Complexity reduction over bi-RNN-based Kerr nonlinearity equalization in dual-polarization fiber-optic communications via a CRNN-based approach

Abtin Shahkarami

INSTITUT POLYTECHNIQUE DE PARIS

TELECOM Paris

IP PARIS

NNT : 2022IPPAT034

Thèse de doctorat

Complexity Reduction over Bi-RNN-Based Kerr Nonlinearity Equalization in Dual-Polarization Fiber-Optic Communications via a CRNN-Based Approach

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à Telecom Paris

École doctorale n°626, Institut Polytechnique de Paris (ED IP Paris)
Spécialité de doctorat: Réseaux, Informations et Communications

Thèse présentée et soutenue à Palaiseau, le 10 octobre 2022, par

Abtin SHAHKARAMI

Composition du Jury :

Catherine LEPERS Professeure, Institut Polytechnique de Paris, Telecom SudParis	Président
Jean-Claude BELFIORE Professeur, Huawei Technologie France	Rapporteur
Raphael Le BIDAN Maître de Conférences, IMT-Atlantique	Rapporteur
Catherine LEPERS Professeure, Institut Polytechnique de Paris, Telecom SudParis	Examinateur
Philippe CIBLAT Professeur, Institut Polytechnique de Paris, Telecom Paris	Examinateur
Walid HACHEM Directeur de recherche CNRS, Université Gustave Eiffel	Examinateur
Erwan PINCEMIN Ingénieur de Recherche, Orange Labs	Examinateur
Yves Jaouën Professeur, Institut Polytechnique de Paris, Telecom Paris	Directeur de thèse
Mansoor Yousefi Maître de Conférences, Institut Polytechnique de Paris, Telecom Paris	Co-Directeur de thèse

Complexity Reduction over Bi-RNN-Based Kerr Nonlinearity Equalization in Dual-Polarization Fiber-Optic Communications via a CRNN-Based Approach

Abtin Shahkarami

Department of Communications and Electronics
Institut Polytechnique de Paris, Telecom Paris
Paris, France

Doctoral dissertation, to be presented for public discussion with the permission of the doctoral school of Institut Polytechnique de Paris, Telecom Paris, in Amphitheater 2, on the 10th of October, 2022

Paris 2022

Author's Address: Institut Polytechnique de Paris, Telecom Paris
Place Marguerite Perey
91120 Palaiseau, France
abtin.shahkarami@telecom-paris.fr

Director: Professor Yves Jaouën
Institut Polytechnique de Paris, Telecom Paris

Co-director: Professor Mansoor Yousefi
Institut Polytechnique de Paris, Telecom Paris

Reviewers: Professor Jean-Claude Belfiore
Huawei Technologies France

Professor Raphael Le Bidan
IMT-Atlantique, Brest Campus

Examiners: Professor Catherine Lepers
Institut Polytechnique de Paris, Telecom SudParis

Professor Philippe Ciblat
Institut Polytechnique de Paris, Telecom Paris

Professor Walid Hachem
University Gustave Eiffel

Dr. Erwan Pincemin
Orange Labs France

Acknowledgements

Special thank you to my supervisors, Prof. Mansoor Yousefi and Prof. Yves Jaouën, for your devotion, invaluable guidance, diligent editing, and precious support. I have significantly profited from your wealth of knowledge and the lessons you taught me. I am immensely grateful for your confidence and trust in me over the years. My deepest gratitude, as well, to my industry advisor, Dr. Vahid Aref, for your insightful guidance and kind-hearted support.

I would cordially thank the members of my thesis jury for their precious time and efforts in evaluating my doctoral manuscript and their valuable comments and suggestions.

I would like to thank the members of my comité de suivi de thèse for their follow-ups on my progress during the doctoral program and for their enlightening feedbacks on my work.

I would like to express my warm gratitude to all the members of the FONTE project for their insightful collaboration and the constructive feedbacks I received from them.

I would like to sincerely thank all my venerable professors and teachers during my academic life and express my deepest gratitude to them.

I would also like hereby to honor the memory of my middle school mathematics teacher, Mr. Davijani, who marked a turning point in my academic life and instilled in me the passion and drive to be a pioneer student.

Last, and above all, I wish to express my love and most profound gratefulness to my dear parents and lovely siblings for their endless love, care, and encouragement in every stage of my life.

