unit receives not only the data that belongs to its user but also the data of other users. This makes the network vulnerable to sniffing even though encryption has been done. To solve this problem, we have successfully proposed and implemented a mode-forming technique in an optical network. In this approach, optical channels from an optical line terminal are switched directly to different user locations in such a way that the optical network unit receives only the data that belongs to its user. A mode forming technique is implemented in an optical network of MMF using a commercially viable fiber coupler to optically route only the desired channel to the desired destination. We have successfully implemented a 2×2 mode forming network and the total channel coefficients related to each user is estimated at their locations. The feasibility of the proposed technique has been confirmed using a conventional graded index Multimode fiber. The total channel matrix coefficients of this network are obtained by processing the transmitted training sequence in their user locations. Pre-Multiple Input Multiple Output (pre-MIMO) equalizer is implemented at the transmitter. The channels are successfully transmitted to their destinations clarifying wavelength independence while maintaining data security to some extent. Furthermore, the exchange of channel in the output port is also supported by this network.

Optimum MIMO detector has always been a challenge in MIMO communication systems. We have designed a novel MIMO detector using a supervised Deep Learning Neural Network (DLNN) and we have implemented it successfully in a MDM optical transmission system. A conventional GI-MMF is used to design an MDM optical transmission system. We have used a DLNN for MIMO detection in MDM optical transmission system and have compared its performance with Zero Forcing (ZF) detector and Semi-Definite Relaxation Row-by-Row (SDR-RBR). The results confirm that our DLNN outruns the performance of traditional MIMO detectors. A supervised Deep Learning neural network which is designed, trained and evaluated using a Keras library and TensorFlow is used for optimum MIMO detection.

A Mode-forming technique in an optical network of multi-mode fiber transmits only the data to the optical network unit that belongs to its user making the network resilient to sniffing. Implementing Deep Learning neural network for MIMO detection provides a new approach to optimum detection in a modern optical transmission system. When the end-users are located at a

different location, computationally intensive MIMO processing is not suitable. Therefore, we have purposed and successfully implemented a Deep Learning neural network for signal detection using only one received signal. We have successfully implemented this approach in a 2×2 MIMO MDM optical transmission system. The feasibility of the proposed optical transmission system has been confirmed using a conventional graded-index multi-mode fiber. A supervised Deep Learning neural network which is designed, trained and evaluated using a Keras library and TensorFlow is used for signal detection in each receiver location.

1.3 Outline of the dissertation

This dissertation is organized into six chapters.

Chapter 1. This chapter presents the research background and problem statement. The contributions of the research are also discussed here.

Chapter 2. This chapter includes a detailed review on describing the non-linear multimode propagation in multi-mode fiber using coupled Manakov equation, the design of a fused fiber coupler for mode division multiplexing/demultiplexing, and the application of deep learning in optical communication.

Chapter 3. In this chapter, a new optical mode forming network, which implements a feedback system, has been proposed and successfully implemented by using a mode division multiplexing in a conventional GI-MMF network using a commercially available mode dependent fused fiber coupler. Detailed information on the design of this network, its implementation technique, and its results are included.

Chapter 4. This chapter presents a new technique for designing a digital MIMO processor. A MIMO processor is designed using a deep learning technique and is implemented for MIMO detection in a mode division multiplexed optical transmission system. Detailed information on the design of a deep learning neural network and its characteristics are presented here. The performance of the newly designed neural network is also compared with the existing MIMO detector.

Chapter 5. This chapter presents a new technique for extracting the desired signal from one composite signal. This method is significant where MIMO processing is not suitable at the receiver

end. Here, a deep learning neural network is designed and implemented at the remote user location for the desired signal detection.

Chapter 6. Finally, a summary is presented in this chapter and some directions for future work are considered.

1.4 References

- [1] Cisco Visual Networking, “The Zettabyte Era—Trends and Analysis,” *Cisco white paper* (2017).
- [2] R. W. Tkach, “Scaling optical communications for the next decade and beyond,” *Bell Labs Tech. J.*, vol. 14, no. 4, pp. 3–9, 2010.
- [3] D. J. Richardson, J. M. Fini, and L. E. Nelson, “Space-division multiplexing in optical fibres,” *Nat. Photonics*, vol. 7, no. 5, pp. 354–362, 2013.
- [4] R. Ryf, R. Essiambre, A. Gnauck, S. Randel, M. A. Mestre, C. Schmidt, P. Winzer, R. Delbue, P. Pupalaikis, A. Sureka, T. Hayashi, T. Taru, and T. Sasaki, “Space-division multiplexed transmission over 4200 km 3-core microstructured fiber,” in *Proc. OFC/NFOEC (2012)*, paper PDP5C.2.
- [5] S. Chandrasekhar, A. Gnauck, X. Liu, P. Winzer, Y. Pan, E. C. Burrows, B. Zhu, T. Taunay, M. Fishteyn, M. Yan, J. M. Fini, E. Monberg, and F. Dimarcello, “WDM/SDM transmission of 10 x 128-Gb/s PDM-QPSK over 2688-km 7-core fiber with a per-fiber net aggregate spectral-efficiency distance product of 40,320 km.b/s/Hz,” in *Proc. ECOC (2011)*, paper Th.13.C.4.
- [6] H. Takahashi, T. Tsuritani, E. L. T. de Gabory, T. Ito, W. R. Peng, K. Igarashi, K. Takeshima, Y. Kawaguchi, I. Morita, Y. Tsuchida, Y. Mimura, K. Maeda, T. Saito, K. Watanabe, K. Imamura, R. Sugizaki, and M. Suzuki, “First demonstration of MC-EDFA-repeated SDM transmission of 40 x 128-Gbit/s PDM-QPSK signals per core over 6,160-km 7-core MCF,” *Opt. Express*, vol. 21, no. 1, pp. 789–795, 2013.
- [7] A. Al Amin, A. Li, X. Chen, and W. Shieh, “LP01/LP11 dual-mode and dual-polarization CO-OFDM transmission on two-mode fibre,” *Electron. Lett.*, vol. 47, no. 10, pp. 606–608, 2011.
- [8] R. Ryf, S. Randel, A. H. Gnauck, C. Bolle, A. Sierra, S. Mumtaz, M. Esmaelpour, E. C. Burrows, R. Essiambre, P. J. Winzer, D. W. Peckham, A. H. McCurdy, and R. Lingle, “Mode-

- division multiplexing over 96 km of fewmode fiber using coherent 6 x 6 MIMO Processing,” *J. Lightwave Technol.*, vol. 30, no. 4, pp. 521–531, 2012.
- [9] E. Ip, G. Milione, M. J. Li, N. Cvijetic, K. Kanonakis, J. Stone, G. Peng, X. Prieto, C. Montero, V. Moreno, and J. Liñares, “SDM transmission of real-time 10GbE traffic using commercial SFP + transceivers over 0.5km elliptical-core few-mode fiber,” *Opt. Express*, vol. 23, no. 13, pp. 17120–17126, 2015.
- [10] S. Iano, T. Sato, S. Sentsui, T. Kuroha, and Y. Nishimura, “Multicore optical fiber,” in *Optical Fiber Communication*, 1979 OSA Technical Digest Series (Optical Society of America, 1979), paper WB1.
- [11] T. Hayashi, T. Sasaki, E. Sasaoka, “Multi-core fibers and their crosstalk characteristics,” in *Proc. ECOC 2010*, Paper We.8.F.6.
- [12] J. M. Fini, B. Zhu, T. F. Taunay, and M. F. Yan, “Statistics of crosstalk in bent multicore fibers,” *Opt. Express*, vol.18, no. 14, pp. 15122–15129, 2010.
- [13] J. M. Fini, B. Zhu, T. F. Taunay, and M. F. Yan, and K. S. Abedin, “Crosstalk in multicore fibers with randomness: gradual drift vs. short-length variations,” *Opt. Express*, vol. 20, no. 2, pp. 949-959, 2012.
- [14] K. Imamura, Y. Tsuchida, K. Mukasa, R. Sugizaki, K. Saitoh and M. Koshiba, “Investigation on multi-core fibers with large A_{eff} and low micro bending loss,” *Opt. Express*, vol. 19, no. 11, pp. 10595-10603, 2011.
- [15] T. Hayashi, T. Taru, O. Shimakawa, T. Sasaki, and E. Sasaoka, “Design and fabrication of ultra-low crosstalk and low-loss multi-core fiber,” *Opt. Express*, vol. 19, no. 17, pp. 16576–16592, 2011.
- [16] K. Imamura, K. Mukasa, and R. Sugizaki, “Trench Assisted Multi-Core Fiber with Large A_{eff} over 100 μm^2 and Low Attenuation Loss,” in *Proc. ECOC2011*, Paper Mo.1.LeCervin.1.
- [17] J. Sakaguchi, Y. Awaji, N. Wada, A. Kanno, T. Kawanishi, T. Hayashi, T. Taru, T. Kobayashi, M. Watanabe, “109-Tb/s (7x97x172-Gb/s SDM/WDM/PDM) QPSK transmission through 16.8-km homogeneous multi-core fiber,” in *Proc. OFC2011*, PDPB6.

- [18] B.J. Puttnam, R.S. Luís, W. Klaus, J. Sakaguchi, J.M.D. Mendinueta, Y. Awaji, N. Wada, Y. Tamura, T. Hayashi, M. Hirano, J. Marciante, "2.15 Pb/s transmission using a 22 core homogeneous single-mode multi-core fiber and wideband optical comb," in *Proc. ECOC2015*, PDP 3-1.
- [19] S. Berdagué and P. Facq, "Mode division multiplexing in optical fibers," *Appl. Opt.*, vol. 21, no. 11, pp. 1950–1955, 1982.
- [20] J. Sakai, K. Kitayama, M. Ikeda, Y. Kato, K. Tatsuya, "Design Considerations of Broadband DualMode Optical Fibers," *IEEE Trans. Microwave Theory Technol.*, vol. 26, no. 9, pp. 658–65, 1978.
- [21] G. Rademacher, R. S. Luís, B. J. Puttnam, T. A. Eriksson, R. Ryf, E. Agrell, R. Maruyama, K. Aikawa, Y. Awaji, H. Furukawa, and N. Wada, "High Capacity Transmission With Few-Mode Fibers," *J. Lightwave Technol.*, vol. 37, no. 2, pp. 425-432, 2019
- [22] D. Qian, E. Ip, M. Huang, M. Li, A. Dogariu, S. Zhang, Y. Shao, Y. Huang, Y. Zhang, X. Cheng, Y. Tian, P. N. Ji, A. Collier, Y. Geng, J. Liñares, C. Montero, V. Moreno, X. Prieto, and T. Wang, "1.05 Pb/s transmission with 109 b/s/Hz spectral efficiency using hybrid single and few-mode cores," in *Proc. Frontiers in Optics*, Rochester, NY, USA, 2012, Paper FW6C.3.
- [23] D. Soma, Y. Wakayama, S. Beppu, S. Sumita, T. Tsuritani, T. Hayashi, T. Nagashima, M. Suzuki, M. Yoshida, K. Kasai, M. Nakazawa, H. Takahashi, K. Igarashi, I. Morita, and M. Suzuki, "Over 6-Mode 19-Core Fiber Across the C+L Band," *J. Lightwave Technol.*, vol. 36, no. 6, pp. 1362-1368, 2018
- [24] T. Sakamoto, T. Matsui, K. Saitoh, S. Saitoh, K. Takenaga, T. Mizuno, Y. Abe, K. Shibahara, Y. Tobita, S. Matsuo, K. Aikawa, S. Aozasa, K. Nakajima, and Y. Miyamoto, "Low-loss and Low-DMD Few-mode Multi-core Fiber with Highest Core Multiplicity Factor," in *Proc. OFC 2016*, paper Th5A.2
- [25] A. Li, A. Al Amin, X.Chen, and W.Shieh, "Transmission of 107Gb/s mode and polarization multiplexed CO-OFDM signal over a two-mode fiber," *Opt. Express*, vol. 19, no. 9, pp. 8808-8814, 2011

- [26] N.Hanzawa, K.Saitoh, T.Sakamoto, T.Matsui, S.Tomita, and M.Koshiba, "Demonstration of mode-division multiplexing transmission over 10km two-mode fiber with mode coupler," in *Proc. Optical Fiber Commun. Conf. (OFC/NFOEC 2011)*, Los Angeles, CA, paperOWA4.
- [27] M. Salsi, C. Koebele, D. Sperti, P. Tran, P. Brindel, H. Mardoyan, S. Bigo, A. Boutin, F. Verluise, P. Sillard, M. Bigot Astruc, L. Provost, F. Cerou, and G. Charlet, "Transmission at 2×100 Gb/s, over Two Modes of 40km-long Prototype Few-Mode Fiber, using LCOS based Mode Multiplexer and Demultiplexer," *Opt. Express*, vol. 19, no. 17, pp. 16593-16600, 2011.
- [28] R.Ryf, S.Randel, A.Gnauck, C.Bolle, R.Essiambre, P.Winzer, D.Peckham, A.McCurdy, and Lingle, "Space-division multiplexing over 10km of three-mode fiber using coherent 6×6 MIMO processing," in *Proc. Optical Fiber Commun. Conf. (OFC/NFOEC 2011)*, Los Angeles, CA, paperPDPB10.
- [29] X. Chen, J. Ye, Y. Xiao, A. Li, J. He, Q. Hu, and W. Shieh, Equalization of two-mode fiber based MIMO signals with larger receiver sets, *Opt. Express* vol. 20, no. 26, (2012) B413–B418.
- [30] L. Fang, L. Xu, and D. Huang, Low Complexity Iterative MMSE-PIC Detection for Medium-Size Massive MIMO, *IEEE Wireless Commun. Letters* vol. 5, no. 1, (2016) 108–111.
- [31] Z. Lei, Y. Dai, and S. Sun, Near Optimal List MIMO Detection, in *Proc. of IEEE conference on 16th International Symposium on Personal, Indoor and Mobile Radio Communications (IEEE, 2005)*, pp. 2441-2446
- [32] C. Jeon, R. Ghods, A. Maleki, and C. Studer, Optimality of large MIMO detection via approximate message passing, in *Proc. of IEEE conference on International Symposium on Information Theory (IEEE, 2015)*, pp. 1227-1231.
- [33] Z. Q. Luo, W. K. Ma, A. M. So, Y. Ye, and S. Zhang, Semidefinite relaxation of quadratic optimization problems, *IEEE Signal Process. Magaz.* vol. 27, no. 3, pp. 20-34, 2010
- [34] J. C. Chen, A Low Complexity Data Detection Algorithm for Uplink Mutiuser Massive MIMO Systems, *J. Selected Areas in Commun.* vol. 35, no. 8, pp. 1701-1714, 2017
- [35] W. K. Ma, C. C. Su, J. Jalden, and C. Y. Chi, Some results on 16-QAM MIMO detection using Semidefinite Relaxation, in *Proc. of IEEE International conference on Acoustics, Speech and Signal Processing (IEEE, 2008)*, pp. 2673-2676

Chapter 2. Literature Review

2.1 MIMO-MDM

Mode division multiplexing (MDM) is a technique of using spatial modes of FMF/MMF as a unique channel to carry independent data stream. Using MDM, the transmission capacity of an optical fiber increases by an amount proportional to the number of modes used [1]. An MDM transmission system consists of a mode multiplexer, an FMF/MMF, a mode de-multiplexer, multiple coherent receivers, and a MIMO processor [2-3]. The basic MDM transmission system is shown in Fig 2.1. The MMF supports several modes, so it is used as the transmission media of MDM system.

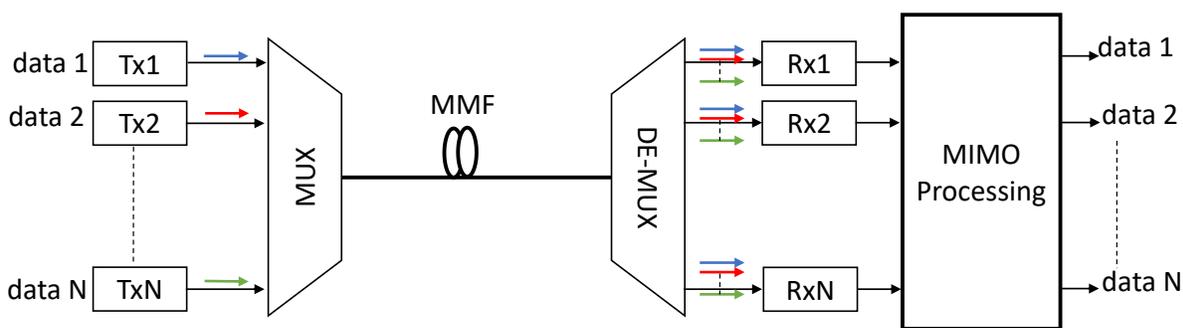


Fig. 2.1: Schematic diagram of an MDM transmission system

Using the concept of MIMO introduced by Foschini et al. [4], Stuart [5] tested the feasibility of MIMO over MMF in a 2×2 channel experiment. Many MDM transmission systems using FMFs have already been proposed [6-9]. Inside a FMF or MMF multiple modes exist as shown in Fig. 2.2.

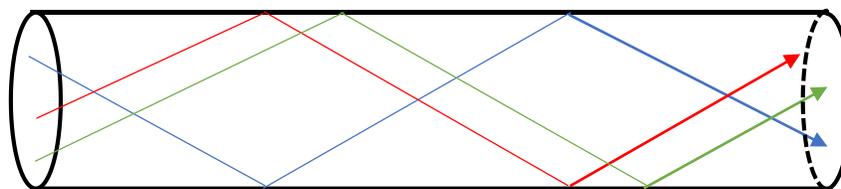


Fig. 2.2: Multiple modes of FMF or MMF

A guided mode is a solution of Maxwell's equation describing the distribution of electro-magnetic field along the fiber. Normalized frequency determines the number of guided modes in an optical fiber [10]. When the value of normalized frequency of a fiber is small then such fiber can only

guide fundamental mode LP01 as in SMF. As the value of normalized frequency exceeds 2.405, the fiber can support LP11 mode group. Guided modes are orthogonal, it means, if the normalized mode profile for LP01 is represented by Ψ_{01} and that of LP11 by Ψ_{11} then

$$\iint \Psi_{01} \cdot \Psi_{11} drd\theta = 0 \quad (1)$$

Equation 1 is the basis of mode-division multiplexing. Different modes in a MMF have different group velocities resulting in mode dependent group delays. The group velocity is defined as

$$v_g = (d\beta/d\omega)^{-1} \quad (2)$$

Where β denotes the propagation constant and ω denotes the frequency. Mode dependent group delay results in modal dispersion. Mode multiplexing (MUX) and De-multiplexing (DEMUX) is a critical operation in MDM transmission system. MUX transforms signals from multiple SMFs to MDM signals. Generally, a MUX performs two operations: mode conversion and combination. First, signals from SMFs are converted to the desired modes and then combined before coupling to MMF. Mode conversion can be achieved using various technique such as directional coupler [11], spatial phase modulator [12], phase plate [13], or fiber Bragg grating [14]. Low loss techniques for MUX/DEMUX are also available [15-16].

To compensate the crosstalk between the propagating modes of a fiber a MIMO Digital Signal Processor (DSP) is required at the receiver side. Optimum MIMO detector does not exist. Only near optimum MIMO detectors are available. In order to achieve a near-optimal detection, a dual data path architecture has been proposed in paper [17]. It uses a signal-vector based list detection method.

Some of the MIMO-less optical transmission system also exist [18-19]. In paper [18], two spatial modes carrying carrier-less amplitude/phase (CAP) modulation signals are transmitted over 1.1km Elliptical-Core Few Mode Fiber (EC-FMF). EC-FMF fibers provides better performance even under extreme bending conditions. Three spatial modes of EC-FMF can transmit 10Gbps data using MIMO-less transmission systems [20]. Even in FMF when the number of spatial modes is increased, modal crosstalk increases, then they need a MIMO processor. By using mode selective couplers for multiplexing/de-multiplexing two spatial modes are transmitted over 10-km long FMF without using a MIMO processor [21-22]. However, when the number of modes increases, the design of mode selective coupler becomes complex. Furthermore, as the name suggests, FMF

supports only few tens of modes as compared to MMF. For MIMO-less transmission, a special fiber must be designed, but the scalability of optical fiber design remains a challenging problem [23].

Recently, MIMO MDM in MMF has been gaining huge interest among researchers [24-27]. It is seen as a technique to increase data capacity [24-31]. The capacity can be increased by exploiting the usable modes of GI-MMF, but the crosstalk between high number of modes limits its capacity. Successful transmission of 6 spatial modes over a 17 km long GIMMF has been reported [32]. The results verify that the GI-MMF support scalable MDM. It also indicates that GI-MMF can support MDM transmission over 300 km. In papers [33-36], the first MIMO MDM experiments over standard MMFs were performed. Some of the challenges possessed by MDM are mode coupling and mixing, Differential Group-Delay (DGD) spread between the modes of the fiber, and differential attenuation of higher order modes nearer to the cladding index. Mode coupling and mixing occurs in optical components and MMFs. Various techniques are purposed to lower the DMGDs of GIMMFs for MIMO MDM [26-27]. Rescaling and optimizing the diameter of a GIMMF can lower the DMGDs for 6 to 36 spatial modes for MIMO MDM transmissions [35]. Optimization of a trench-assisted 50 μ m diameter GIMMF for MIMO MDM transmission at 1550nm has also been done [36]. It helps to selectively excite and detect the 20 lowest-order LP modes for optimum MIMO MDM transmission. It also increases the effective areas of fiber and lowers the DMGDs, fiber bend losses, attenuations of those LP modes.

2.2 Non-linear propagation in multi-mode fibers

In a fiber propagation, optical non-linearity and its effect on data transmission are of great concern. Strong random coupling occurs between modes that have similar wavenumbers. Whereas, weak coupling occurs between modes that have significantly different wavenumbers. Non-linearity in fiber is due to random mode coupling [37]. The non-linear multi-mode propagation is described in terms of coupled generalized multi-component Manakov Equations (ME) [38] as

$$\begin{aligned}\frac{\partial \vec{E}_a}{\partial z} &= i\beta_a \vec{E}_a - \beta_a' \frac{\partial \vec{E}_a}{\partial t} - i \frac{\beta_a''}{2} \frac{\partial^2 \vec{E}_a}{\partial t^2} + i\gamma(\kappa_{aa} |\vec{E}_a|^2 + \kappa_{ab} |\vec{E}_b|^2) \vec{E}_a. \\ \frac{\partial \vec{E}_b}{\partial z} &= i\beta_b \vec{E}_b - \beta_b' \frac{\partial \vec{E}_b}{\partial t} - i \frac{\beta_b''}{2} \frac{\partial^2 \vec{E}_b}{\partial t^2} + i\gamma(\kappa_{ba} |\vec{E}_a|^2 + \kappa_{bb} |\vec{E}_b|^2) \vec{E}_b.\end{aligned}\tag{3}$$

Where a and b represents the group of modes. \vec{E}_a and \vec{E}_b represents the vectors that describe the modes in the group a and b respectively. The group velocity dispersion is represented by $\beta_{a,b}''$. The group velocity is represented by $\beta_{a,b}'$. And the non-linear coupling coefficients by κ_{uv} [24]. The value of κ_{uv} is given as

$$\kappa_{uv} = \sum_{k,m} \sum_{j \in u} \sum_{h \in v} C_{jhkm} \frac{\xi \left[E_h^* E_k E_m E_j \right]}{|\vec{E}_u|^2 |\vec{E}_v|^2}. \quad (4)$$

where C_{jhkm} is the dimensionless constant. ξ represents statistical averaging. Depending on which of the coefficients is evaluated 'v' and 'u' takes the value 'b' and 'a'. If we set 'v=b' and 'u=a', then the indices 'h' and 'j' in the summation takes all the corresponding values in group 'b' and 'a' respectively. Other indices 'm' and 'k' run through all the modes.