Abtin Shahkarami
Paris, October 2022

Abstract

The impairments arising from Kerr nonlinearity in optical fiber is a major hindering effect in optical fiber communications, limiting the achievable information rate. Unlike linear effects, such as chromatic and polarization-mode dispersion (PMD), which can be compensated via linear equalization in a low-complexity regime, nonlinear equalizers are required to be implemented in digital signal processing (DSP) unit of the receivers in fiber-optic transmission systems to mitigate nonlinear channel impairments. This requirement substantially escalates the computational complexity of the DSP unit. With the conventional nonlinear equalization methods, such as digital backpropagation (DBP), this computational overhead becomes so significant that it may impede the practical implantation of the method in real-time systems.

With regard to this matter, considering the universality of neural networks, deep learning-based methods have recently attracted attention for nonlinearity mitigation in fiber-optic communications. In these methods, a neural network structure is adopted to learn the nonlinearity patterns within the received signal. These sorts of approaches have shown success in this task by attaining a comparable performance to conventional methods while having lower computational complexity.

Subsequent to the advent of this research area, as a matter of course, there is ever-increasing attention in the community over the concept of the computational efficiency in the adopted neural structures for equalization; in pursuit of investigating more efficient neural networks capable of attaining comparable or higher performance relative to other networks while incurring lower or equivalent complexity.

This topic founds the subject of this doctoral manuscript, which has been prepared within the framework of European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 766115.

In this manuscript, following having a concise overview of the fields of transmission systems, fiber-optic communication, and neural networks, we investigate a variety of state-of-the-art neural network-based equalization methods, in particular bidirectional recurrent neural network (bi-RNN) -based approaches, considering their superior performance, with their implementation on a natively-developed polarization-multiplexed

fiber-optic transmission system model.

In continuation of these studies, subsequently, through identifying and addressing the sources of inefficiency in the adopted networks, we propose a more computationally efficient neural structure formed as a hybrid convolutional recurrent neural network (CRNN) comprising a convolutional neural network (CNN) -based encoder and a unidirectional many-to-one vanilla RNN working in tandem, each best capturing specific types of channel impairments while compensating for the shortcomings of the other.

We show that for a 64 GBd dual-polarization 16-QAM optical transmission over 14×80 km standard single-mode fiber, the proposed hybrid CRNN-based method achieves superior or comparable performance to recently-proposed multilayer perceptron (MLP), CNN+MLP, bi-RNN, bidirectional gated recurrent unit network, bidirectional long short-term memory (bi-LSTM) network, and CNN+bi-LSTM -based equalizers in the literature, with considerably lower complexity measured by the number of floating-point operations (FLOPs). In particular, we demonstrate that the suggested model approaches the performance of the state-of-the-art bidirectional recurrent-based methods with $> 50\%$ lower computational complexity compared to them.

In addition, by taking into account the polarization-mode dispersion effects, including frequent rotation of the state-of-polarization and variation of differential group delay along the fiber, and also by considering the laser phase noise in supplement to the previous studies, we further investigate a more coherent way in dealing with these effects when adopting a neural network-based equalization solution. We demonstrate that a solitary neural network-based nonlinearity mitigation approach implemented at the end of the linear equalization chain could deliver higher efficiency in terms of performance and complexity relative to incorporating the polarization-mode dispersion compensation and carrier phase estimation task into the neural network's task.

Résumé

Les dégradations dues à la non-linéarité de Kerr dans les fibres optiques constituent une entrave majeure aux communications sur fibre optique, limitant le débit d'information. Contrairement aux effets linéaires, tels que la dispersion chromatique et la dispersion en mode de polarisation (PMD), qui peuvent être compensés par une égalisation linéaire dans un régime de faible complexité, les égaliseurs non linéaires doivent être mis en œuvre dans l'unité de traitement numérique du signal (DSP) des récepteurs dans les systèmes de transmission sur fibre optique pour atténuer les dégradations non linéaires des canaux. Cette exigence augmente considérablement la complexité de calcul de l'unité DSP. Avec les méthodes d'égalisation non linéaires classiques, telles que Digital backpropagation (DBP), cette surcharge de calcul devient si importante qu'elle peut entraver l'implantation pratique de la méthode dans les systèmes en temps réel.