2.3 Design of a fused fiber coupler for mode multiplexing/demultiplexing

The implementation technique of the mode-dependent fused fiber couplers for two channel is shown in Fig. 2.3. In the fused fiber coupler 1, Signal 1 with power $P1_{in}$ and Signal 2 with power $P2_{in}$ are coupled to the input ports. The blue arrow represents optical Signal 1 and the red arrow represents optical Signal 2. The fused fiber coupler consists of two closely placed parallel optical fiber that is stretched and fused together so that their cores are very close to each other. The length of the Coupling region, the separation of the cores, the wavelength of operation and other fiber parameters determine the coupling coefficients. The coupling coefficients quantify how efficiently power couple from one mode to another.

The two optical signals, signal 1 and signal2, are of the same wavelength, but they are propagating with different modes and different propagation constant. When these signals enter the coupling region of the fused fiber coupler, coupling begins. The strength of mode coupling between these two propagating modes is determined by the dimensionless ratio of the coupling coefficient (per unit length) to the difference between the propagation constant of these propagating modes. Then the output port of fused fiber coupler 1 is connected to the GI-MMF. The power of Signal 1 concentrate inside the dotted blue circle, whereas the power of Signal 2 concentrate inside the dotted red circle. The modes in the overlapping region of these circles carry both the signals. Only

by using the overlapping modes, the signals cannot be detected because the channel matrix

becomes $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ and its determinant becomes zero.

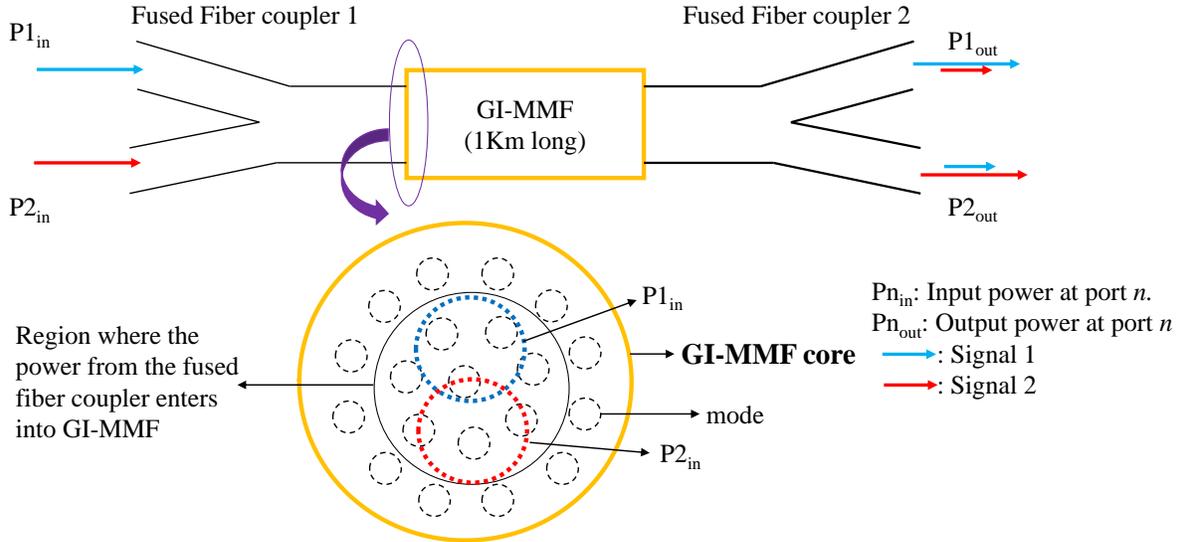


Fig. 2.3: Mode-dependent fused fiber coupler for multiplexing

During the detection process, we need a non-zero value of the determinant of the channel matrix. In the non-overlapped region of two dotted circles, multiple modes exist. Signal 1 and Signal 2 are propagated with multiple different modes.

2.4 Deep learning in optical communication

In recent years, the use of machine learning in communication systems is getting a lot of attention [39-40]. In an optical communication, machine learning has been used to monitor the performance, to mitigate fiber non-linearity, to recover the carrier, to recognize the modulation format [41-43]. In optical communication system, the major factors for limiting the data rate are chromatic dispersion and nonlinear Kerr effects [44]. So researchers are attracted to use an artificial neural network for channel equalization [45-49]. In paper [50], a Deep learning technique [51] has been used to design an artificial neural network to mitigate Nonlinear Interference by digital backpropagation [52]. An artificial neural network has also been used in PAM8 intensity modulation with a direct-detection (IMDD) system for the purpose of equalization [53]. Deep learning neural networks are also being used in the detection block of IMDD system for the mitigation of linear and non-linear impairments [54].

Existing communication systems are designed based on splitting the entire signal processing into a chain of multiple independent blocks. In such a system, each block performs a pre-defined function. These functions, such as source coding, channel coding, modulation, channel estimation, and equalization, are completed isolated from each other. This technique of individual block optimization does not provide the best end-end performance [55]. However, DLNN can do so even in complex communications scenarios that is difficult to represent using tractable mathematical models. The use of (Deep Learning) DL in MIMO communication systems jointly optimize the system [56]. In adaptive channel equalization, a training sequence is transmitted and the errors in the receiver side are calculated. Using these errors an optimum value for the relevant parameters of the equalizer are searched. However, most of the linear adaptive equalizer cannot efficiently equalize a non-linear channel and do not perform well on channels with deep spectral nulls. Non-linearity in fiber is due to intra- and intermodal nonlinear coupling. GI-MMF supports many modes, so there exist non-linear multimode propagation. Non-linear multi-mode propagation can be described in terms of coupled generalized multi-component Manakov equation [38]. This non-linearity can be addressed using Deep learning. Deep learning neural networks are highly non-linear and capable of forming arbitrarily non-linear decision boundaries.

2.5 Sub-carrier multiplexing in MDM transmission system

In optical SCM, multiple signals are multiplexed in the radiofrequency domain and they are transmitted using a light of single wavelength. The advantages of using SCM are that the subcarrier has an extremely longer wavelength than optical carrier; therefore, the effect of phase change in sub-carrier system is small compared to the optical carrier system. This makes sub carrier system more stable. The technique of implementing SCM in MDM transmission system is shown in Fig. 2.4. Here, in each transmitter, n independent high speed channels are mixed by n different radio frequencies f_i . After combining, these signals are modulated onto an optical carrier λ_1 . The optical carrier in each transmitter uses the same wavelength. This results in N modes multiplexed MDM signal at the output of mode multiplexer (MUX). This N mode signal propagate through a MMF.

The objective of using SCM is to generate a single sideband modulated signal with an optical carrier, as shown in fig. 2.5 so that self-heterodyne is possible by using the square-law detection technique. Figure 2.5 shows the modulation technique with 2 mode transmission. By using a radio frequency ω_m , two channels are first electrical modulated and then they are intensity modulated.

The spectrum after intensity modulation is shown in Fig. 2.5 (a), but after controlling the modulation current of Intensity modulator we get the spectrum which consists of only one sideband and an optical carrier as in Fig. 2.5 (b). This is the spectrum we have used in our research. This spectrum can self-heterodyne.

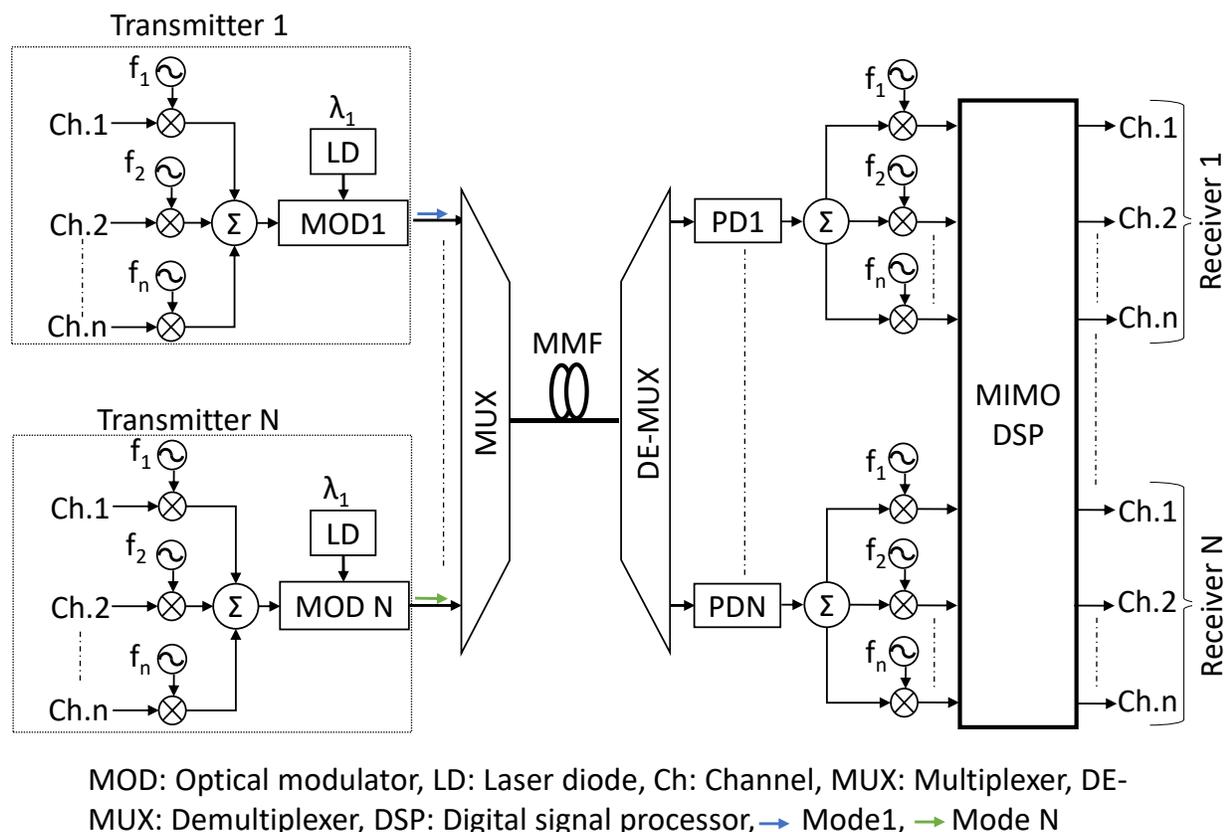
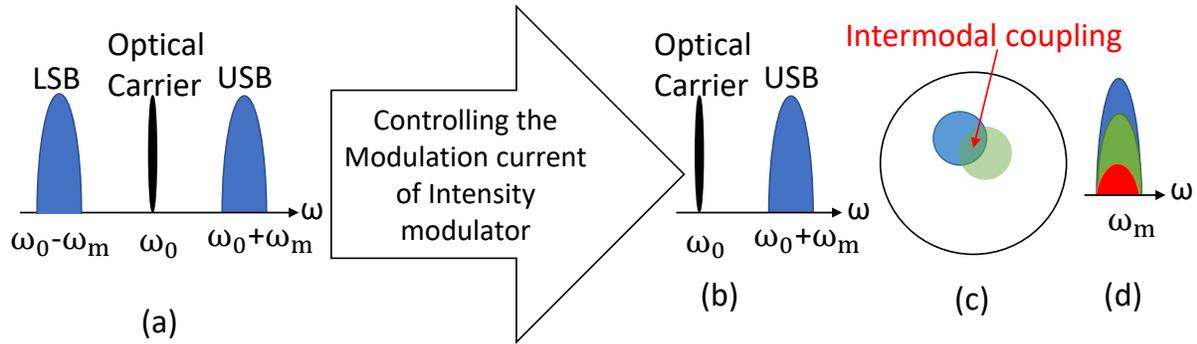


Fig. 2.4: SCM in MDM transmission system

Two data channels generates two SCM channels. Their optical spectrum consists of an optical carrier and the upper sideband of the modulated signal. The optical spectrum shown in Fig. 2.5 (b) is coupled into a GI-MMF. Figure 2.5 (b) is the optical spectrum of channel 1 signal. Similarly, channel 2 signal also has a similar spectrum. The two optical channels are coupled to two different modes of GI-MMF as in Fig. 2.5 (c). At the receiver, these signals are detected using photodiodes. The electrical spectrum of these two signals is shown in Fig. 2.5 (d).

In direct detection technique, the optical phase of the transmitted signal is lost after square-law photo-detection. However, if the SCM technique is performed, both the amplitude and phase information of the subcarrier, Fig 2.5 (d), can be recovered after direct detection from the beating

between optical carrier and subcarrier signal. Due to this, the phase information is preserved [57,58].



ω_0 : Optical carrier frequency, ω_m : RF sub-carrier frequency. ●: channel 1 signal, ●: channel 2 signal, ●: impairment.

Fig. 2.5: Illustration of SCM when 2 modes are used

The interference between signal-signal beat products and the desired signal-carrier beat terms is also called as ripple. When the spread spectrum technique is used along with SCM, noise enhancement can be compensated [57].

2.6 References

- [1] T. Morioka, Y. Awaji, R. Ryf, P. Wizer, D. Richardson, and F. Poletti, "Enhancing Optical Communications with Brand New Fibers," *IEEE Commun. Magazine*, S31-S42, Feb. 2012.
- [2] S. Randel, R. Ryf, A. Sierra, P. J. Winzer, A. H. Gnauck, C. A. Bolle, R. J. Essiambre, D. W. Peckham, A. McCurdy, and R. Lingle, Jr., "6×56-Gb/s mode-division multiplexed transmission over 33-km few-mode fiber enabled by 6×6 MIMO equalization," *Opt. Express* 19(17), 16697–16707 (2011).
- [3] R. Ryf, S. Randel, A. H. Gnauck, C. Bolle, A. Sierra, S. Mumtaz, M. Esmaeelpour, E. C. Burrows, R. Essiambre, P. J. Winzer, D. W. Peckham, A. H. McCurdy, and R. Lingle, Jr., "Mode-division multiplexing over 96 km of few-mode fiber using coherent 6×6 MIMO processing," *J. Lightwave Technol.* 30(4), 521–531 (2012).
- [4] G.J. Foschini and M.J. Gans, "On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas," *Wirel. Pers. Comm.*, vol. 6, pp. 311–335, Aug. 1998.

- [5] H. R. Stuart, "Dispersive Multiplexing in Multimode Optical Fiber," *Science*, vol. 289, no. 5477, pp. 281-283, Jul. 2000
- [6] N.Hanzawa, K.Saitoh, T.Sakamoto, T.Matsui, S.Tomita, and M.Koshiha, "Demonstration of mode-division multiplexing transmission over 10 km two-mode fiber with mode coupler," OFC/NFOEC, Los Angeles, CA, USA, 2011, Paper OWA4.
- [7] M. Salsi, C. Koebele, G. Charlet, and S. Bigo, "Mode division multiplexed transmission with a weakly-coupled few-mode fiber," OFC/NFOEC, Los Angeles, CA, USA, 2012, Paper OTU2C.5.
- [8] R. Ryf, S. Randel, N. K. Fontaine, M. Montoliu, E. Burrows, S. Corteslli, S. Chandrasekhar, A. H. Gnauck, C. Xie, R.-J.Essiambre, P. J. Winzer, R. Delbue, P. Pupalaikis, A. Sureka, Y. Sun, L. Gruner-Nielsen, R.V.Jensen, and R.Lingle, "32-bit/s/Hz spectral efficiency WDM transmission over 177-km few-mode-fiber, OFC/NFOEC, Anaheim, CA, USA, 2013, Paper PDP5A.1.
- [9] E. Ip, M. J. Li, K. Bennett, Y. K. Huang, A. Tanaka, A. Korolev, K. Koreshkov, W. Wood, E. Mateo, J. Hu, and Y. Yano, "146×6×19-Gbaudwavelength-and-mode-division multiplexed transmission over 10×50-km spans of few-mode fiber with a gain-equalized few-mode EDFA," OFC/NFOEC, Anaheim, CA, USA, 2013, Paper PDP5A.2.
- [10] D. Gloge, "Weakly Guiding Fibers," *Appl. Opt.*, vol. 10, pp. 2252-2258, 1971
- [11] B. Huang, C. Xia, G. Matz, N. Bai, and G. Li, "Structured Directional Coupler Pair for Multiplexing of Degenerate Modes," in *Optical Fiber Communication Conference/National Fiber Optic Engineers Conference*, 2013, p. JW2A.25.
- [12] C. Koebele, M. Salsi, D. Sperti, P. Tran, P. Brindel, H. Mardoyan, et al., "Two mode transmission at 2x100Gb/s, over 40km-long prototype few-mode fiber, using LCOSbased programmable mode multiplexer and demultiplexer," *Opt. Express*, vol. 19, pp. 16593-16600, 2011
- [13] R. Ryf, S. Randel, A. H. Gnauck, C. Bolle, R.-J. Essiambre, P. Winzer, et al., "Space division multiplexing over 10 km of three-mode fiber using coherent 6 × 6 MIMO processing," in *Optical Fiber Communication Conference/National Fiber Optic Engineers Conference*, 2011, p. PDPB10.

- [14] A. Li, A. Al Amin, X. Chen, and W. Shieh, "Reception of Mode and Polarization Multiplexed 107-Gb/s CO-OFDM Signal over a Two-Mode Fiber," in Optical Fiber Communication Conference/National Fiber Optic Engineers Conference, 2011, p. PDPB8.
- [15] R. Ryf, M. A. Mestre, A. Gnauck, S. Randel, C. Schmidt, R. Essiambre, et al., "LowLoss Mode Coupler for Mode-Multiplexed transmission in Few-Mode Fiber," in Optical Fiber Communication Conference/National Fiber Optic Engineers Conference, 2012, p. PDP5B.5.
- [16] N. K. Fontaine, R. Ryf, S. G. Leon-Saval, and J. Bland-Hawthorn, "Evaluation of Photonic Lanterns for Lossless Mode-Multiplexing," in Optical Communication (ECOC), European Conference and Exhibition on, 2012, p. Th.2.D.6.
- [17] G. Lee, H. Yun, T. Kim, "A near-optimal detector for spatial modulation MIMO systems," 2016 IEEE International Symposium on Circuits and Systems (ISCAS), Montreal, QC, Canada, 2016.
- [18] Wang L, Ai J, Zhu L, Wang A, Fu S, Du C, Mo Q, Wang J, "MDM transmission of CAP-16 signals over 1.1- km anti-bending trench-assisted elliptical-core few-mode fiber in passive optical networks," *Optics Express*, vol. 25, no. 19, pp. 22991–23002, Sept. 2017
- [19] Z. Wu, J. Li, D. Ge, F. Ren, P. Zhu, Q. Mo, Z. Li, Z. Chen, and Y. He, "Demonstration of all-optical MDM/WDM switching for short-reach networks," *Optics Express*, vol. 24, no. 19, pp. 21609–21618, Sept. 2016
- [20] F. Parmigiani, Y. Jung, L. Grüner-Nielsen, T. Geisler, P. Petropoulos, and D. J. Richardson, "MIMO-less Space Division Multiplexing Transmission over 1 km Elliptical Core Few-Mode Fiber," in Proc. at CLEO: Science and Innovations 2017, San Jose, California United States, 14–19 May 2017, paper SW11.1
- [21] R. Fang, L. Juhao, W. Zhongying, H. Tao, Y. Jinyi, M. Qi, H. Yongqi, C. Zhangyuan and L. Zhengbin, "Three-mode mode-division-multiplexing passive optical network over 12-km low mode-crosstalk FMF using all-fiber mode MUX/DEMUX," *Optics Comm.*, vol. 383, pp. 525–530, Jan. 2017.
- [22] J. Li, F. Ren, T. Hu, Z. Li, Y. He, Z. Chen, Q. Mo and G. Li, "Recent progress in mode-division multiplexed passive optical networks with low modal crosstalk," *Optical Fiber Technology*, vol. 35, pp. 28–36, Feb. 2017.

- [23] S. LaRochelle, A. Corsi, R. M. Nejad, L. Wang, X. Guan, J.-H. Chang, J. Lin, and L. A. Rusch, "Optical fiber designs for MIMO-less SDM," in Proc. at Conference on Lasers and Electro-Optics (CLEO), San Jose, CA, USA, May 2018
- [24] S. O. Arik, J. M. Kahn, K. Ho, "MIMO Signal Processing for Mode-Division Multiplexing: An overview of channel models and signal processing architectures," *IEEE Signal Processing Magazine*, vol. 31, no. 2, 2014
- [25] J. M. Kahn, S. Ö. Arik, E. L. Ginzton, K. Po. Ho, "MIMO channel statistics and signal processing in mode-division multiplexing systems," in proc. 2015 IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Stockholm, Sweden
- [26] P. Sillard, D. Molin, M. Bigot-Astruc, K. de Jongh, F. Achten, "Multimode Fibers for Mode Division Multiplexing," in proc. CLEO 2017, San Jose, CA, USA
- [27] P. Sillard, D. Molin, M. Bigot-Astruc, A. Amezcua-Correa, K. de Jongh, and F. Achten, "50 μm Multimode Fibers for Mode Division Multiplexing," *Journal of Lightwave Technology*, vol. 34, no. 8, pp. 1672 - 1677, 2016
- [28] H. R. Stuart, "Dispersive multiplexing in multimode fiber," presented at the Optical Fiber Communication Conf., Baltimore, MD, USA, 2000, Paper ThV2.
- [29] C. P. Tsekrekos, A. Martinez, F. M. Huijskens, and A. M. J. Koonen, "Design consideration for a transparent mode group diversity multiplexing link," *IEEE Photon. Technol. Lett.*, vol. 18, no. 22, pp. 2359–2361, Nov. 2006.
- [30] S. Schollmann and W. Rosenkranz, "Experimental equalization of crosstalk in 2×2 MIMO system based on mode group diversity multiplexing in MMF systems @10.7 Gbps," presented at the 33rd European Conf. Optical Communication, Berlin, Germany, 2007, Paper 7.4.2.
- [31] B. Franz, D. Suikat, R. Dischler, F. Buchali, and H. Buelow, "High speed OFDM transmission over 5 km of GI-Multimode fiber using spatial multiplexing with 2×2 MIMO," presented at the 36th European Conf. Exhibition Optical Communication, Torino, Italy, 2010, Paper Tu.3.C.4.
- [32] R. Ryf, N. K. Fontaine, H. Chen, B. Guan, B. Huang, M. Esmaelpour, A. H. Gnauck, S. Randel, S.J.B. Yoo, A.M.J. Koonen, R. Shubochkin, Y. Sun, and R. Lingle, "Mode-

multiplexed transmission over conventional graded-index multimode fibers," *Optics Express* Vol. 23, n. 1, pp. 235-246, Jan 2015

- [33] H. R. Stuart, "Dispersive multiplexing in multimode fiber," in *Proc. Optical Fiber Communication Conf.*, Baltimore, MD, USA, 2000, Paper ThV2.
- [34] C. P. Tsekrekos, A. Martinez, F. M. Huijskens, and A. M. J. Koonen, "Design consideration for a transparent mode group diversity multiplexing link," *IEEE Photon. Technol. Lett.* vol. 18, no. 22, pp. 2359–2361, 2006.
- [35] S. Schollmann and W. Rosenkranz, "Experimental equalization of " crosstalk in 2×2 MIMO system based on mode group diversity multiplexing in MMF systems @10.7 Gbps," in *Proc. 33rd European Conf. Optical Communication*, Berlin, Germany, 2007, Paper 7.4.2.
- [36] P. Sillard, D. Molin, M. Bigot-Astruc, A. Amezcua-Correa, K. d. Jongh, and F. Achten, "50 μ m multimode fibers for mode division multiplexing," in *Proc. European Conference on Optical Communication (ECOC)*, Valencia, Spain, Oct. 2015.
- [37] A. Mecozzi, C. Antonelli, and M. Shtaif, "Coupled Manakov equations in multimode fibers with strongly coupled groups of modes," *Opt. Express*, vol. 20, no. 21, pp. 23436-23441 (2012)
- [38] S. V. Manakov, "On the theory of two-dimensional stationary self-focusing of electromagnetic waves," *Sov. Phys. JETP* 38, 248–253 (1974).
- [39] C. Jiang, H. Zhang, Y. Ren, Z. Han, K.-C. Chen, and L. Hanzo, "Machine learning paradigms for next-generation wireless networks," *IEEE Wireless Commun.*, vol. 24, no. 2, pp. 98–105, 2017.
- [40] F. Khan, C. Lu, and A. Lau, "Machine learning methods for optical communication systems," in *Proc. Adv. Photon. IPR, NOMA, Sensors, Netw., SPPCom, PS, OSA Tech. Dig.*, 2017, Paper SpW2F.3.
- [41] J. Thrane, J. Wass, M. Piels, J. C. M. Diniz, R. Jones, and D. Zibar, "Machine learning techniques for optical performance monitoring from directly detected PDM-QAM signals," *J. Lightw. Technol.*, vol. 35, no. 4, pp. 868–875, 2017.
- [42] D. Zibar, M. Piels, R. Jones, and C. Schäffer, "Machine learning techniques in optical communication," *J. Lightw. Technol.*, vol. 34, no. 6, pp. 1442–1452, 2016.

- [43] F. Khan, Y. Zhou, A. Lau, and C. Lu, “Modulation format identification in heterogeneous fiber-optic networks using artificial neural networks,” *Opt. Express*, vol. 20, no. 11, pp. 12422–12431, 2012.
- [44] R.-J. Essiambre, G. Kramer, P. Winzer, G. Foschini, and B. Goebel, “Capacity limits of optical fiber networks,” *J. Lightw. Technol.*, vol. 28, no. 4, pp. 662–701, 2010.
- [45] K. Burse, R. Yadav, and S. Shrivastava, “Channel equalization using neural networks: A review,” *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 40, no. 3, pp. 352–357, 2010.
- [46] M. Ibnkahla, “Applications of neural networks to digital communications—A survey,” *Elsevier Signal Process.*, vol. 80, no. 7, pp. 1185–1215, 2000.
- [47] M. Jarajreh et al., “Artificial neural network nonlinear equalizer for coherent optical OFDM,” *IEEE Photon. Technol. Lett.*, vol. 27, no. 4, pp. 387–390, 2014.
- [48] E. Giacomidis et al., “Fiber nonlinearity-induced penalty reduction in CO-OFDM by ANN-based nonlinear equalization,” *Opt. Lett.*, vol. 40, no. 21, pp. 5113–5116, 2015.
- [49] T. Eriksson, H. Bülow, and A. Leven, “Applying neural networks in optical communication systems: Possible pitfalls,” *IEEE Photon. Technol. Lett.*, vol. 29, no. 23, pp. 2091–2094, 2017.
- [50] C. Häger and H. Pfister, “Nonlinear interference mitigation via deep neural networks,” in *Proc. Opt. Fiber Commun. Conf., OSA Tech. Dig.*, 2018, Paper W3A.4.
- [51] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
- [52] E. Ip and J. Kahn, “Compensation of dispersion and nonlinear impairments using digital backpropagation,” *J. Lightw. Technol.*, vol. 26, no. 20, pp. 3416–3425, 2008.
- [53] S. Gaiarinet et al., “High speed PAM-8 optical interconnects with digital equalization based on neural network,” in *Proc. Asia Commun. Photon. Conf.*, 2016, Paper AS1C-1.
- [54] J. Estarànet et al., “Artificial neural networks for linear and non-linear impairment mitigation in high-baudrate IM/DD systems,” in *Proc. 42nd Eur. Conf. Opt. Commun.*, 2016, pp. 106–108.
- [55] T. O’Shea and J. Hoydis, *An Introduction to Deep learning for the Physical Layer*, *IEEE Transactions on Cognitive Commun. and Netw.* vol. 3, no. 4, (2017) 563-575.
- [56] T. O’Shea, T. Erpek, and T. C. Clancy, *Physical Layer Deep Learning of Encodings for the MIMO Fading Channel*, in *Proceedings of 55th Annual Allerton Conference on*

Communication, Control, and Computing, N. Hovakimyan and N. Kiyavash (Allerton, 2017) 76-80.

- [57] Yang Zhang, Junya Matsushita, Tomoyasu Nishimori, Katsushi Iwashita “SCM-SS scheme for optical MIMO transmission using multimode fibers,” IEICE Communications Express, vol.2, no. 6, pp.268-273, 2013
- [58] Z. Li, M. S. Erkılınç, S. Pachnicke, H. Griesser, R. Bouziane, B. C. Thomsen, P. Bayvel, and R. I. Killey, "Signal-signal beat interference cancellation in spectrally-efficient WDM direct-detection Nyquist-pulse-shaped 16-QAM subcarrier modulation," Optics Express, Vol. 23, No. 18, pp. 23694-23709, 2015

Chapter 3. Passive Optical Delivering Network using Conventional Graded-Index Multi-Mode Fiber with MDM and SCM

3.1 Mode forming network

The main objective of our proposed mode forming network is to deliver the desired signal only to the desired destination, making the delivering process independent of wavelength, data rate, and modulation format.

In wireless communications, beamforming technique is used to focus the beam to the desired area by controlling the transmitted signals. We have proposed to apply the same beamforming concept to MMF systems using sub-carrier signals instead of the optical carrier. This mode forming network consists of transmitters, MMF or few-mode fiber, mode dependent fiber couplers and receivers. Multiplexing and de-multiplexing are achieved through the use of fused fiber couplers. The transmitter (Tx_n) transmits a superimposed data. These signals propagate through a fiber experiencing a different delay with the different mode. By controlling the amplitude and the phase of the transmitted signal, only the desired signal is received by the receiver (Rx_n) whereas other undesired signals get canceled being out of phase. The proposed N channel mode forming network with a feedback system is shown in Fig. 3.1. Here, \mathbf{H} is the channel matrix of the MMF and $\mathbf{H}_{ov} = \mathbf{H}\mathbf{W}$ is the total channel matrix including the weighting network. This system uses SCM. The advantages of using SCM are that the subcarrier has an extremely longer wavelength than optical carrier; therefore, the effect of phase change in sub-carrier system is small compared to the optical carrier system. This makes sub carrier system more stable. These channels $d_1, d_2 \dots d_N$ are superimposed to generate N transmitter signals $x_1, x_2 \dots x_N$ and are transmitted through an MMF. By using another fused fiber coupler, these transmitter signals are divided and are taken to N destinations that are separately placed. As these optical signals are at different locations, MIMO processing at the remote user cannot be implemented either to get the channel matrix or to separate the channels. Therefore, a new approach has been used to calculate channel matrix.

In this approach, signal processing is done independently on all receivers at a remote location to calculate their respective total channel matrix coefficients. These total channel coefficients are feedback to the transmitter in the uplink using a multiple access technique such as Time Division

Multiple Access (TDMA). Weight controller computes the channel matrix and controls the weighting network. The coefficients of the channel matrix are calculated as follows.

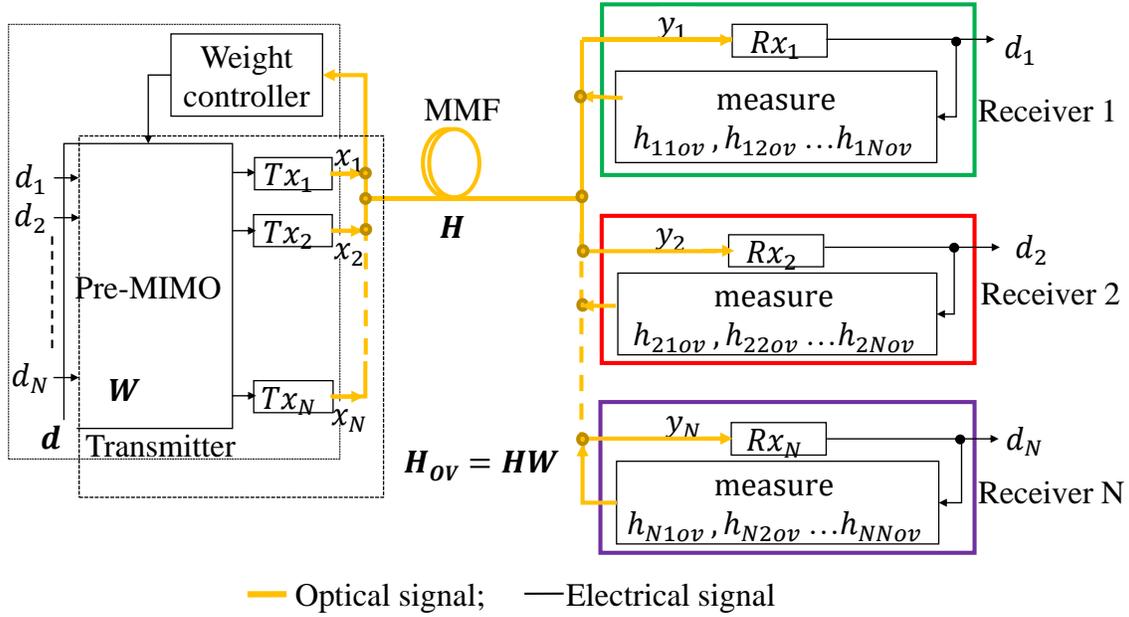


Fig. 3.1: Proposed mode forming network

The frame is shown in Fig. 3.2. It consists of training sequence and data. The training sequence is used to acquire the total channel matrix coefficients for each terminal. During the training sequence transmission, when $d_1 = '1'$, all other channels have '0', then d_1 is available on all output ports. By using these received signals, total channel matrix coefficients associated with the channel d_1 are calculated. Similarly, total channel matrix coefficients associated with the channel d_2 are calculated when, $d_2 = '1'$ and all other channels have '0'. This process is repeated to calculate all the total channel matrix coefficients. To implement this principle a training sequence is used.

The characteristics of the mode forming network are described as follows.

If $y_i(t)$ represents the signal received by the i th receiver, the total received signal

$$\mathbf{y}(t) = (y_1(t) \ y_2(t) \ \dots \ y_N(t))^T \text{ is given as:}$$

$$\begin{aligned}
\mathbf{y}(t) &= \mathbf{H}\mathbf{W}\mathbf{d}(t) = \mathbf{H}_{ov}\mathbf{d}(t) \\
&= \begin{pmatrix} h_{11} & h_{12} & \cdots & h_{1N} \\ h_{21} & h_{22} & \cdots & h_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N1} & h_{N2} & \cdots & h_{NN} \end{pmatrix} \begin{pmatrix} w_{11} & w_{12} & \cdots & w_{1N} \\ w_{21} & w_{22} & \cdots & w_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ w_{N1} & w_{N2} & \cdots & w_{NN} \end{pmatrix} \begin{pmatrix} \mathbf{d}_1(t) \\ \mathbf{d}_2(t) \\ \vdots \\ \mathbf{d}_N(t) \end{pmatrix}, \quad (1) \\
&= \begin{pmatrix} (h_{11}w_{11} + \cdots h_{1N}w_{N1})\mathbf{d}_1(t) + \cdots (h_{11}w_{1N} + \cdots h_{1N}w_{NN})\mathbf{d}_N(t) \\ (h_{21}w_{11} + \cdots h_{2N}w_{N1})\mathbf{d}_1(t) + \cdots (h_{21}w_{1N} + \cdots h_{2N}w_{NN})\mathbf{d}_N(t) \\ \vdots \\ (h_{N1}w_{11} + \cdots h_{NN}w_{N1})\mathbf{d}_1(t) + \cdots (h_{N1}w_{1N} + \cdots h_{NN}w_{NN})\mathbf{d}_N(t) \end{pmatrix}
\end{aligned}$$

where, \mathbf{H} is the channel matrix, \mathbf{H}_{ov} is the total channel matrix, $\mathbf{d}(t)$ is the input signal matrix and \mathbf{W} is the weighting network matrix.

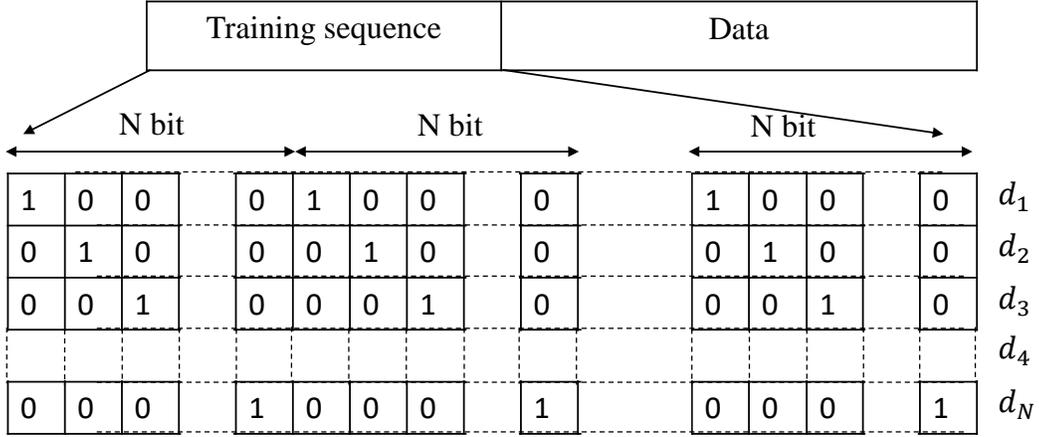


Fig. 3.2: Frame structure

The weighting network matrix is defined as $\mathbf{W} = (\mathbf{w}_1 \ \mathbf{w}_2 \ \cdots \ \mathbf{w}_N)$ where

$$\mathbf{w}_k = (w_{1k} \ w_{2k} \ \cdots \ w_{Nk})^T.$$

The total channel matrix is defined as:

$$\mathbf{H}_{ov} = \mathbf{H}\mathbf{W} = \begin{pmatrix} h_{11ov} & h_{12ov} & \cdots & h_{1Nov} \\ h_{21ov} & h_{22ov} & \cdots & h_{2Nov} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N1ov} & h_{N2ov} & \cdots & h_{NNov} \end{pmatrix}, \quad (2)$$

The training sequence of length M is used to obtain the coefficients of the weighting network. Let

we take the first N bits of the training sequence to compute the coefficients of the total channel matrix.

When only $d_1(T)$ is transmitted, where T is the bit period, the input signal matrix is

$$\mathbf{d}(T) = \begin{pmatrix} d_1(T) \\ d_2(T) \\ \vdots \\ d_N(T) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \text{ Then the received signal is:}$$

$$\mathbf{y}(T) = \begin{pmatrix} y_1(T) \\ y_2(T) \\ \vdots \\ y_N(T) \end{pmatrix} = \begin{pmatrix} h_{11}w_{11} + h_{12}w_{21} + \dots + h_{1N}w_{N1} \\ h_{21}w_{11} + h_{22}w_{21} + \dots + h_{2N}w_{N1} \\ \vdots \\ h_{N1}w_{11} + h_{N2}w_{21} + \dots + h_{NN}w_{N1} \end{pmatrix} = \mathbf{H}\mathbf{w}_1, \quad (3)$$

When only $d_2(2T)$ is transmitted, the input signal matrix is $\mathbf{d}(2T) = \begin{pmatrix} d_1(2T) \\ d_2(2T) \\ \vdots \\ d_N(2T) \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}$. Then the

received signal is:

$$\mathbf{y}(2T) = \begin{pmatrix} y_1(2T) \\ y_2(2T) \\ \vdots \\ y_N(2T) \end{pmatrix} = \begin{pmatrix} h_{11}w_{12} + h_{12}w_{22} + \dots + h_{1N}w_{N2} \\ h_{21}w_{12} + h_{22}w_{22} + \dots + h_{2N}w_{N2} \\ \vdots \\ h_{N1}w_{12} + h_{N2}w_{22} + \dots + h_{NN}w_{N2} \end{pmatrix} = \mathbf{H}\mathbf{w}_2, \quad (4)$$

In general, when only $d_i(iT)$ is transmitted the received signal is

$$\mathbf{y}(iT) = (y_1(iT) \ y_2(iT) \ \dots \ y_N(iT))^T.$$

Similarly, all other remaining equations are obtained and the total channel matrix of the overall network, \mathbf{H}_{ov} , is computed by using the Eq. (1). These total channel matrix coefficients are uplinked to the weight controller. The weight controller calculates the channel matrix as:

$$\mathbf{H} = \mathbf{H}_{ov}\mathbf{W}^{-1}, \quad (5)$$

If we want to send $d_i(t)$ only to $y_i(t)$, for $1 \leq i \leq N$ then all other branch signals must be set to zero, so we obtain next relation from Fig. 3.3 as:

$$\sum_{k=1}^N h_{ik} w_{kj} = 0, \text{ for } i \neq j \text{ and } 1 \leq i, j \leq N, \quad (6)$$

When channel matrix coefficients are obtained by using Eq. (5), we can determine the weighting network value.

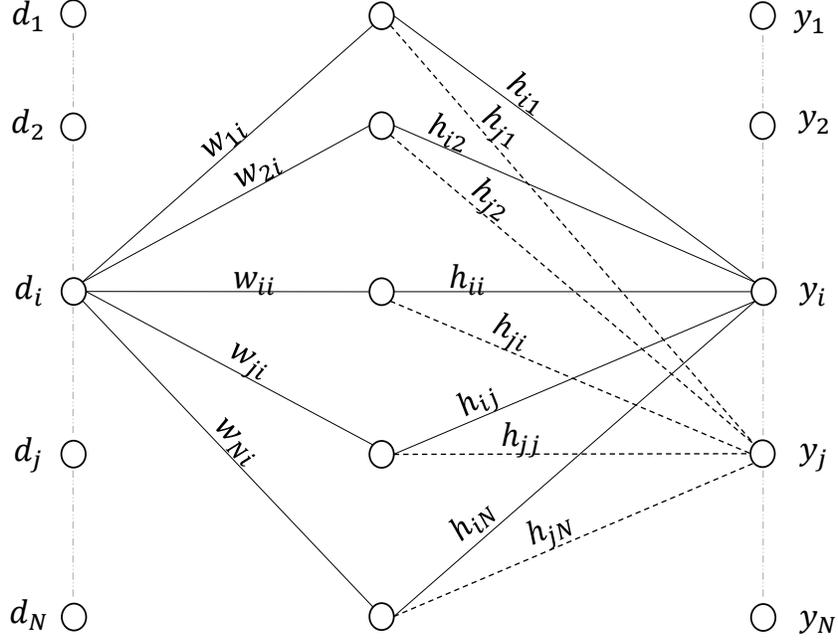


Fig. 3.3: Signal receiving method

However, the Eq. (6) is an indeterminate equation, so we assume the fixed value of w_{kk} and change the other value. This can be written using matrix form as

$$\mathbf{H}_k \mathbf{w}_k^{new} = -\mathbf{h}_k w_{kk}, \quad (7)$$

where,

$$\mathbf{H}_k = \begin{pmatrix} h_{1,1} & \cdots & h_{1,k-1} & h_{1,k+1} & \cdots & h_{1,N} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ h_{k-1,1} & \cdots & h_{k-1,k-1} & h_{k-1,k+1} & \cdots & h_{k-1,N} \\ h_{k+1,1} & \cdots & h_{k+1,k-1} & h_{k+1,k+1} & \cdots & h_{k+1,N} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ h_{N,1} & \cdots & h_{N,k-1} & h_{N,k+1} & \cdots & h_{N,N} \end{pmatrix}$$

$$\mathbf{w}_k^{new} = \left(w_{1,k}^{new} \quad \cdots \quad w_{k-1,k}^{new} \quad w_{k+1,k}^{new} \quad \cdots \quad w_{N,k}^{new} \right)^T$$

$\mathbf{h}_k = \left(h_{1,k} \quad \cdots \quad h_{k-1,k} \quad h_{k+1,k} \quad \cdots \quad h_{N,k} \right)^T$ and $1 \leq k \leq N$. w_k^{new} is the new value obtained by the

calculation. Solving this, the new weight vector is obtained:

$$\mathbf{w}_k^{new} = -(\mathbf{H}_k)^{-1} \mathbf{h}_k w_{kk}, \quad (8)$$

In this way, all the coefficients of the weighting network matrix \mathbf{W}^{new} are calculated at the weight controller and feed to the weighting network to control the transmitter's signal. The process from Eq. (1) to Eq. (8) is repeated continuously to update the weighting network. The received signals can be expressed as:

$$\begin{pmatrix} \mathbf{y}_1(t) \\ \mathbf{y}_2(t) \\ \vdots \\ \mathbf{y}_N(t) \end{pmatrix} = \begin{pmatrix} (h_{11}w_{11} + h_{12}w_{21} + \dots h_{1N}w_{N1})\mathbf{d}_1(t) \\ (h_{21}w_{12} + h_{22}w_{22} + \dots h_{2N}w_{N2})\mathbf{d}_2(t) \\ \vdots \\ (h_{N1}w_{1N} + h_{N2}w_{2N} + \dots h_{NN}w_{NN})\mathbf{d}_N(t) \end{pmatrix}, \quad (9)$$

We'll consider the case for $N = 2$. When the transmitter sends $\mathbf{d}(T) = (1 \ 0)^T$, receiver 1 gets $y_1(T)$ and receiver 2 gets $y_2(T)$. Similarly, when the transmitter sends $\mathbf{d}(2T) = (0 \ 1)^T$, receiver 1 and receiver 2 gets $y_1(2T)$ and $y_2(2T)$ respectively. These four values are the four coefficients of the total channel matrix. This process is repeated for the entire length of the training sequence and the average value of the coefficients is taken. The four coefficients of the total channel matrix are h_{11ov} , h_{12ov} , h_{21ov} and h_{22ov} .

These values are then uplinked to the weight controller. The weight controller calculates the channel matrix coefficients from Eq. (5) as

$$\mathbf{H} = \frac{1}{\Delta_w} \begin{pmatrix} h_{11ov}w_{22} - h_{12ov}w_{21} & -h_{11ov}w_{12} + h_{12ov}w_{11} \\ h_{21ov}w_{22} - h_{22ov}w_{21} & -h_{21ov}w_{12} + h_{22ov}w_{11} \end{pmatrix}, \quad (10)$$

where, $\Delta_w = w_{11}w_{22} - w_{12}w_{21}$. The weighting network matrix coefficients can be obtained by assuming $k= 1$ and 2 in Eq. (7) as:

$$\begin{aligned} \mathbf{w}_1^{new} &= (w_{21}^{new}), \mathbf{w}_2^{new} = (w_{12}^{new}) \\ \mathbf{H}_1 &= (h_{22}), \mathbf{H}_2 = (h_{11}) \\ \mathbf{h}_1 &= (h_{21}), \mathbf{h}_2 = (h_{12}) \end{aligned}$$

Then from Eq. (8) we get,

$$\begin{aligned} w_{21}^{new} &= -\frac{h_{21}}{h_{22}} w_{11} = -\frac{h_{21ov} w_{22} - h_{22ov} w_{21}}{-h_{21ov} w_{12} + h_{22ov} w_{11}} w_{11}; \text{ and} \\ w_{12}^{new} &= -\frac{h_{12}}{h_{11}} w_{22} = -\frac{-h_{11ov} w_{12} + h_{12ov} w_{11}}{h_{11ov} w_{22} - h_{12ov} w_{21}} w_{22} \end{aligned} \quad (11)$$

The received signals from Eq. (9) and Eq. (11) are:

$$\mathbf{y}_1(t) = \frac{\Delta}{h_{22}} w_{11} \mathbf{d}_1(t), \text{ and } \mathbf{y}_2(t) = \frac{\Delta}{h_{11}} w_{22} \mathbf{d}_2(t) \quad (12)$$

where, $\Delta = h_{11}h_{22} - h_{12}h_{21}$. By renewing weighting network matrix coefficients according to these values $\mathbf{y}_1(t)$ signal contains only $\mathbf{d}_1(t)$ and $\mathbf{y}_2(t)$ signal contains only $\mathbf{d}_2(t)$. Furthermore, the output port can be exchanged by controlling the weighting network. The $\mathbf{d}_2(t)$ channel is made available to $\mathbf{y}_1(t)$ and $\mathbf{d}_1(t)$ channel to $\mathbf{y}_2(t)$ by taking the weighting network matrix coefficients as:

$$w_{21}^{new} = -\frac{h_{11}}{h_{12}} w_{11} \text{ and } w_{12}^{new} = -\frac{h_{22}}{h_{21}} w_{22}, \quad (13)$$

When someone tries to sniff another port signal, it is not easy by this configuration. For example, if receiver 2 tries to sniff receiver 1's signal, receiver 1 can control the coefficients of w_{12} and w_{22} . On the other hand, receiver 2 can control w_{21} and w_{11} . Therefore, w_{12} and w_{22} are determined by receiver 1 signal. w_{21} and w_{11} are controlled without a cancellation and can receive channel 1 signal as:

$$\mathbf{y}_1(t) = (h_{11}w_{11} + h_{12}w_{21})\mathbf{d}_1(t), \quad (14)$$

$$\mathbf{y}_2(t) = (h_{21}w_{12} + h_{22}w_{22})\mathbf{d}_2(t), \quad (15)$$

This signal affects the sniffing signal and is not easy to sniff other signals. In our system, first, we use SCM, and then, sub-carrier signal amplitudes and phases are adjusted according to the weighting network matrix before optical amplitude modulations. A pre-MIMO equalization using one tap is used at the transmitter. If we use an optical carrier instead of a subcarrier, the weighting controller should be constructed by optical amplitude and phase controller such as optical vector

modulator.