À ce sujet, compte tenu de l'universalité des réseaux neuronaux, les méthodes basées sur l'apprentissage profond ont récemment attiré l'attention pour l'atténuation de la non-linéarité dans les communications sur fibre optique. Dans ces méthodes, une structure de réseau neuronal est adoptée pour apprendre les modèles de non-linéarité dans le signal reçu. Ces types d'approche ont montré leur succès dans cette tâche en atteignant une performance comparable aux méthodes conventionnelles tout en ayant une complexité de calcul plus faible.

Après l'avènement de ce domaine de recherche, la communauté s'intéresse de plus en plus au concept d'efficacité de calcul dans les structures neuronales adoptées pour l'égalisation, dans le but d'étudier des réseaux neuronaux plus efficaces capables d'atteindre des performances comparables ou supérieures à celles d'autres réseaux tout en présentant une complexité inférieure ou équivalente.

Ce sujet constitue le thème de ce manuscrit de doctorat, qui a été préparé dans le cadre du programme de recherche et d'innovation Horizon 2020 de l'Union européenne, sous la convention de subvention Marie Skłodowska-Curie n° 766115.

Dans ce manuscrit, après avoir eu un aperçu concis des domaines des systèmes de transmission, de la communication sur fibre optique et des réseaux neuronaux, nous étudions une variété de méthodes d'égalisation basées sur les réseaux neuronaux, en particulier les approches basées sur les réseaux neuronaux récurrents bidirectionnels (bi-RNN), compte tenu de leur performance supérieure, avec leur mise en œuvre sur

un modèle de système de transmission sur fibre optique multiplexé par polarisation développé nativement.

Dans le cadre de ces études, par la suite, en identifiant et en traitant les sources d'inefficacité des réseaux adoptés, nous proposons une structure neuronale plus efficace en termes de calcul, sous la forme d'un réseau neuronal récurrent convolutionnel (CRNN) hybride comprenant un encodeur basé sur un réseau neuronal convolutif (CNN) et un réseau de neurones récurrents vanille unidirectionnel travaillant en tandem, chacun capturant au mieux des types spécifiques de dégradations de canaux tout en compensant les défauts de l'autre.

Nous montrons que pour une transmission optique 16-QAM à double polarisation de 64 GBd sur une fibre monomode standard de 14×80 km, la méthode hybride CRNN proposée atteint des performances supérieures ou comparables à celles des égaliseurs basés sur un perceptron multicouche (MLP), CNN+MLP, bi-RNN, réseau d'unités récurrentes gated bidirectionnel, réseau de mémoire à long terme bidirectionnel (bi-LSTM) et CNN+bi-LSTM récemment proposés dans la littérature, avec une complexité considérablement inférieure mesurée par le nombre d'opérations en virgule flottante (FLOP). En particulier, nous démontrons que le modèle suggéré s'approche de la performance des méthodes bidirectionnelles récurrentes de pointe avec une complexité de calcul inférieure de $> 50\%$ par rapport à celles-ci.

En outre, en prenant en compte les effets de dispersion du mode de polarisation, y compris la rotation fréquente de l'état de polarisation et la variation du retard de groupe différentiel le long de la fibre, et en considérant également le bruit de phase du laser en complément des études précédentes, nous étudions la manière plus cohérente de traiter ces effets lors de l'adoption d'une solution d'égalisation basée sur un réseau neuronal. Nous démontrons qu'une approche d'atténuation de la non-linéarité basée sur un réseau neuronal solitaire, mise en œuvre à la fin de la chaîne d'égalisation linéaire, pourrait offrir une efficacité supérieure en termes de performance et de complexité par rapport à l'intégration de la compensation de la dispersion en mode de polarisation et de l'estimation de la phase de la porteuse dans la tâche du réseau neuronal.