3.2 Experimental setup

The experimental setup to confirm efficient operation of the mode forming network using conventional GI-MMF is shown in Fig. 3.4. The pulse pattern generator generates the signal at 100Mbps and this signal has training sequence appended to the data. While generating the signal, the data has the voltage level of +5V and -5V, whereas, the training sequence has the voltage level of +5V and 0V. d_1 and d_2 are the two channels. The d_2 signal is a 3-bit delay over d_1 .

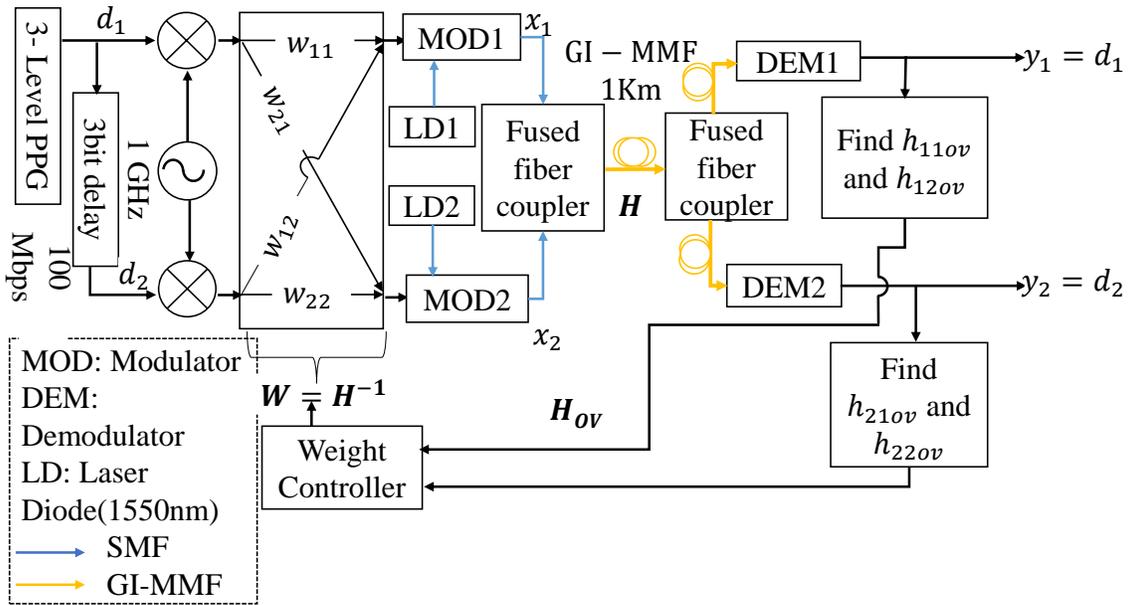


Fig. 3.4: 2×2 mode forming network

SCM is done using 1GHz microwave carrier and it is optically modulated using a Mach-Zehnder Modulator. The kind and method of implementation of SCM are as described in the paper [23]. The difference between the wavelengths of these optical carriers is kept 5GHz in order to avoid appearing of the product term of two signals, during square-law detection, in the same band with the desired signal. Two optical signals, each propagating through a single mode fiber, are coupled to ‘1km’ conventional GI-MMF, OM2 with 50 μm core diameter, by using a mode dependent fiber coupler. The coupling ratio of a multimode fiber coupler depends on the excited modes of MMF because each mode has a different propagation constant. If all the modes are excited uniformly, the coupling ratio of the fiber coupler will be constant. If this happens, there will be difficulty in MIMO transmission. To achieve non-uniform mode excitation, we have used an SMF just before

the fiber coupler so that we can excite only a few modes. Therefore, the fiber coupler used here is called a mode dependent. This fiber supports about 40 modes. At the other end of GI-MMF, another fiber coupler is used to optically divide the signal into two branches and each branch is transmitted to two different locations by using GI-MMF. Fiber coupler is used to perform multiplexing and de-multiplexing. In each location, the signals are first detected using photodetectors and are demodulated.

RF signal carrying d_1 and a fraction of RF signal carrying d_2 are superimposed and then fed to the modulator MOD1. Similarly, RF signal carrying d_2 and a fraction of RF signal carrying d_1 are superimposed and then fed to the modulator MOD2. Although we have used different modes of GI-MMF, their types are not exactly identified. The channel coefficients are estimated in two different locations independently by using a training sequence, which is shown in Fig. 3.5. The transmitting signal is a combination of two modulation formats; BPSK and OOK. Signal generated by the pulse pattern generator has three levels; 1, -1 and 0. Data is represented by 1 and -1 while, the training sequence is represented by 0 and 1. Therefore, training sequence uses OOK.

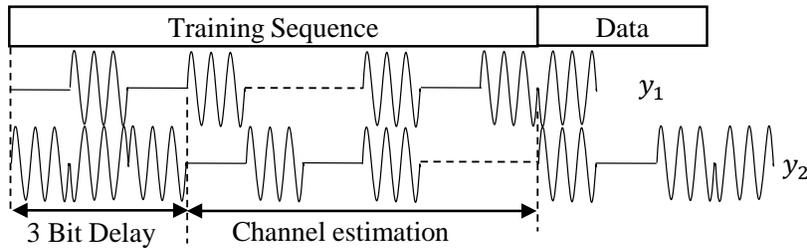


Fig. 3.5: Data format

During the transmission of the training sequence, when the upper branch transmits bit '1' lower branch transmits level '0' and vice versa. It resembles as when the upper branch transmits bit '1' (i.e. ON) of the training sequence, the lower branch is OFF and when the lower branch transmits bit '1' (i.e. ON) of the training sequence, the upper branch is OFF. A frame of 152-bit is used and it contains 25 bits of the training sequence starting at (010...1) from 1 to 25 positions. We can increase this frame length to a value less than the interval after which the fiber characteristic changes. From Eq. (3) & (4), the receiver 1 in the remote location receives $y_1(T)$ and $y_1(2T)$. Whereas, receiver 2 receives $y_2(T)$ and $y_2(T)$. The coefficients are calculated for the entire length of the training sequence and the average value is taken. In the received signal $y_1(t)$, the average

value of even bits from 4 to 25 position during the training sequence transmission gives h_{11ov} and the average value of odd bits from 4 to 25 position during the training sequence transmission gives h_{12ov} . The first three bits are neglected because we are using a 3-bit delay in the second branch. These two total channel coefficients are sent to the weight controller using uplink. Similarly, at receiver 2, remaining total channel coefficients h_{21ov} and h_{22ov} are calculated. These coefficients are also uplinked to the weight controller. In this experiment, we have used an Ethernet for uplink. The weight controller process these four coefficients h_{11ov} , h_{12ov} , h_{21ov} and h_{22ov} in order to calculate the channel coefficients h_{11} , h_{12} , h_{21} and h_{22} . The weight controller then uses these channel coefficients to generate control signals in order to update the weighting network to W^{new} . The weighting network consists of magnitude and phase controller in each branch. Only one channel is present in one destination when all the weighting values are updated.

3.3 Results

The cross-correlation between receiving signals and the reference signal is shown in Fig. 3.6. The cross-correlation between y_1 and the reference signal reveal that both channel d_1 and d_2 are present in y_1 before implementing feedback, which is shown in Fig. 3.6(a). Similarly, the cross-correlation between y_2 and the reference signal also shows that y_2 contains both channels before implementing feedback, as shown in Fig. 3.6(d). However, after implementing continuous feedback only one channel is present in each location. y_1 contains only d_1 and y_2 contains only d_2 as shown in Fig. 3.6(b) and Fig. 3.6(e) respectively. Due to the implementation of weighting network, each receiver receives two copies of each transmitted signal. The signals which are intended for that user are in phase, so they get added. Whereas, the signals which are not intended for that user are 180 degrees out of phase, so they get canceled. In this system, perfect cancellation is not suitable because if perfect cancellation occurs, channel matrix coefficients cannot be calculated. Therefore, we can see in the above figures that a very weak portion of an unwanted signal is present in both locations due to crosstalk. If we increase the length of the training sequence, SNR increases and the crosstalk can be reduced.

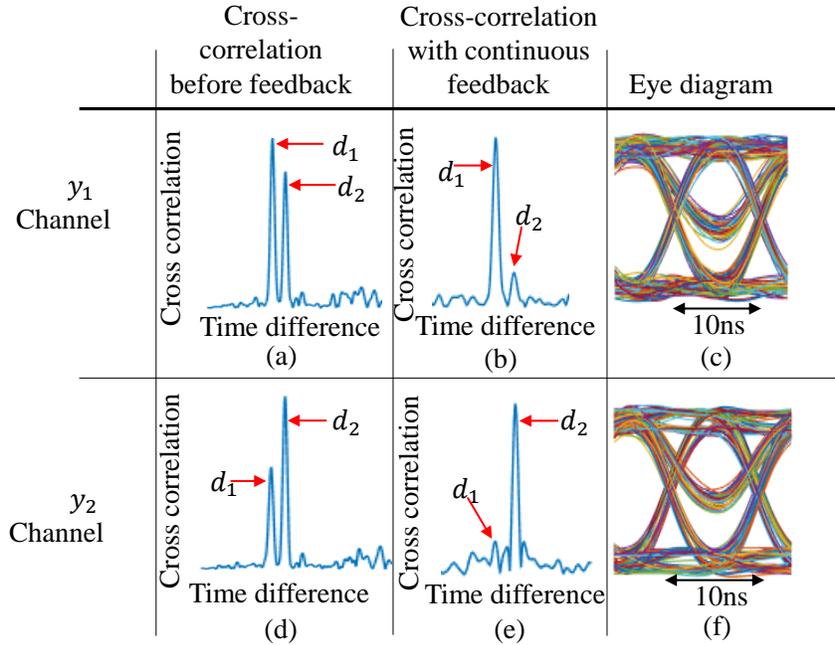


Fig. 3.6: y_1 channel; (a) correlation before feedback, (b) correlation with continuous feedback, (c) eye diagram. y_2 channel; (d) correlation before feedback, (e) correlation with continuous feedback, (f) eye diagram.

The eye diagram of y_1 and y_2 after feedback is shown in Fig. 3.6(c) and Fig. 3.6(f) respectively. These eye diagrams have three levels. The middle level is due to the training sequence, so it can be removed before the decision threshold. The output port can be exchanged by controlling the weight controller. After continuous feedback, channels have been successfully exchanged making d_2 available to y_1 and d_1 to y_2 as shown Fig. 3.7.

In Fig. 3.6, at the port y_1 , the cross-correlation value of d_1 is larger compared to the cross-correlation value of d_2 . Similarly, at the port y_2 , the cross-correlation value of d_2 is larger compared to that of d_1 . The weighting matrix coefficients when continuous feedback is implemented are shown in Table. 3.1. The magnitude and phase of these weighting matrix coefficients are shown in Fig. 3.8 and it can be seen that the change in magnitude of these coefficients is less than 2dB.

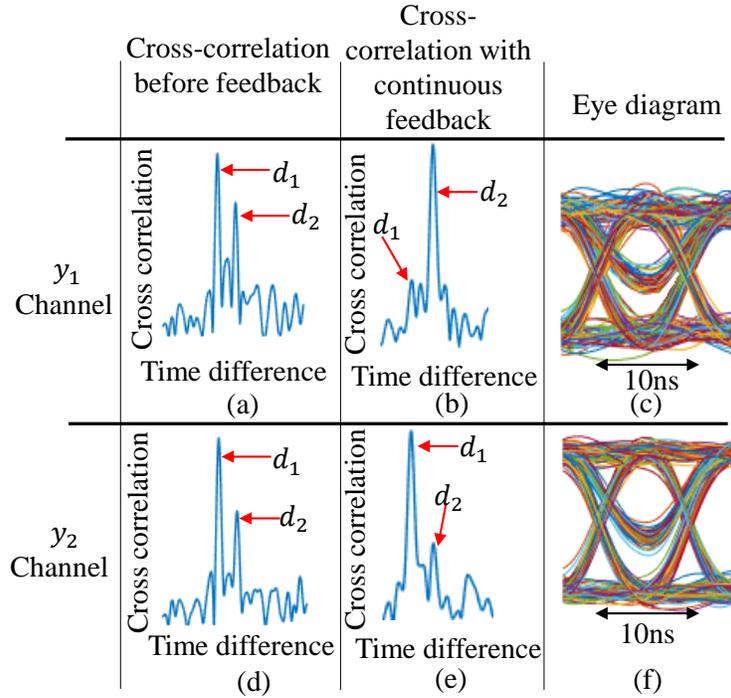


Fig. 3.7: Output port exchange. y_1 channel; (a) correlation before feedback, (b) correlation with continuous feedback, (c) eye diagram. y_2 channel; (d) correlation before feedback, (e) correlation with continuous feedback, (f) eye diagram.

3.4 Protection against sniffing

Receiver 1 controls the coefficients w_{12} and w_{22} while receiver 2 controls the coefficients w_{21} and w_{11} . If receiver 2 tries to sniff the signal of receiver 1 without attacking the value of channel coefficients, then it is difficult to detect. Similarly, if receiver 1 tries to sniff the signal of receiver 2 it is difficult to detect. This can be clearly seen in Fig. 3.9. By increasing the SNR level of the transmitting signal, the level of crosstalk can be further reduced. However, if the attacker attacks the value of the channel coefficients and alters their values, then both the signal might be detected at each location. Nevertheless, the occurrence of an attack can be known because the SNR level in another receiver decreases.

Therefore, this network clearly reveals security against sniffing by making us aware of sniffing. The input signals are successfully delivered to their desired destination. In this way, a 2×2 optical mode forming network has been successfully realized.

3.5 Discussion

The proposed passive optical mode forming network using conventional GI-MMF with mode division multiplexing and sub-carrier multiplexing has been successfully implemented.

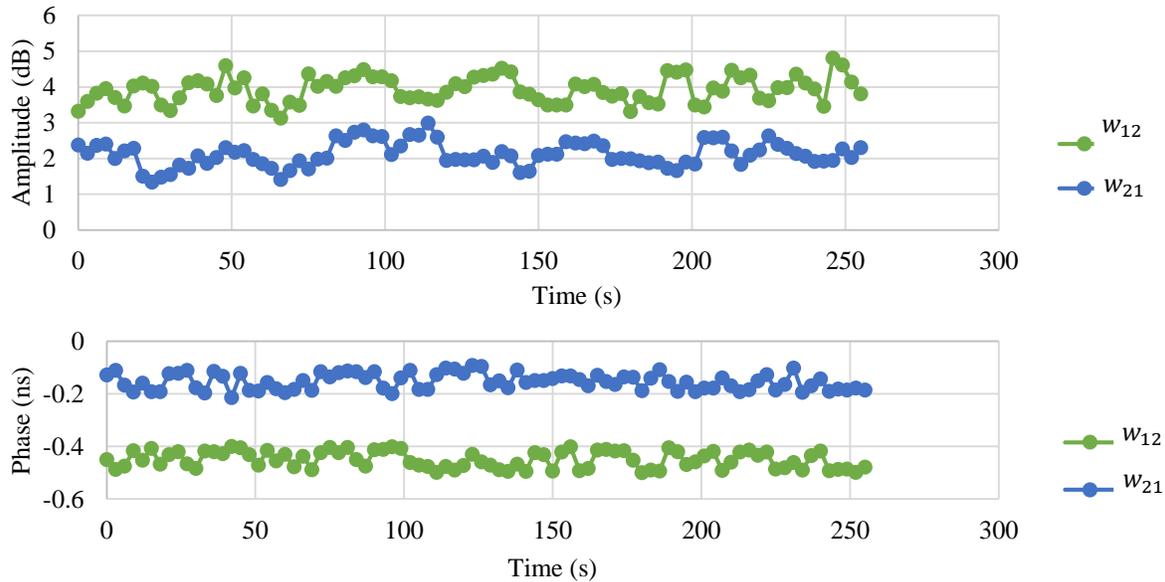


Fig. 3.8: Weighting matrix. (a) magnitude. (b) phase

Table 3.1: weighting matrix coefficients

S.N	w_{12}	w_{21}
1	0.6250-0.4957i	0.6353-0.0469i
2	0.5837-0.5108i	0.5619-0.2582i
3	0.6411-0.4738i	0.5821-0.2032i
4	0.6441-0.4793i	0.6123-0.1751i
5	0.6184-0.0593i	0.6184-0.0593i
6	0.6801-0.4363i	0.6314-0.0403i

We have chosen a symbol rate of 100Mbps because the purpose of our experiment is to test the feasibility of our proposed technique. The GI-MMF fiber that we have used in our experiment has a bandwidth-distance product of 500MHzkm. If the data rate is increased, ISI increases error. However, for commercial use, the fiber length should be less than 10km with data rate above 10Gbps. This higher data rate can be achieved by increasing the number of sub-carriers and combining SCM. We must use OOK format for the training sequence; however, we can use more

complicated format for data, such as QAM, to increase the data rate. The vibration in fiber and changes in temperature changes the channel matrix. Therefore, in order to update the weighting matrix, the measured channel matrix should be fed back to the transmitter.

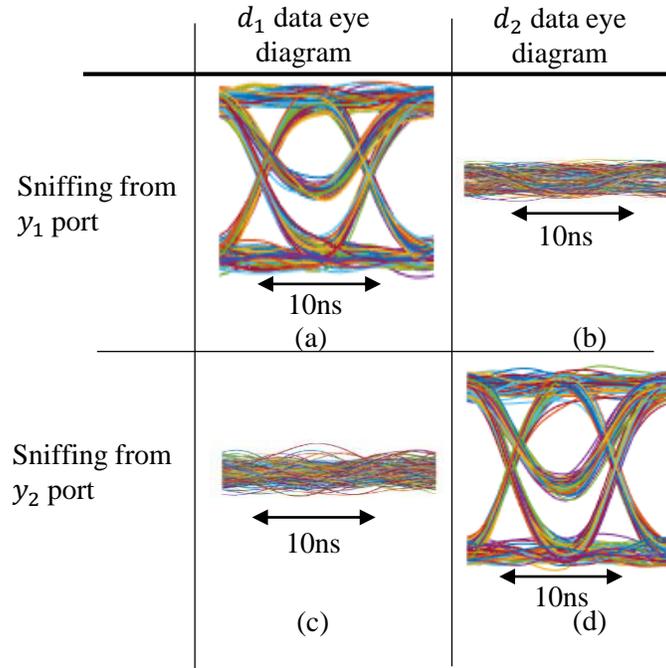


Fig. 3.9: Eye pattern when sniffing is done.

The spectral and temporal properties of the acoustic noise induced in the fiber by temperature and vibration are more than 10ms [24]. So the fiber characteristic changes in a time period not less than 10ms. But, the round trip time for a 10km long fiber is about 100 μ s. It does not cause a considerable problem, until the feedback time from each receiver is less than 10ms, even if the receivers located at different remote location feedback the total channel matrix coefficients at a different delay. Therefore, the process of measuring the transfer matrix and sending it back to the transmitter to update the weighting matrix should be faster than the temporal change of the channel matrix.

3.6 Conclusion

This technique can be used to realize any $N \times N$ channel optical mode delivering network. N data channels on the transmitter side can be modulated by a sub-carrier and superimposed, the proportion of superimposition is determined by the fiber characteristics, to generate N different signals. These signals are optically modulated and then transmitted as N different modes. However,

for the $N \times N$ system, the length of the training sequence becomes $(N-1)$ times longer than that of the 2×2 system and the undesired signal might not be completely canceled out.

3.7 References

- [1] S. H. S. Newaz, W. S. H. Suhaili, G. M. Lee, M. R. Uddin, A. F. Y. Mohammed, and J. K. Choi, "Towards Realizing the Importance of Placing Fog Computing Facilities at the Central Office of a PON," in *proc. ICACT*, Phoenix Park, PyeongChang, Korea, 2017, pp. 152–157.
- [2] Derek Nasset, "PON Roadmap," *J. Opt. Commun. Netw.*, vol. 9, no. 1, pp. 71-76, Jan. 2017.
- [3] K. Kondepu, A. Sgambelluri, L. Valcarenghi, F. Cugini, and P. Castoldi, "Exploiting SDN for Integrating Green TWDM-PONs and Metro Networks Preserving End-to-End Delay," *J. Opt. Commun. Netw.*, vol. 9, no. 1, pp. 67-74, Jan. 2017.
- [4] H. Yao, W. Li, Q. Feng, J. Han, Z. Ye, Q. Hu, Q. Yang, and S. Yu, "Ring-Based Colorless WDM-PON With Rayleigh Backscattering Noise Mitigation," *J. Opt. Commun. Netw.*, vol. 9, no. 1, pp. 27-35, Jan. 2017.
- [5] Q. Wang, K. H. Tse, L. K. Chen and S. C. Liew, "Physical-layer network coding for VPN in TDM-PON," *IEEE Photon. Technol. Lett.*, vol. 24, no. 23, pp. 2166–2168, Dec. 2012.
- [6] T. Kodama, N. Wada, G. Cincotti and K. I. Kitayama, "Asynchronous OCDM-based 10 G-PON using cascaded multiport E/Ds to suppress MAI noise," *J. Lightwave. Technol.*, vol. 31, no. 20, pp. 3258–3266, Oct. 2013.
- [7] C. W. Chow, C. H. Yeh, C. H. Wang, F. Y. Shih and S. Chi, "Signal modulation of OFDM-QAM for long reach Carrier distributed passive optical networks," *IEEE Photon. Technol. Lett.*, vol. 21, no. 11, pp. 715–717, Jun. 2009.
- [8] W. Lee, M. Y. Park, S. H. Cho, J. Lee, B. W. Kim, G. Jeong and B. W. Kim, "Bidirectional WDM-PON based on gain-saturated reflective semiconductor optical amplifiers," *IEEE Photon. Technol. Lett.*, vol. 17, no. 11, pp. 2460–2462, Nov. 2005.
- [9] Y. Luo, X. Zhou, F. Effenberger, X. Yan, G. Peng, Y. Qian and Y. Ma, "Time- and wavelength-division multiplexed passive optical network (TWDM-PON) for next-generation PON stage 2 (NG-PON2)," *J. Lightwave. Technol.*, vol. 31, no. 4, pp. 587–593, Feb. 2013.

- [10] G. Das, B. Lannoo, A. Dixit, D. Colle, M. Pickavet and P. Demeester, “Flexible hybrid WDM/TDM PON architectures using wavelength selective switches,” *Optical Switching Network*, vol. 9, no. 2, pp. 156–169, Apr. 2012.
- [11] H. Yang, J. Li, B. Lin, Y. Wan, Y. Guo, L. Zhu, L. Li, Y. He and Z. Chen, “DSP-based evolution from conventional TDM-PON to TDM-OFDM-PON,” *J. Lightwave Technol.*, vol. 31, no. 16, pp. 2735–2741, Aug. 2013.
- [12] D. J. Richardson, J. M. Fini and L.E. Nelson, “Space-division multiplexing in optical fibers,” *Nat. Photonics* 7, pp. 354–362, Apr. 2013.
- [13] G. Li, N. Bai, N. Zhao and C. Xia, “Space-division multiplexing: the next frontier in optical communication,” *Adv. Opt. Photonics* 6, pp. 413–487, Dec. 2014.
- [14] H. Wen, C. Xia, A. M. V. Benitez, N. Chand, J. E. A. Lopez, B. Huang, H. Liu, H. Zheng, P. Sillard, X. Liu, F. Effenberger, R. A. Correa, and G. Li, “First Demonstration of Six-Mode PON Achieving a Record Gain of 4 dB in Upstream Transmission Loss Budget,” *J. of Lightwave Technol.*, vol. 34, no. 8, pp. 1990–1996, April. 2016.
- [15] C. Xia, N. Chand, A. M. V. Benitez, Z. Yang, X. Liu, J. E. A. Lopez, H. Wen, B. Zhu, N. Zhao, F. Effenberger, R. A. Correa, and G. Li, “Time-division-multiplexed few-mode passive optical network,” *Opt. Exp.*, vol. 23, no. 2, pp. 1151–1158, Jan. 2015.
- [16] R. Fang, L. Juhao, W. Zhongying, H. Tao, Y. Jinyi, M. Qi, H. Yongqi, C. Zhangyuan and L. Zhengbin, “Three-mode mode-division-multiplexing passive optical network over 12-km low mode-crosstalk FMF using all-fiber mode MUX/DEMUX,” *Optics Comm.*, vol. 383, pp. 525–530, Jan. 2017.
- [17] J. Li, F. Ren, T. Hu, Z. Li, Y. He, Z. Chen, Q. Mo and G. Li, “Recent progress in mode-division multiplexed passive optical networks with low modal crosstalk,” *Optical Fiber Technology*, vol. 35, pp. 28–36, Feb. 2017.
- [18] Wang L, Ai J, Zhu L, Wang A, Fu S, Du C, Mo Q, Wang J, “MDM transmission of CAP-16 signals over 1.1- km anti-bending trench-assisted elliptical-core few-mode fiber in passive optical networks,” *Optics Express*, vol. 25, no. 19, pp. 22991–23002, Sept. 2017

- [19] Z. Wu, J. Li, D. Ge, F. Ren, P. Zhu, Q. Mo, Z. Li, Z. Chen, and Y. He, "Demonstration of all-optical MDM/WDM switching for short-reach networks," *Optics Express*, vol. 24, no. 19, pp. 21609–21618, Sept. 2016
- [20] F. Ren, J. Li, T. Hu, R. Tang, J. Yu, Q. Mo, Y. He, Z. Chen and Z. Li, "Cascaded Mode Division Multiplexing and Time Division Multiplexing Passive Optical Network Based on Low Mode-Crosstalk FMF and Mode MUX/DEMUX," *IEEE Photon. J.*, vol. 8, no. 3, Oct. 2015, Art. no. 7903509.
- [21] Y. Feng, J. Luo, Y. Li, J. Li and Z. Li, "Mode Multiplexing and High Efficient Switching in Few-Mode Fiber Based on Modeled Blazed Grating," *IEEE Photon. J.*, vol. 8, no.3, Jun. 2016, Art. no. 7803207.
- [22] Y. Morizumi, H. Kobayashi and K. Iwashita, "Arbitrary Output Port Selection in Multimode Fiber Networks using Mode Division Multiplexing," in *Proc. OECC/PS*, Niigata, Japan, 2016.
- [23] Y. Morizumi, H. Kobayashi and K. Iwashita, "User Selective mode-forming remote switch in multi-mode fiber distributing networks using MIMO processing," *Optics Commun.*, vol. 392, pp. 234-238, Jun. 2017.
- [24] D. S. Waddy, P. Lu, L. Chen and X. Bao, "Fast State of Polarization Changes in Aerial Fiber Under Different Climatic Conditions," *IEEE Photon. Technol. Lett.*, vol. 13, no. 9, pp. 1035-1037, Sept. 2001.

Chapter 4. MIMO Detection using a Deep Learning Neural Network in a MDM Optical Transmission System.

4.1 DLNN in MDM optical transmission system

The objective of our MDM optical transmission system is to implement Deep Learning Neural Network for MIMO detection. We have used an MDM on GI-MMF systems using sub-carrier signals instead of an optical carrier. The advantages of using sub-carrier signals are that they have an extremely longer wavelength than the optical carrier. Because of this, the effect of phase change in sub-carrier signal is small compared to that in the optical carrier signal, so sub-carrier systems are more stable than an optical system. Moreover, optical sub-carrier signals can be converted directly to electrical signals without using complicated coherent detection.

The MDM optical transmission system with N number of input channels and N number of output channels is shown in Fig. 4.1. It consists of transmitters, GI-MMF, mode-dependent fused fiber coupler, receivers, and a DLNN. Mode dependent fused fiber couplers are used for achieving mode multiplexing and de-multiplexing operation.

Data $d_1, d_2, d_3, \dots, d_N$ are the random data stream. These data stream should be random in order to avoid the DLNN learning this data pattern. Data are sub-carrier multiplexed and then feed into the Mach-Zehnder Modulator (MZM). In MZM, an intensity of the optical carrier is modulated. After optical modulation, these optical signals are mode division multiplexed with the help of a mode-dependent fused fiber coupler.

In our experiment, we have used a commercially available mode-dependent fused fiber coupler. The fused fiber coupler consists of two closely placed parallel optical fiber that are stretched and fused together so that their cores are very close to each other. The length of the Coupling region, the separation of the cores, the wavelength of operation and other fiber parameters determine the coupling coefficients. The coupling coefficients quantify how efficiently power couple from one mode to another. The strength of mode coupling between the two propagating modes is determined by the dimensionless ratio of the coupling coefficient (per unit length) to the difference between the propagation constant of these propagating modes. Then the output port of fused fiber coupler is connected to the GI-MMF.

Bends, stress and index perturbations (whether random or intentional) can induce mode coupling in an GI-MMF. Strong coupling occurs when the propagation constants of two modes are nearly equal, otherwise limited coupling appears. To classify mode coupling as strong or weak, a correlation length is defined. A correlation length is defined as the length along a fiber over which the local eigenvectors [19] remains constant. When the length of fiber is shorter than the correlation length, weak coupling occurs. In this case, there might be a remarkable level of coupling between modes that have approximately equal propagation constant, but the coupling between modes that have an unequal propagation constant is limited. As a result, in different spatial mode group, GI-MMF shows weak mode coupling [20]. When the length of fiber is longer than the correlation length, coupling becomes stronger. In this case, remarkable coupling occurs between all the possible modes.

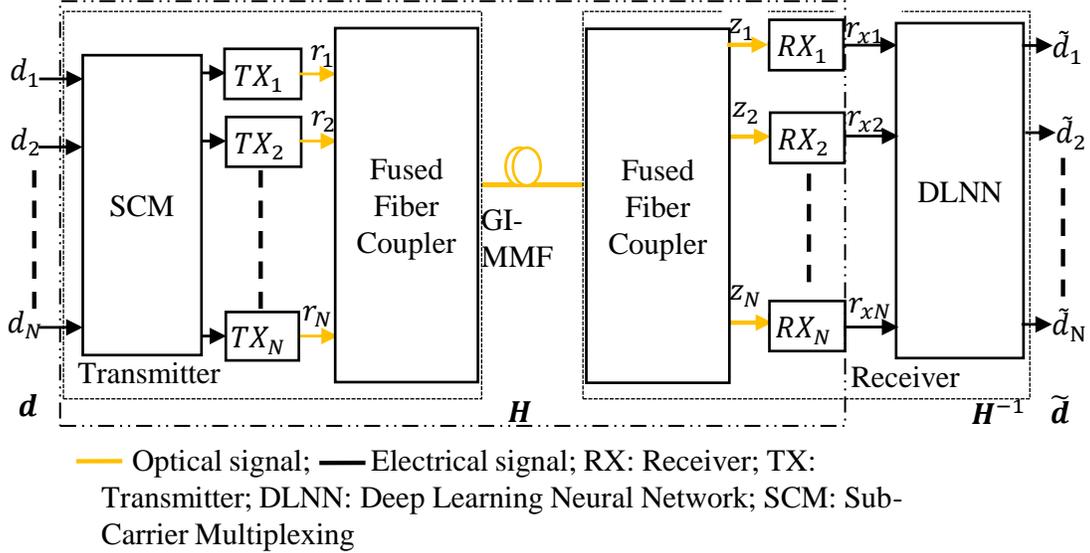


Fig.4.1: MDM optical transmission system that uses DLNN for MIMO detection.

In the regime of strong mode coupling, let \mathbf{P} denote a propagation operator which includes both modal losses and gains. Mode dependent loss and gain (MDL) can be described by taking the log of the eigenvalues of $\mathbf{P}\mathbf{P}^H$. In [21], under the strong coupling, two fundamental propositions were verified. The statistics of MDL is described by the eigenvalue distribution of a zero-trace Gaussian unitary ensemble. The standard deviation of the overall MDL is given as $\sigma_{mdl} = \chi \sqrt{1 + \frac{\chi^2}{12}}$, where χ represents the standard deviation of accumulated MDL. If K is the number of MDL sources each

with MDL variance σ_g^2 , then $\chi = \sqrt{K}\sigma_g$. At another end of GI-MMF fiber, another mode-dependent fused fiber coupler divides the signal into two ports. Both the signals are present in each output port of the fused fiber coupler and these signals have different coefficients.

At the receiver side, in Fig. 4.1, the optical signal is divided into N branches using another mode-dependent fused fiber coupler, converted to electrical signal using photodiode, sub-carrier synchronized, and demodulated. The received signals are fed to the DLNN to perform MIMO detection. This DLNN, which is a supervised neural network, should be first trained with the input variables and output variables. Then, by using a suitable algorithm, it learns the mapping from the input to the output. The accuracy of DLNN depends on the number of data used to train it. In order to deploy DLNN at the diverse environmental condition, it must be trained by data at various channel conditions. A well trained DLNN generates the response \mathbf{H}^{-1} , where \mathbf{H} represents the channel matrix of the system.

4.2 Design of DLNN

A N mode MDM optical transmission system can be represented as a $N \times N$ MIMO system. For a $N \times N$ MIMO system with input $\mathbf{d}(t) = [d_1(t), d_2(t), d_3(t), d_4(t), \dots, d_N(t)]^T \in \mathbf{B}^N$, the output $\tilde{\mathbf{d}}(t) \in \mathbf{B}^N$ is

$$\tilde{\mathbf{d}}(t) = \mathbf{H}^{-1} \mathbf{H} \mathbf{d}(t). \quad (1)$$

where \mathbf{B} denotes the set of binary numbers; $\mathbf{H} \in \mathbb{C}^{N \times N}$ denotes a channel matrix, where \mathbb{C} denotes the set of complex numbers; and \mathbf{H}^{-1} denotes the response of DLNN. The entire signal processing technique for an i th channel of Fig. 4.1 from input to just before DLNN is shown in Fig. 4.2. In the SCM block, there is a bank of N Quadrature Phase Shift Keying (QPSK) modulators with sub-carrier frequency f_c . The TX_i block consists of MZM (Mach-Zehnder Modulator) and an optical carrier. The n th input bit on the i th data channel is denoted as $d_{i,n}$, where $n = 0, 1, 2, \dots, \infty$. The output of an i th electrical QPSK modulator is

$$a_i(t) = q_i e^{j2\pi f_c t}. \quad (2)$$

where f_c represents the sub-carrier frequency and $q_i = (d_{i,2n} + jd_{i,2n+1})$ represents a QPSK symbol. This QPSK modulated signal drives an MZM optical intensity modulator. All the optical carriers used here has the same wavelength. The output of the optical modulator is

$$r_i(t) = P_{bias} + P_i \text{Re}[a_i(t)] . \quad (3)$$

where P_{bias} represents the optical power due to a bias voltage, P_i represents the amplitude of the optical carrier, and $\text{Re}[a_i(t)]$ represents the real part of $a_i(t)$. The bias voltage adjusts the MZM operating point to maintain the linear relations between the applied electrical signal and modulated optical power. These modulated optical lights are coupled to different modes of GI-MMF using a mode-dependent fused fiber coupler. This fused fiber coupler provides mode division multiplexing. After mode division multiplexing, all N optical signals are propagated through a GI-MMF. During propagation, there is a partial coupling between different modes.

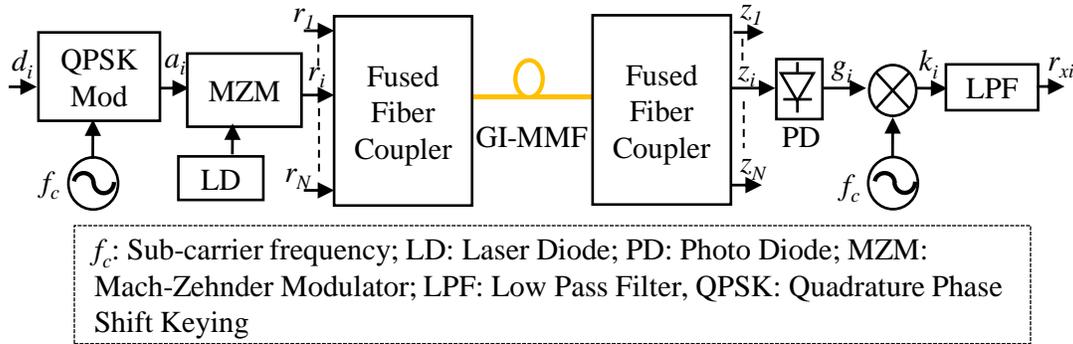


Fig. 4.2: Signal processing technique on the i th channel.

At the receiver side, the optical signal is divided optically into N branches by using another mode-dependent fused fiber coupler. The signal on the i th output port of the fused fiber coupler is

$$z_i(t) = \sum_{j=1}^N r_j(t) h_{ij}(t) . \quad (4)$$

where $h_{ij}(t)$ represents the (i,j) th element of \mathbf{H} and is complex. Absolute of $h_{ij}(t)$ are power coupling and phases of $h_{ij}(t)$ denote time differences of subcarrier. Equation 4 not only contains the signal transmitted by the TX_i but also the signal from all other transmitters. This is because of mode coupling during transmission. The current generated by the i th photodetector is

$$I_i(t) \propto z_i(t) = \sum_{j=1}^N r_j(t) h_{ij}(t) . \quad (5)$$

This current produces a voltage signal $g_i(t)$, which is given as

$$g_i(t) = \sum_{j=1}^N P_j h_{ij}(t) [a_j(t)]. \quad (6)$$

The term P_{bias} in Eq. (3) is a dc value, so it vanishes after passing through a band-pass filter, so it is not present in the Eq. (6). This signal is multiplied by a sub-carrier frequency f_c to generate $k_i(t)$ as

$$k_i(t) = g_i(t) e^{-j2\pi f_c t} = \left[\sum_{j=1}^N P_j h_{ij}(t) [a_j(t)] \right] e^{-j2\pi f_c t}. \quad (7)$$

After low pass filtering, the signal on the i th port is

$$r_{xi}(t) = \sum_{j=1}^N P_j h_{ij}(t) q_j. \quad (8)$$

The received signal $r_{xi}(t)$ is the summation of all the N transmitted signals. The output from N ports is $\mathbf{R}_x(t) = [r_{x1}(t), r_{x2}(t), r_{x3}(t), r_{x4}(t), \dots, r_{xN}(t)]^T \in \mathbb{C}^N$. In order to build a training dataset and test dataset, a continuous synchronous transmission of \mathbf{d} is done and \mathbf{R}_x is received.

Figure 4.3 shows the structure of our DLNN. An input vector to the DLNN is

$$\mathbf{x}_0 = [\text{Re}(r_{x1}(t)), \text{Im}(r_{x1}(t)), \text{Re}(r_{x2}(t)), \text{Im}(r_{x2}(t)), \dots, \text{Re}(r_{xN}(t)), \text{Im}(r_{xN}(t))]. \quad (9)$$

A neuron in an artificial neural network is a mathematical approximation of a biological neuron. It takes a vector of inputs, performs a transformation on them, and output a single scalar value. The length of \mathbf{x}_0 is $D=2N$. If \mathbb{R} denote the set of real numbers, a deep feedforward neural network with M layers describes a mapping operation $f(\mathbf{x}_0; \boldsymbol{\theta}) : \mathbb{R}^{D_0} \mapsto \mathbb{R}^{D_M}$ of an input vector $\mathbf{x}_0 \in \mathbb{R}^{D_0}$ to an output vector $\mathbf{x}_M \in \mathbb{R}^{D_M}$ through M iteration as

$$\mathbf{x}_m = f_m(\mathbf{x}_{m-1}; \boldsymbol{\theta}_m), \quad m = 1, \dots, M \quad (10)$$

where $f_m(\mathbf{x}_{m-1}; \boldsymbol{\theta}_m) : \mathbb{R}^{D_{m-1}} \mapsto \mathbb{R}^{D_m}$ is the mapping operation performed by the m th layer. This operation relies on a set of parameters $\boldsymbol{\theta}_m$ and output vector \mathbf{x}_{m-1} from the preceding layer. This mapping can be a stochastic process. In order to represent the set of all parameters of the neural network, $\boldsymbol{\theta} = \{\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_M\}$ is used. All the layers of our network are fully connected. An m th layer becomes a fully connected layer if it has the form

$$f_m(\mathbf{x}_{m-1}; \theta_m) = \sigma(\mathbf{W}_m \mathbf{x}_{m-1} + \mathbf{b}_m). \quad (11)$$

where $\mathbf{W}_m \in \mathbb{R}^{D_m \times D_{m-1}}$, $\mathbf{b}_m \in \mathbb{R}^{D_m}$ and $\sigma(\cdot)$ is an activation function [22]. Each element of the input vector is passed through the activation function just as $[\sigma(\mathbf{u})]_i = \sigma(u_i)$. The parameters for this m th layer is $\theta_m = \{\mathbf{W}_m, \mathbf{b}_m\}$. Whenever any layer is called, a new random mapping is generated. The activation function is very important for the expressive power of the neural network because it introduces nonlinearity. This nonlinearity provides an advantage of using multiple hidden layers.

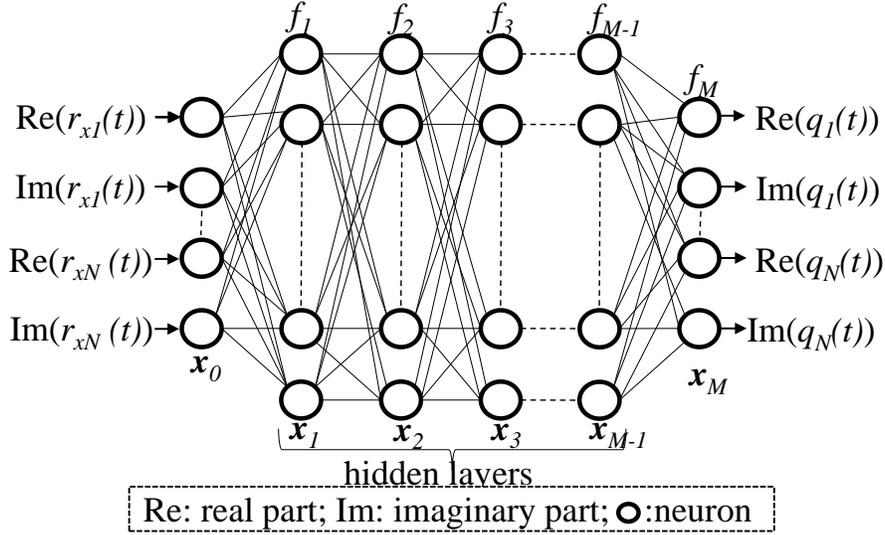


Fig. 4.3: Our deep learning neural network structure.

For training DLNN, a dataset $(\mathbf{x}^{(i)}, \mathbf{y}^{(i)})$, $i = 1, \dots, S$, where $\mathbf{y}^{(i)}$ is the desired output for $\mathbf{x}^{(i)}$ input and S is the number of training examples, is first developed and used. A continuous synchronous transmission of \mathbf{d} is done and \mathbf{R}_x is received at different channel condition number. Channel CN is a measure that indicates channel correlation property and signifies the difficulties of recovering MIMO signals. The CN of channel matrix \mathbf{H} is defined as $20 \log_{10}(\|\mathbf{H}\| \cdot \|\mathbf{H}^{-1}\|) \geq 0\text{dB}$, where $\|\cdot\|$ is the norm. The closer the CN approaches 0dB, the better the quality of the MIMO channel. The matrix \mathbf{H} represents the characteristics of two fused fiber coupler and GI-MMF. The characteristics of GI-MMF gives information about modal dispersion (MD), MDL, other mode-dependent effects, and crosstalk due to mode coupling in the fiber. For the real-time implementation of DLNN for MIMO detection, it must be trained by data at different CNs. Training minimizes the cost

$$J(\boldsymbol{\theta}) = \frac{1}{S} \sum_{i=1}^S L(f(\mathbf{x}^{(i)}; \boldsymbol{\theta}), \mathbf{y}^{(i)}). \quad (12)$$

with respect to $\boldsymbol{\theta}$, where $L(\mathbf{u}, \mathbf{v})$ represents the per-example cost function and $f(\mathbf{x}^{(i)}; \boldsymbol{\theta})$ represents the actual output of the neural network at the input $\mathbf{x}^{(i)}$. A cost function helps us to know how accurate our DLNN model is at making predictions for a given set of parameters. Mean Square Error (MSE) is a convex function so it has only one global minimum value. If MSE is used as a cost function, then Eq. (12) becomes

$$J(\boldsymbol{\theta}) = \frac{1}{S} \sum_{i=1}^S \|f(\mathbf{x}^{(i)}; \boldsymbol{\theta}) - \mathbf{y}^{(i)}\|_2^2. \quad (13)$$

Various types of optimization algorithm exist for the training of neural networks [22]. First-order optimization algorithms use only the gradient of a cost function, whereas the second-order uses a matrix of second derivatives of a cost function called a Hessian matrix. Even though second-order optimization algorithms are able to provide faster convergence than first-order optimization algorithms, we have used a first-order optimization algorithm in our experiment because of the high cost of computing the second-order information [22].

First order optimization algorithms are simple, cheap, and its complexity is nearly independent of dimension. Gradient descent is a relatively efficient optimization method if the objective function is differentiable with respect to its parameters because the computational complexity for calculating first-order partial derivatives with respect to all the parameters is the same as just evaluating the function. Usually, objective functions are stochastic. Stochastic gradient descent (SGD) is a popular algorithm for optimization. However, it uses a common learning rate for all parameters. For functions with a huge number of parameters, this might be problematic and this problem is even more pronounced in a higher dimension. SGD is noisier, it oscillates around the minimum giving some variation in accuracy, and it is very difficult to tune the learning rate. Other than data subsampling noise, objective functions may have noise sources such as dropout regularization [23]. Using adaptive learning rate along different axes can solve these issue [24].