Contents

1 Introduction **19**

2 Optical fiber transmission systems: from modulation to nonlinearity mitigation **24**

2.1 Optical fiber as a data transmission medium 25
2.1.1 Attenuation 26
2.1.2 Chromatic dispersion 27
2.1.3 Polarization mode dispersion 28
2.1.4 Nonlinearity 30
2.2 Transmission in fiber-optic communications 31
2.2.1 Modulation 31
2.2.1.1 Mathematical foundation 31
2.2.1.2 Fundamental modulation techniques for digital communication systems 33
2.2.1.3 Pulse shaping 35
2.2.1.4 Signal power 38
2.2.2 Polarization-division multiplexing 40
2.2.3 Laser phase noise 41
2.3 Channel model of dual-polarization optical fiber transmission systems . 42
2.3.1 Signal propagation in one span of single-mode optical fiber . . . 42
2.3.2 Optical amplification 45
2.3.2.1 Erbium-doped fiber amplifiers 46
2.3.2.2 Amplified spontaneous emission noise effect 47
2.4 Optical receiver and digital signal processing 48

2.4.1 Optical detection . . . 48
2.4.2 Digital signal processing in dual-polarization optical fiber communication . . . 51
2.4.2.1 Matched filtering . . . 51
2.4.2.2 Chromatic dispersion compensation . . . 52
2.4.2.3 MIMO-based equalization . . . 53
2.4.2.4 Carrier phase estimation . . . 55
2.4.2.5 Nonlinearity mitigation through digital backpropagation 56
2.5 Conclusion . . . 58

3 Learning by neural networks: principles and fundamental models 59
3.1 Multi-layer perceptrons and universal approximation theorem . . . 64
3.2 Convolutional neural networks . . . 66
3.3 Recurrent neural networks . . . 68
3.3.1 Vanilla RNNs . . . 68
3.3.2 Long-Short Term Memory Networks . . . 69
3.3.3 Gated Recurrent Unit Networks . . . 70
3.3.4 Different topologies . . . 71
3.4 Dimensionality reduction and latent space . . . 71
3.5 Conclusion . . . 76

4 Neural network-based equalization in optical fiber communication 77
4.1 Neural network-based equalizers in the recent literature . . . 78
4.2 Complexity reduction using latent space of CNNs . . . 84
4.3 Conclusion . . . 88

5 Implementation and numerical results 89
5.1 Implemented polarization-multiplexed fiber-optic transmission system model . . . 89
5.1.1 Transmitter model . . . 90
5.1.2 Channel model . . . 91
5.1.3 Receiver model . . . 91
5.2 Developed software . . . 93

5.3 Training methodology 93
5.3.1 RX model 1 95
5.3.2 RX model 2 98
5.4 Complexity measurement 98
5.5 Evaluation 100
5.5.1 RX model 1 100
5.5.2 RX model 2 102
5.5.3 RX 1 versus RX 2 105
5.6 Conclusion 106

Conclusions **108**

Future studies **110**

References **115**

List of Figures

1.1 Operating submarine optical fiber cables by May 2022 21

2.1 General anatomy of an optical fiber 26
2.2 Attenuation of optical fiber as a function of the wavelength 27
2.3 Spliced simple birefringent fiber segments at random rotational angles introducing PMD effects in cabled optical fibers 29
2.4 An illustration of four common modulation techniques: OOK, ASK, FSK, BPSK, to modulate a binary data stream over a sine carrier wave. 33
2.5 Sample constellations for 8-QAM and 16-QAM 35
2.6 BER performance of different QAM orders over an AWGN channel. . . 36
2.7 Receiver constellation for different QAM orders over AWGN channel . . 36
2.8 Polarization-multiplexed signal propagation 40
2.9 Phase noise effect on the constellation based on different laser linewidths 41
2.10 Schematic of an optical link 43
2.11 Spontaneous and Stimulated emission 45
2.12 Schematic structure of the EDFA amplifier 46
2.13 Erbium excitation diagram for stimulated emission 47
2.14 Non-coherent detection 49
2.15 Coherent detection 50
2.16 Conventional DSP chain in long-haul optical fiber communication system considered in this manuscript. 51
2.17 Effect of matched filtering in resulting constellation at receiver 52
2.18 Radii of constellations in QPSK and QAM signals 55
2.19 Underlying approach of DBP 56
2.20 BER performances of the receiver using CD and DBP. 57

3.1 Schematic of an MLP 65
3.2 Schematic of the process flow in an LSTM cell 69
3.3 Schematic of the process flow in an GRU cell 70
3.4 Common different topologies for recurrent networks 72
3.5 d-dimensional unit sphere inscribed in d-dimensional unit cube. 73
3.6 Ideal MLP autoencoder 75