The Adam [24], is the most popular adaptive learning rate optimization algorithm to update the parameters $\boldsymbol{\theta}$. By using the estimates of first and second moments of the gradient, it calculates individual adaptive learning rates for all the parameters. It optimizes the first-order gradient of stochastic objective functions based on adaptive estimates of low-order moments. It is easy in

terms of implementation, and efficient in terms of computation. Furthermore, it requires little memory, it does not vary to a diagonal rescaling of the gradients, it is suitable for problems that involve larger data and/or parameters. In this method, hyper-parameters require little tuning and have intuitive interpretations. Our cost function $J(\boldsymbol{\theta})$ is the noisy objective function, and it is differentiable with respect to $\boldsymbol{\theta}$. Our objective is to minimize the expected value of this function with respect to $\boldsymbol{\theta}$. At first, a random value is initialized for the parameter $\boldsymbol{\theta}$ as $\boldsymbol{\theta} = \boldsymbol{\theta}_0$. The first and second-moment variables are initialized to $\mathbf{s} = \mathbf{0}$ and $\boldsymbol{\beta} = \mathbf{0}$ respectively. Other variables such as exponential decay rates for first and second-moment estimates, ρ_1 and ρ_2 in $[0,1)$, are also initialized to their default values 0.9 and 0.999 respectively. Step-size and small constant are set to their default values $\varepsilon = 0.001$ and $\delta = 10^{-8}$ respectively. The selection of these hyperparameters value has a large influence on the quality of solutions. Time step t is initialized to $t=0$. We sampled a mini-batch of ξ examples from the training set $\{\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(\xi)}\}$ with corresponding desired outputs $\{\mathbf{y}^{(1)}, \mathbf{y}^{(2)}, \dots, \mathbf{y}^{(\xi)}\}$. Then, the gradient of the cost function with respect to $\boldsymbol{\theta}$ is computed at time-step t as

$$\mathbf{g} = \frac{1}{\xi} \nabla_{\boldsymbol{\theta}} \sum_i L(f(\mathbf{x}^{(i)}; \boldsymbol{\theta}), \mathbf{y}^{(i)}). \quad (14)$$

Now the time step is updated as, $t \leftarrow t + 1$. The updated value for the biased first-moment estimate is

$$\mathbf{s}^{new} = \rho_1 \mathbf{s} + (1 - \rho_1) \mathbf{g}. \quad (15)$$

Similarly, the updated value for the biased second-moment estimate is

$$\boldsymbol{\beta}^{new} = \rho_2 \boldsymbol{\beta} + (1 - \rho_2) \mathbf{g} \odot \mathbf{g}. \quad (16)$$

where \odot represents the elementwise product. The correct bias in the first-moment is given as

$$\hat{\mathbf{s}} = \frac{\mathbf{s}^{new}}{1 - \rho_1^t}. \quad (17)$$

where ρ_1^t is the exponential decay rate of the first-moment at time-step t . Similarly, the correct bias in the second-moment is given as

$$\hat{\boldsymbol{\beta}} = \frac{\boldsymbol{\beta}^{new}}{1 - \rho_2^t}. \quad (18)$$

where ρ_2^t is the exponential decay rate of the second-moment at time-step t . Now the parameter update is computed, with an element-wise operation, as

$$\Delta \boldsymbol{\theta} = -\varepsilon \frac{\hat{\boldsymbol{s}}}{\sqrt{\hat{\boldsymbol{\beta}} + \delta}}. \quad (19)$$

Therefore, the new parameter becomes

$$\boldsymbol{\theta}^{new} = \boldsymbol{\theta} + \Delta \boldsymbol{\theta}. \quad (20)$$

The processes from Eq. (14) to Eq. (20) are repeated continuously until the parameter $\boldsymbol{\theta}$ converges. The converged DLNN is then implemented in the MDM optical transmission system for MIMO detection as in Fig. 4.1.

The computation complexity of our feedforward neural network during forward-propagation, assuming there are A numbers of neuron in each layer, is $O(MA^3)$. The complexity for backpropagation, assuming the gradient descent runs for I iterations, is $O(MIA^4)$, so it is essential to split the computation up into a training and inference phase. Although higher baud rate increases the speed of data availability at the input of the DLNN, this baud speed cannot outrun the processing speed of multiple Central Processing Units (CPUs) during the inference phase. Because of splitting the training and inference phase, there is no significant impact of higher baud rate on the performance and complexity of DLNN. For MIMO processing, we split the received signal into the real and imaginary parts. For $N \times N$ MIMO system, there are $2N$ input neurons and the output layer also has $2N$ output neurons. This shows that the complexity of DLNN increases with polynomial time complexity with MIMO dimension.

4.3 Training of DLNN

Nowadays, a variety of tools and algorithms are available to build and train large neural networks. Numerous tools ranging from a high-level language to a massive parallel Graphics Processing Units (GPU) architectures are the widely used for the training purpose. Out of many existing Deep Learning (DL) libraries [25-28], we have used TensorFlow [27] for designing and training our neural network. TensorFlow supports automatic differentiation of training cost function through arbitrarily large networks. It also allows us to distribute computation across different computers, multiple Central Processing Units (CPUs) and GPUs within a single machine.

In our DLNN, the output of each neuron should be permitted to assume both positive and negative values in the interval $[-1,1]$. A hyperbolic tangent (\tanh) function squashes real-valued number to the range between -1 and +1, i.e., $\tanh(x) \in (-1,1)$ [22,29]. Here, the negative and positive inputs are mapped strongly negative and positive respectively, and the zero inputs near zero in the \tanh graph. The function is monotonic while its derivative is not monotonic. Therefore, we have used a hyperbolic tangent as an activation function in our DLNN.

4.4 Experimental setup

The experimental setup to confirm efficient operation of MDM optical transmission system which uses DLNN for MIMO detection is shown in Fig. 4.4.

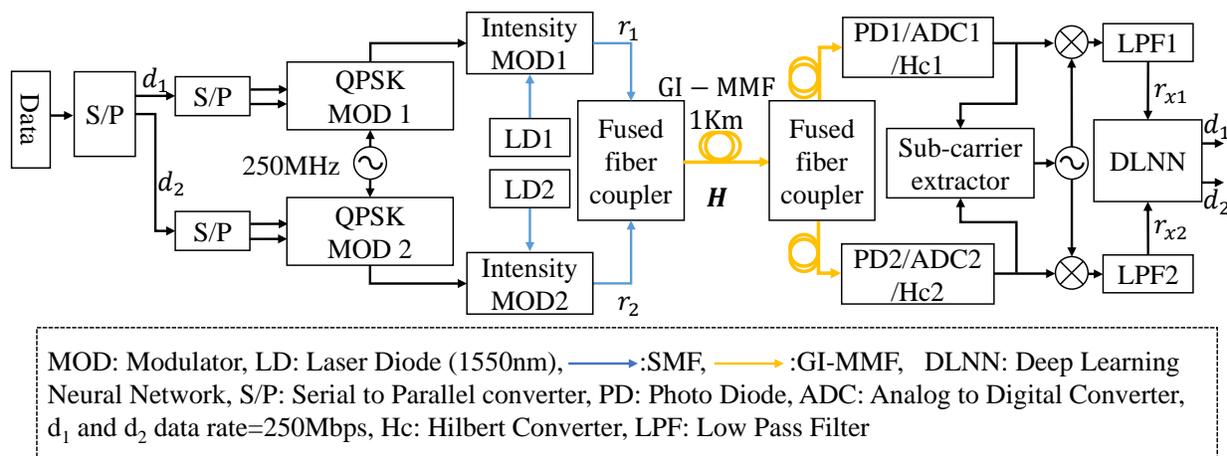


Fig. 4.4: A 2×2 MDM optical transmission system that uses DLNN for MIMO detection

MATLAB is used to generate random data. Using a serial to parallel converter, we divide this data into two channels: d_1 and d_2 . Both the channels operate at 125MBaud with Quadrature Phase Shift Keying (QPSK). SCM is done on both the channels using a 250MHz sub-carrier frequency, and then they are optically modulated using a Mach-Zehnder Modulator (MZM). Two Distributed Feedback Laser Diode (DFB-LD) operating at $1.55\mu\text{m}$, are used. The center frequency of these DFB-LDs is stabilized by controlling their temperature. A Single Mode Fiber (SMF) is used just before the mode-dependent fused fiber coupler, in the transmitter, to achieve non-uniform mode excitation. Two optical signals are coupled to a 1km long conventional GI-MMF OM2 with $50\mu\text{m}$ core diameter by using a mode dependent fused fiber coupler. At the other end, another mode-dependent fused fiber coupler optically divides the signal into two branches. These couplers are commercially available fused fiber couplers.

Then, signals are detected using photodetectors and sampled at the rate of 2Gsamples/s using analog to digital converter. The digital signals are Hilbert converted and sub-carrier is extracted. After then the signals are demodulated. Each demodulated signal includes two channel signals. In order to confirm both the excited modes are received at the receivers, we did some pre-testing experiment. Only for this purpose, we modify the Fig. 4.4 so that the random sequence d_2 is the 3-bit shifted version of d_1 . We performed a cross-correlation between the received signal and the reference signal on both the ports: r_{x1} and r_{x2} . Figure 4.5 confirms that both the modes that are excited are received.

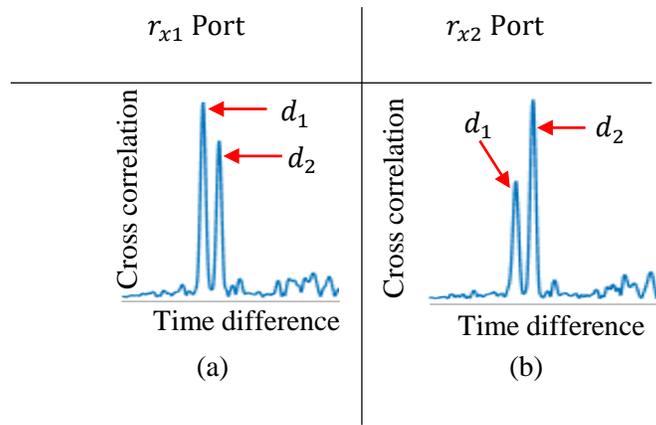


Fig. 4.5: Confirms both the modes that are excited are received.

Before designing a DLNN, we developed two datasets. The datasets are built from a continuous synchronous transmission of \mathbf{d} and reception of \mathbf{R}_x . Continuous generation of \mathbf{d} helped to develop \mathbf{y} vector, while \mathbf{R}_x helped to develop \mathbf{x} vector. This results in a dataset (\mathbf{x}, \mathbf{y}) , where \mathbf{y} is the desired output of our DLNN for \mathbf{x} input. Two datasets are built, one for training and another for testing purpose.

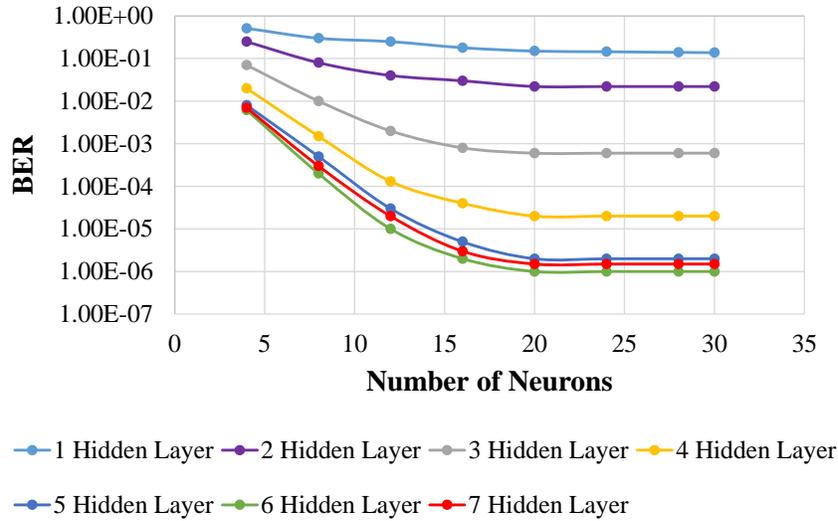


Fig. 4.6: Performance of DLNN at a different number of hidden layers and neurons.

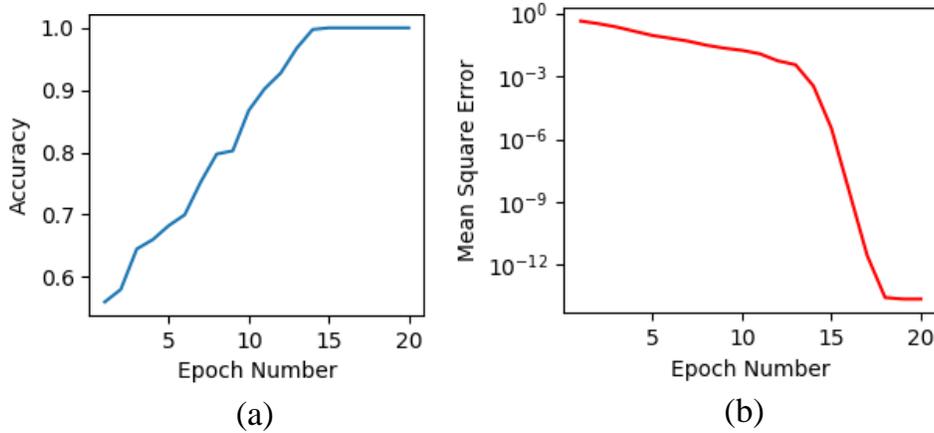


Fig. 4.7: Convergence behavior of DLNN: (a) accuracy versus epoch number, (b) mean square error versus epoch number.

In order to find the optimum number of hidden layers and neurons required for our DLNN, we developed numerous neural networks with a different number of hidden layers and neurons, and trained them. The performance of different neural networks designed with a different number of hidden layers and neurons is shown in Fig. 4.6. The BER is improved with increasing hidden neurons because the behavior of multimode fiber is highly non-linear. To represent this non-linearity, we need to increase the number of hidden neurons. The capacity to learn the underlying pattern increases with the increase in hidden neurons. We can see the performance of a neural network that has 6 hidden layers with 20 neurons in each hidden layer is acceptable. We can also see that the performance starts to degrade when we increase the number of hidden layers beyond

6. The overall performance of DLNN for MIMO detection is acceptable when it has six hidden layers with 20 neurons in each hidden layer.

Before training, data is received at ten different channel condition numbers (CN): 3.5dB, 5.1dB, 8dB, 10.9dB, 13dB, 14.9dB, 16.3dB, 17.5dB, 18.6dB, 19.5dB. The received vector \mathbf{R}_x is categorized into 10 different group based on these channel CNs. While assigning the group, the nearest CN value out of these ten values is selected. 2 million symbols, 200,000 symbols from each group of channel CNs, is used to develop a training dataset (\mathbf{x}, \mathbf{y}) to learn the mapping function from the input to the output. In our experiment, we changed the channel CNs by moving and bending the fiber. We have used ZF technique to calculate channel matrixes.

During training, we have information about the transmitted symbols; therefore, we know to which value each received symbols should be mapped. This mapping information is stored in a codebook. In this experiment we have used a 2×2 MDM optical transmission system; as a result, there are four input variables and four output variables. For example, if a symbol transmitted from QPSK Modulator 1 is $q_1 = +1 - j1$ and from QPSK Modulator 2 is $q_2 = -1 + j1$, the codebook maps the received symbols $r_{x1}(t) = a + jb$ and $r_{x2}(t) = c + jd$ to the expected value $+1-1-1+1$. This expected value, called a code word, is the desired output of our DLNN.

The convergence behavior of our DLNN having 6 hidden layers with 20 neurons in each hidden layer is also analyzed. Figure 4.7 shows its learning characteristics. When our DLNN is trained using a mini-batch size of 50, it converges around epoch number 15. The converged model is saved to perform MIMO detection. By training the DLNN at various channel CNs, its effectiveness for MIMO detection can be improved further. The computational complexity of our DLNN during forward-propagation and backpropagation are $O(n^2m + m^3 \cdot (l - 2))$ and $O(I \cdot (n^3m + m^4 \cdot (l - 2)))$ respectively, where n is the length of the input vector, m is the number of neurons in each hidden layer, l is the number of layers in DLNN and I is the number of gradient iterations.

The converged network is then tested by implementing it in the MDM optical transmission system as shown in Fig. 4.4. The received vector \mathbf{R}_x with a received optical power above -3.0dBm is used for testing. Testing is done at CN=13dB because it is approximately the middle value in the CN vector that we have taken.

4.5 Results

The converged DLNN model is implemented in the MDM optical transmission system to perform MIMO detection at various channel CNs. We have used different datasets for training and testing purpose. These datasets are built from a continuous synchronous transmission and reception. The continuous black line in Fig. 4.8 shows the Bit Error Rate (BER) versus channel CN for DLNN.

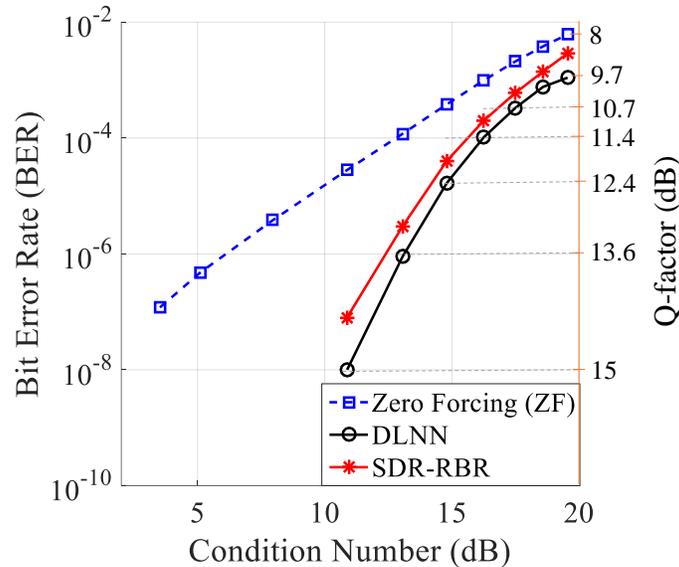


Fig. 4.8: Bit error rate versus channel condition number for our DLNN, ZF, and SDR-RBR.

The performance of our 2×2 MDM optical transmission system is also evaluated at different channel CN using a traditional ZF detector and an advanced SDR. Specifically, we have used Row-by-Row (RBR) method, called as SDR-RBR with 10 iterations. To make a performance comparison between the ZF, SDR-RBR and our DLNN, we have used the same input vector \mathbf{d} and received vector \mathbf{R}_x that we have used for building test dataset for testing the performance of our DLNN. Performance comparison is also done in terms of Quality factor (Q-factor) in dB as: $Q = 20 \log \left[\sqrt{2} \operatorname{erfc}^{-1} (2 * BER) \right]$, where erfc^{-1} is the inverse complementary error function. The relationship between Q-factor and Error Vector Magnitude (EVM) for QPSK is $Q = EVM^{-1}$.

The dotted blue line in Fig. 4.8 shows the BER versus channel CN for ZF detector. At CN=3.5dB, the minimum Quality factor (Q-factor) of this system is 14.3dB. At CN=19.5dB, the maximum Q-factor is 9.7dB using DLNN. Similarly, the continuous red line in Fig. 4.8 shows the BER versus

channel CN for SDR-RBR detector. The performance of SDR-RBR cannot catch up with the performance of DLNN.

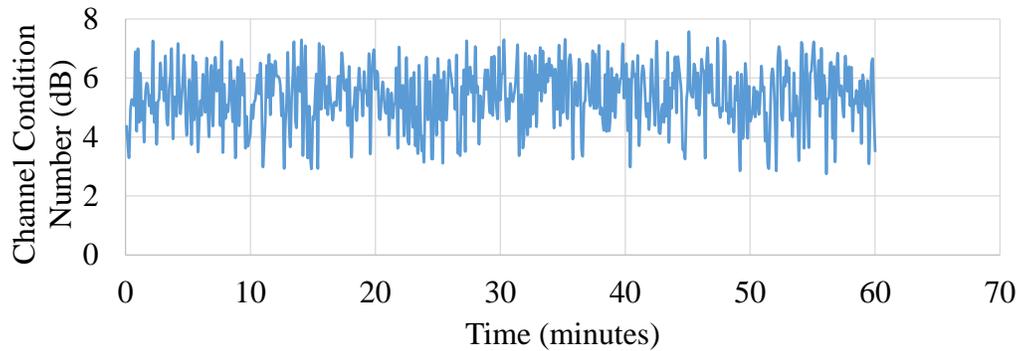


Fig. 4.9: The excursion of the channel condition number

At CN=10.9dB, the minimum Q-factor achieved with SDR-RBR detector is 14.4dB; however, the use of DLNN virtually increases the Q-factor of the MDM optical transmission system to greater than 15dB. This virtual increment in the Q-factor is the unique advantage of using DLNN. This result shows that DLNN based MIMO detection has better performance over the ZF detector and SDR-RBR.

4.6 Excursion of channel condition number

Furthermore, the performance of our DLNN model is also tested against the time-varying MIMO channel, where the channel CNs varies randomly with time. The excursion of the channel condition number is shown in Fig. 4.9. This figure shows that the CNs of MIMO channel lies within the range of (2.5dB, 8dB). To test the performance of MIMO system above channel condition number 8dB, the fiber is moved randomly. The result in Fig. 4.10 shows that our DLNN model has satisfactory performance over time-varying MIMO channel.

These results confirm that our DLNN for MIMO detection outruns the performance of traditional ZF detector and an advanced SDR-RBR. These are the powerful results which hold enormous promise for use in future optical communication systems. This technique provides a novel MIMO detection scheme and significantly improve the BER performance as compared to current day MIMO detectors. The training time of our DLNN was around 4 minutes using Intel(R) Core (TM) i7-4790CPU @3.6GHz.

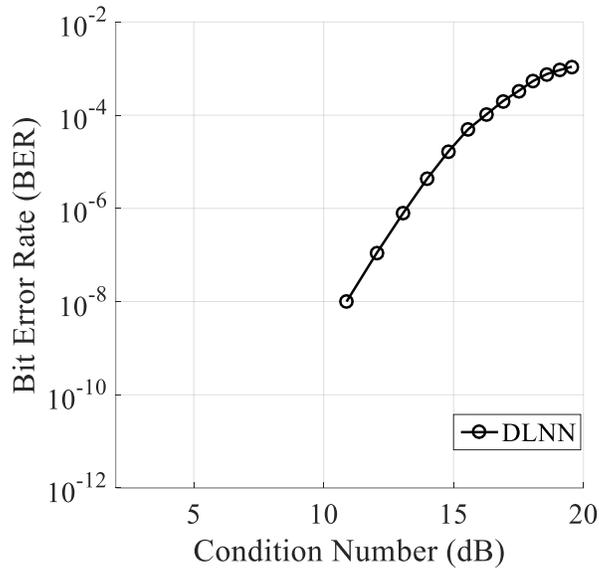


Fig. 4.10: Performance of DLNN at different channel condition.

4.7 Conclusion

The results shown here provides a new approach to novel MIMO detection for MDM optical transmission systems. This experiment proves that we can use the Deep Learning Neural Network for MIMO detection in the real environment.

Our results can be a remarkable achievement towards the use of a neural network for MIMO optical communication systems. The above results clearly show that the performance of DLNN can be very competitive with existing MIMO processing techniques. This model can be extended to realize higher order MDM optical transmission system. The use of DLNN for MIMO detection jointly optimizes the MDM optical transmission systems and provides better results compared to the existing systems.

4.8 References

- [1] R. W. Tkach, Scaling optical communications for the next decade and beyond, Bell Labs Tech. J. vol. 14, no. 4, (2010) 3-9.
- [2] D. J. Richardson, J. M. Fini, and L. E. Nelson, Space-division multiplexing in optical fibres, Nat. Photonics vol. 7, no. 5, (2013) 354–362.
- [3] R. Ryf, R. Essiambre, A. Gnauck, S. Randel, M. A. Mestre, C. Schmidt, P. Winzer, R. Delbue, P. Pupalaiakis, A. Sureka, T. Hayashi, T. Taru, and T. Sasaki, Space-division multiplexed

transmission over 4200 km 3-core microstructured fiber, in Proc. OFC/NFOEC (2012), paper PDP5C.2.

- [4] S. Chandrasekhar, A. Gnauck, X. Liu, P. Winzer, Y. Pan, E. C. Burrows, B. Zhu, T. Taunay, M. Fishteyn, M. Yan, J. M. Fini, E. Monberg, and F. Dimarcello, WDM/SDM transmission of 10 x 128-Gb/s PDM-QPSK over 2688-km 7-core fiber with a per-fiber net aggregate spectral-efficiency distance product of 40,320 km.b/s/Hz, in Proc. ECOC (2011), paper Th.13.C.4..
- [5] H. Takahashi, T. Tsuritani, E. L. T. de Gabory, T. Ito, W. R. Peng, K. Igarashi, K. Takeshima, Y. Kawaguchi, I. Morita, Y. Tsuchida, Y. Mimura, K. Maeda, T. Saito, K. Watanabe, K. Imamura, R. Sugizaki, and M. Suzuki, First demonstration of MC-EDFA-repeated SDM transmission of 40 x 128-Gbit/s PDM-QPSK signals per core over 6,160-km 7-core MCF, *Opt. Express* vol. 21, no. 1, (2013) 789–795.
- [6] A. Al Amin, A. Li, X. Chen, and W. Shieh, LP01/LP11 dual-mode and dual-polarization CO-OFDM transmission on two-mode fibre, *Electron. Lett.* vol. 47, no. 10, (2011) 606–608.
- [7] R. Ryf, S. Randel, A. H. Gnauck, C. Bolle, A. Sierra, S. Mumtaz, M. Esmaelpour, E. C. Burrows, R. Essiambre, P. J. Winzer, D. W. Peckham, A. H. McCurdy, and R. Lingle, Mode-division multiplexing over 96 km of fewmode fiber using coherent 6 x 6 MIMO Processing, *J. Lightwave Technol.* vol. 30, no. 4, (2012) 521–531.
- [8] E. Ip, G. Milione, M. J. Li, N. Cvijetic, K. Kanonakis, J. Stone, G. Peng, X. Prieto, C. Montero, V. Moreno, and J. Liñares, SDM transmission of real-time 10GbE traffic using commercial SFP + transceivers over 0.5km elliptical-core few-mode fiber, *Opt. Express* vol. 23, no. 13, (2015) 17120–17126.
- [9] X. Chen, J. Ye, Y. Xiao, A. Li, J. He, Q. Hu, and W. Shieh, Equalization of two-mode fiber based MIMO signals with larger receiver sets, *Opt. Express* vol. 20, no. 26, (2012) B413–B418.
- [10] L. Fang, L. Xu, and D. Huang, Low Complexity Iterative MMSE-PIC Detection for Medium-Size Massive MIMO, *IEEE Wireless Commun. Letters* vol. 5, no. 1, (2016) 108–111.
- [11] Z. Lei, Y. Dai, and S. Sun, Near Optimal List MIMO Detection, in Proceedings of IEEE conference on 16th International Symposium on Personal, Indoor and Mobile Radio Communications (IEEE, 2005), pp. 2441-2446
- [12] C. Jeon, R. Ghods, A. Maleki, and C. Studer, Optimality of large MIMO detection via approximate message passing, in Proceedings of IEEE conference on International Symposium on Information Theory (IEEE, 2015) 1227-1231.

- [13] Z. Q. Luo, W. K. Ma, A. M. So, Y. Ye, and S. Zhang, Semidefinite relaxation of quadratic optimization problems, *IEEE Signal Process. Magaz.* vol. 27, no. 3, (2010) 20-34.
- [14] J. C. Chen, A Low Complexity Data Detection Algorithm for Uplink Mutiuser Massive MIMO Systems, *J. Selected Areas in Commun.* vol. 35, no. 8, (2017) 1701-1714.
- [15] W. K. Ma, C. C. Su, J. Jalden, and C. Y. Chi, Some results on 16-QAM MIMO detection using Semidefinite Relaxation, in *Proceedings of IEEE International conference on Acoustics, Speech and Signal Processing (IEEE, 2008)* 2673-2676
- [16] T. O'Shea and J. Hoydis, An Introduction to Deep learning for the Physical Layer, *IEEE Transactions on Cognitive Commun. and Netw.* vol. 3, no. 4, (2017) 563-575.
- [17] T. O'Shea, T. Erpek, and T. C. Clancy, Physical Layer Deep Learning of Encodings for the MIMO Fading Channel, in *Proceedings of 55th Annual Allerton Conference on Communication, Control, and Computing*, N. Hovakimyan and N. Kiyavash (Allerton, 2017) 76-80.
- [18] A. Mecozzi, C. Antonelli, and M. Shtaif, "Coupled Manakov equations in multimode fibers with strongly coupled groups of modes," *Optics Express* vol. 20, no. 21, (2012), 23436-23441.
- [19] D. Marcuse, *Theory of Dielectric Optical Waveguides* (Academic Press, 1974)
- [20] K. Balemarchy, A. Polley and S. E. Ralph, Electronic Equalization of Multikilometer 10-Gb/s Multimode Fiber Links: Mode-Coupling Effects, *J. Lightw. Technol.* vol. 24, no. 12, (2006) 4885- 4894.
- [21] K.-P. Ho and J. M. Kahn, Mode-dependent loss and gain: statistics and effect on mode-division multiplexing, *Opt. Express* vol. 19, no. 17, (2011)16612-16635.
- [22] I. Goodfellow, Y. Bengio and A. Courville, *Deep Learning* (MIT Press, 2016).
- [23] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, Improving neural networks by preventing co-adaptation of feature detectors, <https://arxiv.org/pdf/1207.0580.pdf>
- [24] D. P. Kingma and J. L. Ba, Adam: A Method for Stochastic Optimization, <https://arxiv.org/abs/1412.6980>
- [25] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, Caffe: Convolutional Architecture for Fast Feature Embedding, in *Proceedings of the 22nd ACM international conference on Multimedia*, K. Hua, ed. (ACM, 2014) 675–678.

- [26] T. Chen, M. Li, Y. Li, M. Lin, N. Wang, M. Wang, T. Xiao, B. Xu, C. Zhang, Z. Zhang, MXNet: A Flexible and Efficient Machine Learning Library for Heterogeneous Distributed Systems, <https://arxiv.org/abs/1512.01274>.
- [27] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mane, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viegas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, X. Zheng, TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems, <https://arxiv.org/abs/1603.04467>
- [28] R. Al-Rfou, G. Alain, A. Almahairi, C. Angermueller, D. Bahdanau, N. Ballas, F. Bastien, J. Bayer, A. Belikov, A. Belopolsky, Y. Bengio, A. Bergeron, J. Bergstra, V. Bisson, J. B. Snyder, N. Bouchard, N. Boulanger-Lewandowski, X. Bouthillier, A. de Brebisson, O. Breuleux, P. Carrier, K. Cho, J. Chorowski, P. Christiano, T. Cooijmans, M. Cote, M. Cote, A. Courville, Y. N. Dauphin, O. Delalleau, J. Demouth, G. Desjardins, S. Dieleman, L. Dinh, M. Ducoffe, V. Dumoulin, S. E. Kahou, D. Erhan, Z. Fan, O. Firat, M. Germain, X. Glorot, I. Goodfellow, M. Graham, C. Gulceehre, P. Hamel, I. Harlouchet, J. Heng, B. Hidasi, S. Honari, A. Jain, S. Jean, K. Jia, M. Korobov, V. Kulkarni, A. Lamb, P. Lambin, E. Larsen, C. Laurent, S. Lee, S. Lefrancois, S. Lemieux, N. Leonard, Z. Lin, J. A. Livezey, C. Lorenz, J. Lowin, Q. Ma, P. Manzagol, O. Mastropietro, R. T. McGibbon, R. Memisevic, B. V. Merrienboer, V. Michalski, M. Mirza, A. Orlandi, C. Pal, R. Pascanu, M. Pezeshki, C. Raffel, D. Renshaw, M. Rocklin, A. Romero, M. Roth, P. Sadowski, J. Salvatier, F. Savard, J. Schluter, J. Schulman, G. Schwartz, I. V. Serban, D. Serdyuk, S. Shabanian, E. Simon, S. Spieckermann, S. R. Subramanyam, J. Sygnowski, J. Tanguay, G. V. Tulder, J. Turian, S. Urban, P. Vincent, F. Vinsin, H. D. Vries, D. W. Farley D. J. Webb, M. Willson, K. Xu, L. Yao, S. Zhang and Y. Zhang, Theano: A Python framework for fast computation of mathematical expressions, <https://arxiv.org/abs/1605.02688>
- [29] N. Srivastava, G. E. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, Dropout: A simple way to prevent neural networks from overfitting, *J. Mach. Learn. Research.* vol. 35, no. 1, (2014) 1929-1958.

Chapter 5. Mode Division Multiplexing MIMO Optical Transmission System that uses Deep Learning Neural Network for Signal Detection using only one Composite Signal

5.1 Proposed MDM optical network

The main objective of our proposed MDM optical network is to extract the desired data from one composite signal without using MIMO processing at the receiver end. Once a DLNN is trained to extract the desired data, other channel data cannot be extracted, so this type of system also provides security against sniffing.

The proposed network consists of a Sub-Carrier Multiplexers, Optical transmitters, fused fiber couplers to perform multiplexing operation, MMF or FMF, optical receivers and DLNNs. The proposed N channel MDM optical network is shown in Fig. 5.1. $d_1, d_2 \dots d_N$ are N different channels each with random data pattern. In each channel, subcarrier multiplexing is done and the multiplexed signal is feed to the Intensity Modulator. The N modulated optical signals are mode division multiplexed using a mode dependent fused fiber coupler and propagated through an MMF.

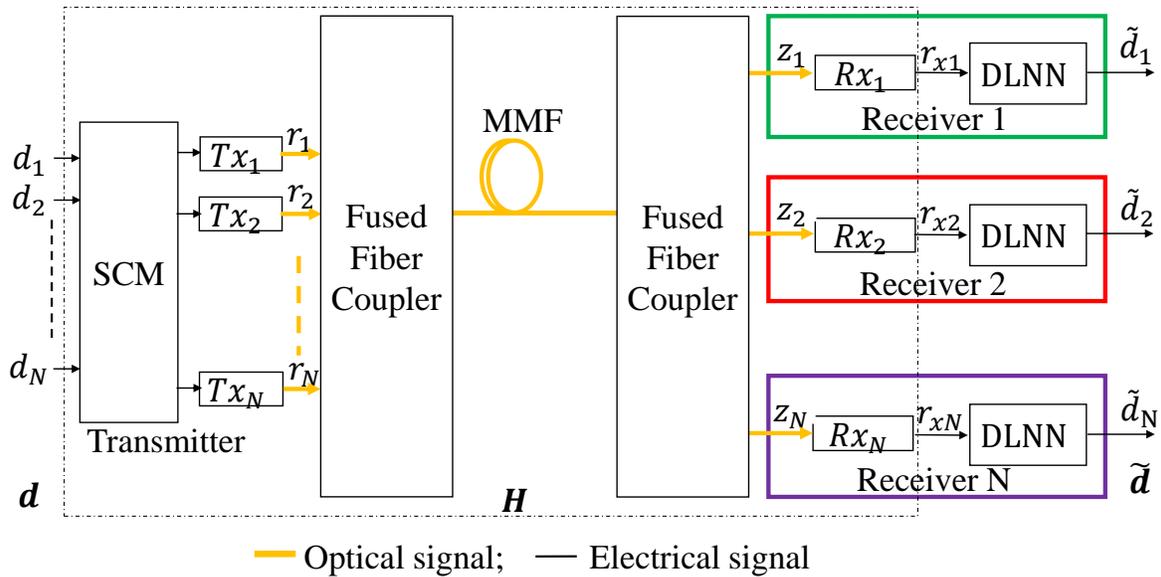


Fig. 5.1: Proposed MDM optical network

At the other end, another fused fiber coupler is used to divide the signal and take them to N different remote user locations. Each user is at a different location, so MIMO processing is not suitable to

separate the channels. Therefore, a new technique has been used to extract the desired channel from one composite signal.

In this technique, signal processing is done in each remote user location to extract the desired channel. In each remote user location, there is a receiver which consists of a DLNN. That particular DLNN is responsible for extracting the desired channel from one composite signal.

5.2 Design of a deep learning neural network

A N mode MDM optical transmission system with input $\mathbf{d}(t)=[d_1(t),d_2(t),d_3(t),d_4(t),\dots,d_N(t)]^T \in \mathbb{B}^N$ produces the output $\tilde{\mathbf{d}}(t) \in \mathbb{B}^N$ where the output on the i th channel is

$$\tilde{d}_i(t) = \mathbf{W}\mathbf{H}^{-1}\mathbf{H}\mathbf{d}(t) \quad (1)$$

where \mathbb{B} represents binary numbers set; $\mathbf{H} \in \mathbb{C}^{N \times N}$ represents a channel matrix, where \mathbb{C} represents the complex numbers set; $\mathbf{W}\mathbf{H}^{-1}$ represents the response of DLNN, and

$$\mathbf{W} = [w_1, w_2, w_3, \dots, w_N], \text{ where } w_k = \begin{cases} 1 & \text{for } k = i \\ 0 & \text{otherwise} \end{cases}.$$

Figure 5.2 shows the signal processing technique on the i th channel of Fig. 5.1 from transmitter to the input of DLNN. Inside the SCM block of Fig. 5.1, there is a bank of N Binary Phase Shift Keying (BPSK) modulators with a sub-carrier frequency generator. f_c is the sub-carrier frequency. The TX_i block consists of an optical carrier and MZM (Mach-Zehnder Modulator). On the i th data channel, the n th input bit is denoted as $d_{i,n}$, where $n= 0, 1, 2, \dots, \infty$. The electrical BPSK modulated signal on the i th channel is

$$a_i(t) = q_i e^{j2\pi f_c t} \quad (2)$$

where $q_i = (d_{i,n})$ denotes a BPSK symbol. This BPSK signal modulates the Intensity of an optical carrier. As we are using MDM, all the optical carriers operate at the same wavelength. The Intensity modulator on the i th channel produces an output

$$r_i(t) = P_{bias} + P_i \text{Re}[a_i(t)] \quad (3)$$

where P_{bias} denotes the optical power due to a dc bias voltage applied on the Intensity Modulator, P_i denotes the amplitude of the optical carrier, and $\text{Re}[a_i(t)]$ denotes the real part of $a_i(t)$. The

bias voltage determines the operating point of Intensity Modulator. A mode dependent fused fiber coupler is used to couple the Intensity modulated signal to different modes of GI-MMF. Here, mode division multiplexing operation is performed by the fused fiber coupler. Partial mode coupling occurs between different modes during propagation through the GI-MMF.

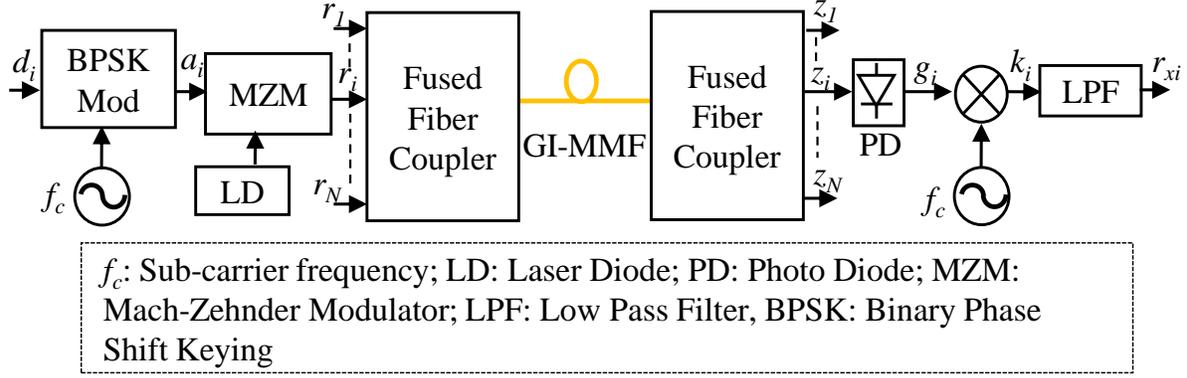


Fig. 5.2: Signal processing on i th channel

At the receiver side, after using another mode dependent fused fiber coupler, the signal received by the i th receiver is

$$z_i(t) = \sum_{j=1}^N r_j(t) h_{ij}(t) \quad (4)$$

where $h_{ij}(t)$ is a complex value and it denotes the (i,j) th element of total channel matrix \mathbf{H} . Absolute value of $h_{ij}(t)$ represents the power coupling and phases of $h_{ij}(t)$ represents the time differences of subcarrier. Equation 4 is a composite signal which contains signal from all the transmitters due to mode coupling during transmission. The photodetector on the i th channel generates a current

$$I_i(t) \propto z_i(t) = \sum_{j=1}^N r_j(t) h_{ij}(t) \quad (5)$$

The voltage $g_i(t)$ produced by the current $I_i(t)$ is

$$g_i(t) = \sum_{j=1}^N P_j h_{ij}(t) [a_j(t)] \quad (6)$$

The dc value P_{bias} in (3) disappears after passing through a band pass filter. The signal $k_i(t)$ is obtained after multiplying (6) by sub-carrier signal as

$$k_i(t) = g_i(t)e^{-j2\pi f_c t} = \left[\sum_{j=1}^N P_j h_{ij}(t) [a_j(t)] \right] e^{-j2\pi f_c t} \quad (7)$$

The signal on the i th port after passing through a low pass filter is

$$r_{xi}(t) = \sum_{j=1}^N P_j h_{ij}(t) q_j \quad (8)$$

This received signal is a composite signal and is a complex value. Training dataset and test datasets are built, after synchronous transmission of d_i and reception of r_{xi} .

The structure of our DLNN with M layers is shown in Fig. 5.3. The input vector to our DLNN is given as

$$\mathbf{x}_0 = [\text{Re}(r_{xi}(t)), \text{Im}(r_{xi}(t))] \quad (9)$$

The length of \mathbf{x}_0 is $D=2$. The mapping operation of this feedforward neural network is the same as described in our previous paper [16]. The difference is in the number of neurons in the input and output layers. In this paper, we have used BPSK modulation, so only two types of symbols are possible; +1 and -1. Since we are extracting only the desired signal in each receiver, one output neuron is sufficient to represent BPSK symbols. So, our feedforward neural network has two neurons in the input layer and only one neuron in the output layer. The output layer uses ReLU [18] as an activation function [19-25]. Whereas, in all other layers tanh [18] is used as an activation function. For training DLNN on the i th receiver, a dataset $(\mathbf{x}_i^k, \mathbf{y}_i^k)$, $k = 1, \dots, S$, where \mathbf{y}_i^k is the desired output for \mathbf{x}_i^k input and S is the number of training examples, is first developed and used. Channel condition number (CN) indicates channel correlation property and gives information about the difficulties of recovering MIMO signals. The CN of a channel with channel matrix \mathbf{H} is $20 \log_{10}(\|\mathbf{H}\| \cdot \|\mathbf{H}^{-1}\|) \geq 0\text{dB}$, where $\|\cdot\|$ is the norm. The quality of the MIMO channel increases when CN approaches 0dB. The matrix \mathbf{H} is the total channel matrix and it includes the characteristics of GI-MMF as well as two fused fiber couplers. During propagation through the fiber, modal dispersion (MD), Mode Dependent loss (MDL), crosstalk due to mode coupling, and other mode-dependent effects exist.

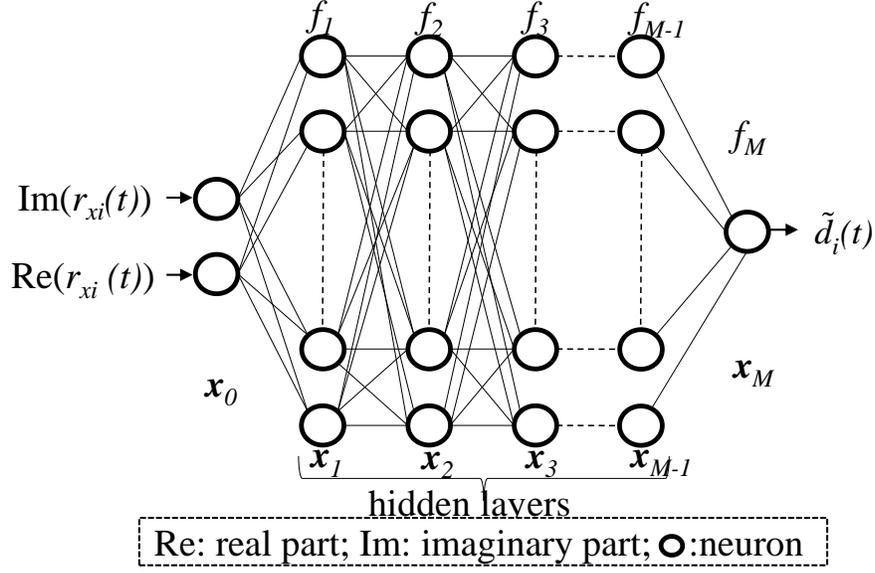


Fig. 5.3: Our deep learning neural network structure

For the detection of the desired signal from one composite signal in real-time, DLNN must be trained at various CNs. The cost of our DLNN is

$$J(\boldsymbol{\theta}) = \frac{1}{S} \sum_{k=1}^S L(f(\mathbf{x}_i^k; \boldsymbol{\theta}), \mathbf{y}_i^k) \quad (10)$$

where $L(\mathbf{u}, \mathbf{v})$ denotes the per-example cost function and $f(\mathbf{x}_i^k; \boldsymbol{\theta})$ denotes the actual output of the neural network at the input \mathbf{x}_i^k . The accuracy of our model is determined by the cost function. Training helps to minimize the cost $J(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$. Mean Square Error (MSE) has only one global minimum value, so if we use it as a cost function, then (10) becomes

$$J(\boldsymbol{\theta}) = \frac{1}{S} \sum_{k=1}^S \|f(\mathbf{x}_i^k; \boldsymbol{\theta}) - \mathbf{y}_i^k\|_2^2 \quad (11)$$

Various type of optimization algorithm exists [18], but the most popular one is Adam [20]. Implementation of Adam and how the parameter $\boldsymbol{\theta}$ is updated is well explained in our previous work [16].

Similarly, all other DLNN in each receiver unit is trained until they converge. All the converged DLNNs are implemented for signal detection as in Fig. 5.1.

The computation complexity of our DLNN is $O(MA^3)$ during forward propagation and $O(MIA^4)$ during backpropagation; where M represents the number of layers, A represents the number of

neurons in each hidden layer, and I represents the iteration number of gradient descent.

5.3 Experimental setup

The experimental setup to separate the desired signal from one composite signal in MDM optical transmission system is shown in Fig. 5.4.

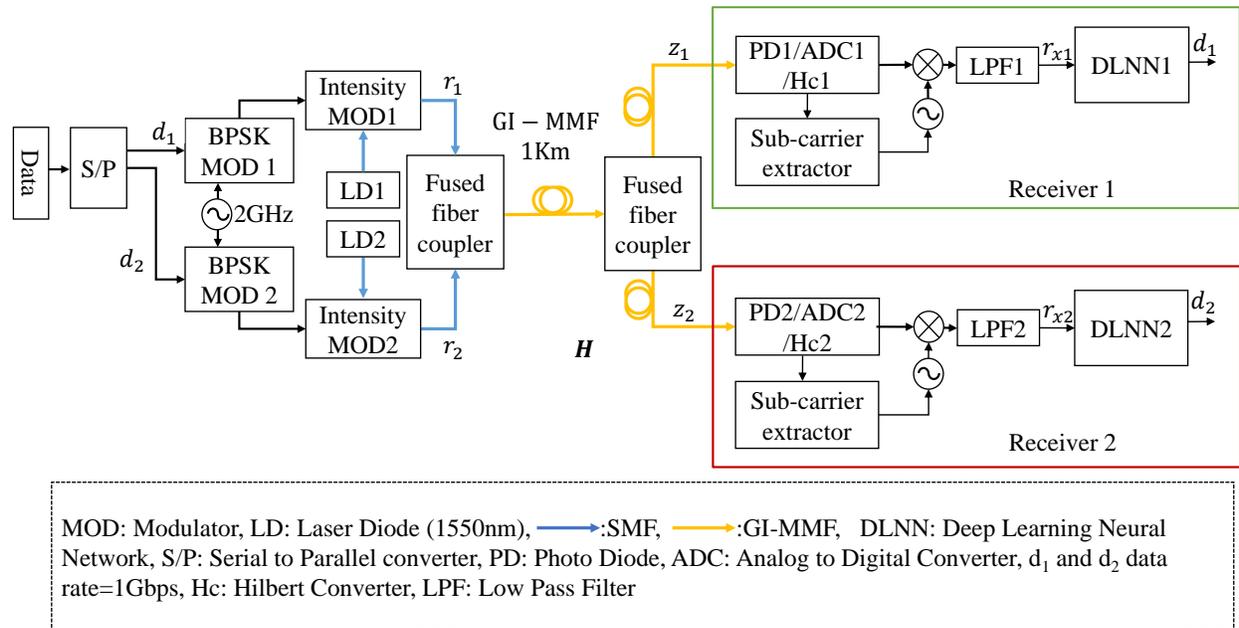


Fig. 5.4: MDM optical transmission system for the desired signal extraction from one composite signal

Random data is generated using MATLAB. Both the channels d_1 and d_2 operate at 1GBaud with Binary Phase Shift Keying (BPSK). 2GHz sub-carrier is used for SCM on both the channels and optical modulation is done using Mach-Zehnder Modulator (MZM). We have used a Distributed Feedback Laser Diode (DFB-LD) to generate an optical carrier of $1.55\mu\text{m}$. Frequency stabilization of DFB-LDs is done by controlling their temperature. The use of a Single Mode Fiber (SMF) in between the MZM and the mode dependent fused fiber coupler provides a non-uniform mode excitation. The two optical channels are coupled to two different modes of a conventional OM2 GI-MMF. Another mode dependent fused fiber on the other end of GI-MMF divides the transmitted optical signals into two signals: z_1 and z_2 . Receiver 1 receives the z_1 signal, whereas, Receiver 2 receives the z_2 . Receiver 1 and 2 are located at two different remote locations. After photo-detection, sampling is done at 20Gsamples/s and then Hilbert conversion is done. Sub-carrier extraction is also done in each receiver. After demodulation, we get a composite signal

which contains both the data d_1 and d_2 . The composite signal in Receiver 1 is r_{x1} , and in Receiver 2 is r_{x2} .

We did some initial experiment to confirm that both the excited modes are properly received at all remote user locations. The input data d_1 and d_2 in Fig. 5.4 is slightly modified only for this purpose. We made d_2 from d_1 by using a 3-bit delay circuit. Cross-correlation between the reference signal and the received signal are performed on both the receiver. Figure 5.5 confirms that both the excited modes are received.

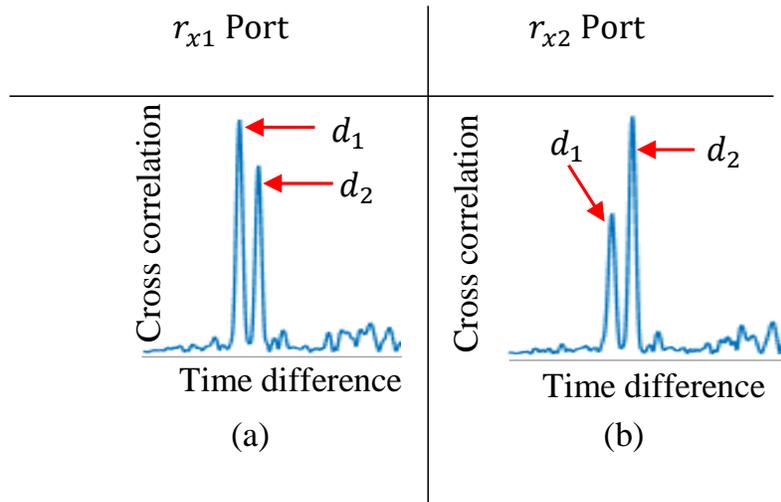


Fig. 5.5: Both the excited modes are received

The two DLNNs in the receiver should be trained separately. Training dataset and testing dataset for DLNN1 are built from a continuous synchronous transmission of d_1 and reception of r_{x1} . Similarly, the other two datasets are built for DLNN2 from d_2 and r_{x2} . Continuous generation of d_i helped to develop y_i vector, while r_{xi} helped to develop x_i vector. This results in a dataset (x_i, y_i) , where y_i is the desired output of the DLNN on the i th receiver for x_i input.

The BER performance of the neural network varies with the number of hidden layers and neurons [16]. We have selected 3 hidden layers with 10 neurons in each hidden layer to design our DLNN for acceptable performance.

For training purpose, we received data at ten different channel condition numbers (CN): 3.5dB, 5.1dB, 8dB, 10.9dB, 13dB, 14.9dB, 16.3dB, 17.5dB, 18.6dB, 19.5dB. From each group, 100,000 symbols are taken to develop a training dataset (x_i, y_i) .

While training the DLNN, each symbol of the received composite signal is mapped to the symbol of the desired signal that was transmitted. A codebook stores this information. There are two input variables and one output variable in each DLNN. If a symbol transmitted from BPSK Modulator 1 is +1 the codebook in the Receiver 1 maps the received symbol $r_{x1}(t) = a + jb$ to the expected value 1. Similarly, if a symbol transmitted from BPSK Modulator 1 is -1, the codebook in the Receiver 1 maps the received symbol to 0. The same process is applied in Receiver 2.

The convergence behavior of our DLNN with 3 hidden layers and 10 neurons in each hidden layer is also analyzed. Figure 5.6 shows its learning characteristics. We have used a mini-batch size of 50. The converged model is saved and implemented for signal detection in the remote user location. Training the DLNN at various channel CNs improves the ability of DLNN to separate the desired signal from the composite signal. Its computational complexity during forward-propagation is $O(n^2m + m^3 \cdot (l - 2))$, and during backpropagation is $O(I \cdot (n^3m + m^4 \cdot (l - 2)))$, where n denotes the length of the input vector, m denotes the number of neurons in each hidden layer, l denotes the number of layers in DLNN and I denotes the number of gradient iterations.

The converged network is implementing as shown in Fig. 5.4 for real-time signal detection. The optical power of the received signal is above -3.0dBm. The performance of our DLNN is tested at CN=13dB. This CN value is approximately the middle value in our CN vector, so we have selected this.

5.4 Results

The designed DLNN is implemented in each remote user location to separate the desired signal separation from a composite signal. For the purpose of training and testing, we have used different datasets in each location. The continuous black line in Fig. 5.7 shows the BER versus channel condition for DLNN. Whereas, the continuous blue line shows the performance of the ZF detector assuming both the users are located at the same place and MIMO detection is possible.

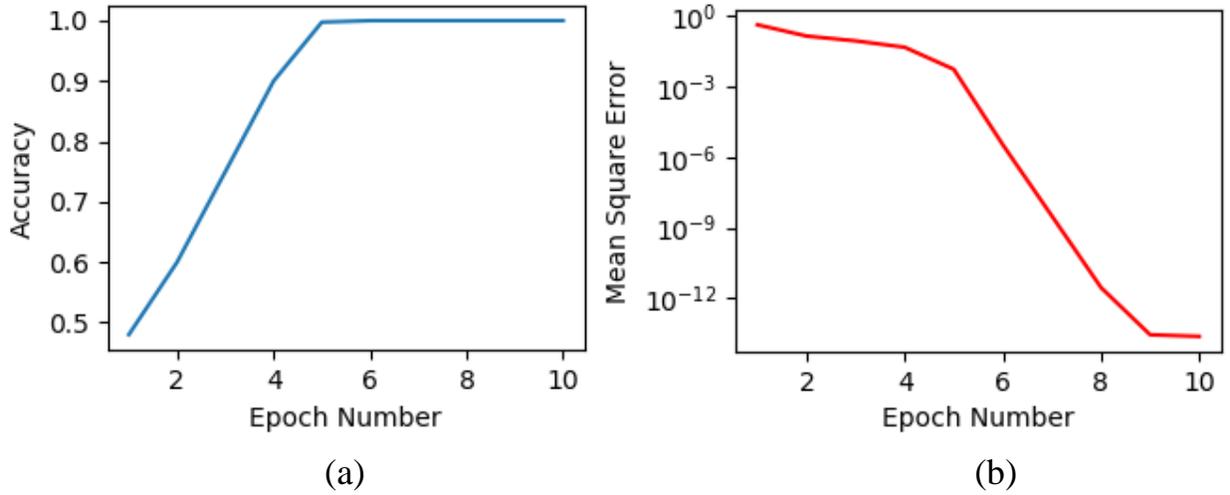


Fig. 5.6: Convergence behavior of DLNN: (a) accuracy versus epoch number, (b) mean square error versus epoch number.

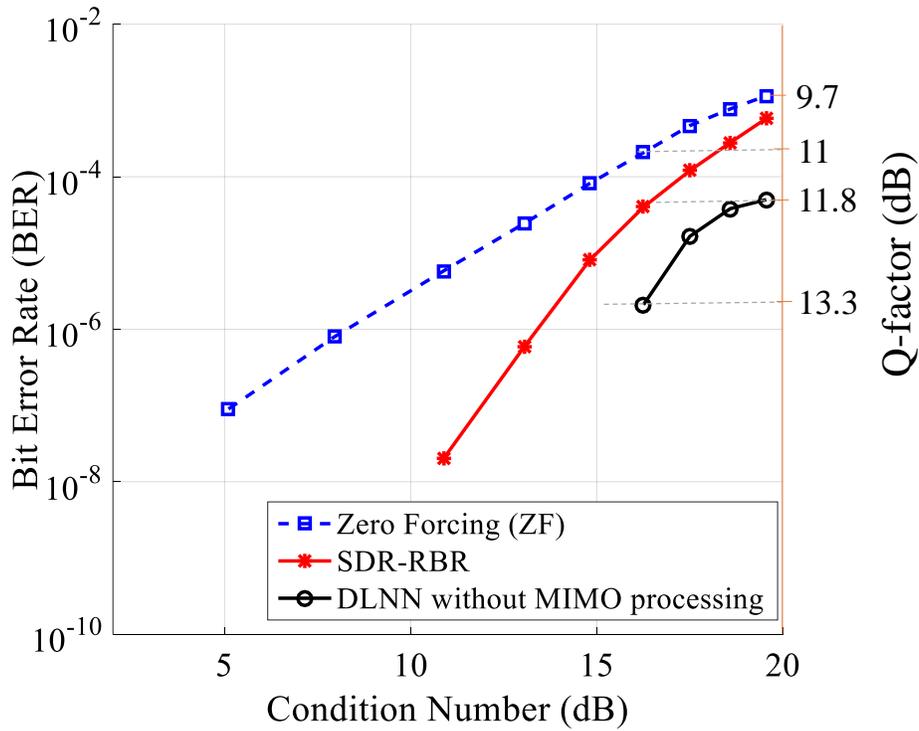


Fig. 5.7: Bit error rate versus channel condition number for our DLNN, and ZF

This clearly shows that we can separate the desired signal from a composite signal using a deep learning technique.

5.5 Conclusion

This result confirms that deep learning neural network can be used to separate the desired signal from a composite signal. This shows the enormous importance of deep learning in optical communication systems.

5.6 References

- [1] Cisco Visual Networking, “The Zettabyte Era—Trends and Analysis,” Cisco white paper (2017).
- [2] Cisco Visual Networking Index, “Forecast and Methodology,” Cisco white paper (2017).
- [3] Cisco Global Cloud Index, “Forecast and methodology 2016–2021,” Cisco Systems Inc., San Jose (2018).
- [4] T. Mizuno, K. Shibahara, H. Ono, Y. Abe, Y. Miyamoto, F. Ye, T. Morioka, Y. Sasaki, Y. Amma, K. Takenaga, S. Matsuo, K. Aikawa, K. Saitoh, Y. Jung, D. J. Richardson, K. Pulverer, M. Bohn, and M. Yamada, “32-core Dense SDM Unidirectional Transmission of PDM-16QAM Signals Over 1600 km Using Crosstalk-managed Single-mode Heterogeneous Multicore Transmission Line,” in *Proc. OFC*, Anaheim, CA, USA, 2016, Th5C.3
- [5] A. Turukhin, O. V. Sinkin, H. G. Batshon, H. Zhang, Y. Sun, M. Mazurczyk, C. R. Davidson, J. X. Cai, M. A. Bolshtyansky, D. G. Foursa, and A. Pilipetskii, “105.1 Tb/s Power-Efficient Transmission over 14350 km using a 12-Core Fiber”, in *Proc. OFC*, Anaheim, CA, USA, 2016, Th4C.1.
- [6] K. Igarashi, T. Tsuritani, I. Morita, Y. Tsuchida, K. Maeda, M. Tadakuma, T. Saito, K. Watanabe, K. Imamura, R. Sugizaki, and M. Suzuki, “1.03-Exabit/s·km Super-Nyquist-WDM transmission over 7326-km seven-core fiber”, in *Proc. ECOC*, London, UK, 2013, PD3-E-3.
- [7] D. J. Richardson, J. M. Fini, L.E. Nelson, “Space-division multiplexing in optical fibres,” *Nat. Photonics*, vol. 7, no. 5, pp. 354-362, 2013
- [8] R. Ryf, R. J. Essiambre, A. H. Gnauck, S. Randel, M. A. Mestre, C. Schmidt, P. J Winzer, R. Delbue, P. Pupalaikis, A. Sureka, T. Hayashi, T. Taru, T. Sasaki, “Space-division multiplexed transmission over 4200 km 3-core microstructured fiber,” in *Proc. OFC/NFOEC*, Los Angeles, CA, USA, March(2012) paper PDP5C.2
- [9] S. Chandrasekhar, A. H. Gnauck, X. Liu, P. J. Winzer, Y. Pan, E. C. Burrows, B. Zhu, M. Fishteyn, M. F. Yan, J. M. Fini, E. M. Monberg, and F. V. Dimarcello, “WDM/SDM transmission of 10 x 128-Gb/s PDM-QPSK over 2688-km 7-core fiber with a per-fiber net

- aggregate spectral-efficiency distance product of 40, 320 km.b/s/Hz,” *Optics Express*, vol. 20, no. 2, pp. 706-711, Jan. 2012
- [10] H. Takahashi, T. Tsuritani, E. L. T. de Gabory, T. Ito, W. R. Peng, K. Igarashi, K. Takeshima, Y. Kawaguchi, I. Morita, Y. Tsuchida, Y. Mimura, K. Maeda, T. Saito, K. Watanabe, K. Imamura, R. Sugizaki, M. Suzuki, “First demonstration of MC-EDFA-repeated SDM transmission of 40 x 128-Gbit/s PDM-QPSK signals per core over 6, 160-km 7-core MCF,” *Opt. Express*, vol. 21, no. 1, pp. 789-795, Jan. 2013
- [11] K. Nakajima, T. Matsui, K. Saito, T. Sakamoto, and N. Araki, “Space division multiplexing: Next generation optical communication strategy,” in *Proc. ITU Kaleidoscope: ICTs for a Sustainable World*, Bangkok, Thailand, Nov. 2016.
- [12] K. Kitayama, N. P. Diamantopoulos, Y. Yoshida, and A. Maruta, “Mode Division Multiplexed Networks,” in *Proc. OECC and PGC*, Singapore, July 2017.
- [13] J. Luo, J. Li, Q. Sui, Z. Li, and C. Lu, “40Gb/s Mode-Division Multiplexed DD-OFDM Transmission Over Standard Multi-Mode Fiber,” *IEEE Photonics Journal*, vol. 8, no. 3, June 2016
- [14] E. IP, G. Milione, M. Li, N. Cvijetic, K. Kanonakis, J. Stone, G. Peng, X. Prieto, C. Montero, V. Moreno, and J. Linares, “SDM transmission of real-time 10GbE Traffic using commercial SFP + transceivers over 0.5 km elliptical-core few-mode fiber,” *Opt. Express*, vol. 23, no. 13, pp. 17120-17126, June 2015.
- [15] A. Mecozzi, C. Antonelli, and M. Shtaif, "Coupled Manakov equations in multimode fibers with strongly coupled groups of modes," *Opt. Express*, vol. 20, no. 21, pp. 23436-23441, Sep. 2012
- [16] B. Poudel, J. Oshima, H. Kobayashi, and K. Iwashita, "MIMO detection using a deep learning neural network in a mode division multiplexing optical transmission system,” *Optics Communications*, vol. 440, pp. 41-48, Feb. 2019.
- [17] B. Poudel, J. Oshima, H. Kobayashi, and K. Iwashita, "Passive Optical Delivering Network Using Conventional Graded-Index Multi-Mode Fiber with Mode Division Multiplexing and Sub-Carrier Multiplexing," *Journal of Optical Communications and Networking*, vol. 10, no. 3, pp. 252-259, March 2018.
- [18] I. Goodfellow, Y. Bengio and A. Courville, *Deep Learning* (MIT Press, 2016).

- [19] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, Improving neural networks by preventing co-adaptation of feature detectors, <https://arxiv.org/pdf/1207.0580.pdf>
- [20] D. P. Kingma and J. L. Ba, Adam: A Method for Stochastic Optimization, <https://arxiv.org/abs/1412.6980>
- [21] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, Caffe: Convolutional Architecture for Fast Feature Embedding, in Proceedings of the 22nd ACM international conference on Multimedia, K. Hua, ed. (ACM, 2014) 675–678.
- [22] T. Chen, M. Li, Y. Li, M. Lin, N. Wang, M. Wang, T. Xiao, B. Xu, C. Zhang, Z. Zhang, MXNet: A Flexible and Efficient Machine Learning Library for Heterogeneous Distributed Systems, <https://arxiv.org/abs/1512.01274>.
- [23] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mane, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viegas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, X. Zheng, TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems, <https://arxiv.org/abs/1603.04467>
- [24] R. Al-Rfou, G. Alain, A. Almahairi, C. Angermueller, D. Bahdanau, N. Ballas, F. Bastien, J. Bayer, A. Belikov, A. Belopolsky, Y. Bengio, A. Bergeron, J. Bergstra, V. Bisson, J. B. Snyder, N. Bouchard, N. Boulanger-Lewandowski, X. Bouthillier, A. de Brebisson, O. Breuleux, P. Carrier, K. Cho, J. Chorowski, P. Christiano, T. Cooijmans, M. Cote, M. Cote, A. Courville, Y. N. Dauphin, O. Delalleau, J. Demouth, G. Desjardins, S. Dieleman, L. Dinh, M. Ducoffe, V. Dumoulin, S. E. Kahou, D. Erhan, Z. Fan, O. Firat, M. Germain, X. Glorot, I. Goodfellow, M. Graham, C. Gulceehre, P. Hamel, I. Harlouchet, J. Heng, B. Hidasi, S. Honari, A. Jain, S. Jean, K. Jia, M. Korobov, V. Kulkarni, A. Lamb, P. Lambin, E. Larsen, C. Laurent, S. Lee, S. Lefrancois, S. Lemieux, N. Leonard, Z. Lin, J. A. Livezey, C. Lorenz, J. Lowin, Q. Ma, P. Manzagol, O. Mastropietro, R. T. McGibbon, R. Memisevic, B. V. Merrienboer, V. Michalski, M. Mirza, A. Orlandi, C. Pal, R. Pascanu, M. Pezeshki, C. Raffel, D. Renshaw, M. Rocklin, A. Romero, M. Roth, P. Sadowski, J. Salvatier, F. Savard, J. Schluter, J. Schulman, G. Schwartz, I. V. Serban, D. Serdyuk, S. Shabanian, E. Simon, S. Spieckermann, S. R.

Subramanyam, J. Sygnowski, J. Tanguay, G. V. Tulder, J. Turian, S. Urban, P. Vincent, F. Vinsin, H. D. Vries, D. W. Farley D. J. Webb, M. Willson, K. Xu, L. Yao, S. Zhang and Y. Zhang, Theano: A Python framework for fast computation of mathematical expressions, <https://arxiv.org/abs/1605.02688>

- [25] N. Srivastava, G. E. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, Dropout: A simple way to prevent neural networks from overfitting, *J. Mach. Learn. Research.* vol. 35, no. 1, (2014) 1929-1958.

Chapter 6. Summary and Future Research Direction

6.1 Summary

In the context of using MDM for designing a spectrally efficient high speed and high capacity optical transmission system, we have studied the under-addressed MIMO-MDM system by adding Deep learning neural network with it.

MIMO processing at the receiver-end is not suitable if all the end-users are located at a separate location. Furthermore, in the existing optical network, each Optical Network Unit (ONU) receives not only the data that belongs to its user but also the data of other users. So even if data is encrypted, the network still remains vulnerable to sniffing. It raises security questions. In this scenario, we need to design an optical transmission network that not only doesn't use MIMO processing at the receiver end but also is resilient to sniffing. In Chapter 3, an MDM optical transmission network is designed using SCM and pre-MIMO processor to address these two issues. The results of Chapter 3 show that by continuously controlling the amplitude and phase of the transmitter signal, mode forming technique can be implemented in optical transmission system so that a MIMO processing is not required in the receiver side. The results also prove that our optical network supports the exchange of signal in the output port. One of the important outcomes of chapter 3 is that our mode forming network is resilient to sniffing.

Remembering the fact that none of the available MIMO detectors is an optimum detector, in Chapter 4 we develop a novel optimal MIMO detector using a Deep Learning technique. A novel optical MIMO detector is designed using a supervised DLNN and implemented in an MDM optical transmission system. We have trained our DLNN with data received at channel $CN_s < 20\text{dB}$. Once the network converges, it is implemented for MIMO processing. In our time-varying MIMO channel, the channel condition number was generally below 10dB. Hence, we did not update the weights until the channel condition number exceeds 10dB. Its weights should be updated much less than the symbol rate. DLNN uses a parallel architecture. The training time of our DLNN was around 4 minutes using Intel(R) Core (TM) i7-4790CPU @3.6GHz. Weights are continuously updated during the whole training process. The results of Chapter 4 provide a new approach to novel MIMO detection for MDM optical transmission systems. It confirms that our DLNN for MIMO detection outruns the performance of traditional ZF detector and an advanced SDR-RBR.

This experiment proves that we can use the Deep Learning Neural Network for MIMO detection in the real environment. The results can be a remarkable achievement towards the use of a neural network for MIMO optical communication systems. This model can be extended to realize higher order MDM optical transmission system. The use of DLNN for MIMO detection jointly optimizes the MDM optical transmission systems and provides better results compared to the existing systems.

We have also designed a DLNN that is capable of extracting the desired signal using only one composite signal in Chapter 5. This DLNN has been successfully implemented in an MDM optical transmission system with two channels operating at 1Gbps. The results of Chapter 5 clearly show that the performance of DLNN can be very competitive with existing MIMO processing techniques. These are the powerful results which hold enormous promise for use of DLNN in future optical communication systems.

6.2 Further works

When remote users are located at a different location, we cannot do MIMO processing because only one composite signal is available. We are designing a deep learning neural network to extract the desired signal by using only one composite signal and trying to improve the performance compared to the system demonstrated in Chapter 4..