4.1 Computational graph of LDBP 79
4.2 MLP-based equalizer by Catanese et al. 81
4.3 CNN+MLP-based equalizer by Chuang et al. 81
4.4 General structure of bi-RNN based equalizer 82
4.5 CNN+bi-LSTM-based equalizer by Freire et al. 82
4.6 Proposed CRNN model 86
4.7 Block diagram of the proposed CRNN model architecture 87

5.1 Block diagram of the considered fiber-optic transmission system 90
5.2 Input-output model of the neural networks in both RX models. 92
5.3 Optimizing the number of units in RNN-based approaches for RX 1 . . 95
5.4 Optimizing the number of units in RNN-based approaches for RX 2 . . 97
5.5 BER performance of the neural networks in RX 1 100
5.6 Q-factor performance of the neural networks in RX 1 101
5.7 BER performance of the neural networks in RX 2 103
5.8 Q-factor performance of the neural networks in RX 2 104
5.9 Comparison of the Q-factor performance of RX 1 and RX 2 106
5.10 Obtained constellation by RX 1 107
5.11 Obtained constellation by RX 2 107
5.12 Pruning and quantization via attention technique for future studies . . 111
5.13 Deep reinforcement learning -based joint transmitter and receiver optimization approach for future studies 114

List of Tables

3.1 Common activation functions adopted in neural networks 62
3.2 Common loss functions adopted in the machine learning tasks 63

5.1 Details of the hidden layers of the implemented neural networks for RX 1 94
5.2 Details of the hidden layers of the implemented neural networks for RX 2. 96
5.3 Neural networks' number of parameters and FLOPs, in RX 1. 102
5.4 Neural networks' number of parameters and FLOPs, in RX 2. 105

Acronyms

Acronym	Full form
AIR	Achievable Information Rate
ASE	Amplified Spontaneous Emission
ASK	Amplitude Shift Keying
BER	Bit Error Ratio
Bi-RNN	Bidirectional Recurrent Neural Network
BPSK	Binary Phase-Shift Keying
CD	Chromatic Dispersion
CMA	Constant Modulus Algorithm
CNLSE	Coupled Nonlinear Schrödinger Equation
CNN	Convolutional Neural Network
CPE	Carrier Phase Estimation
CRNN	Convolutional Recurrent Neural Network
DBP	Digital Backpropagation
DFT	Discrete Fourier transform
DGD	Differential Group Delay
DSP	Digital Signal Processing
EC	European Commission
EDFA	Erbium-Doped Fiber Amplifier
ELU	Exponential Linear Unit
FEC	Forward Error Correction
FIR	Finite Impulse Response
FLOPs	Floating-Point Operations
FNN	Feedforward Neural Network
FONTE	Fibre Optic nonlinear Technologies project [EC GA766115]
FPGA	Field-Programmable Gate Array
FSK	Frequency-Shift Keying
GRU	Gated Recurrent Unit
GVD	Group Velocity Dispersion
HPC	High-performance computing
IDFT	Inverse Discrete Fourier transform

Acronym	Full form
IDFT	Inverse Discrete Fourier transform
LDBP	Learned Digital Backpropagation
LSTM	Long-Short Term Memory
MAE	Mean Absolute Error
MIMO	Multiple-Input Multiple-Output
ML	Machine Learning
MLP	Multilayer Perceptron
MSCA	Marie Sklodowska Curie Action
MSE	Mean Square Error
NF	Noise Figure
NFT	Nonlinear Fourier Transform
NLS	Nonlinear Schrödinger
NLSE	Nonlinear Schrödinger Equation
OOK	On-Off Keying
PDM	Polarization-Division Multiplexing
PMD	Polarization-Mode Dispersion
PSK	Phase-Shift Keying
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase Shift Keying
RC	Raised-Cosine
RDE	Radius Directed Equalization
ReLU	Rectified Linear Unit
RL	Reinforcement Learning
RNN	Recurrent Neural Network
RRC	Root-Raised Cosine
RX	Receiver
SGD	Stochastic Gradient Descent
SMF	Single-Mode Fiber
SNR	Signal-to-Noise Ratio
SOP	State-of-Polarization
SpS	Samples per Symbol
SSFM	Split-Step Fourier Method
SSMF	Standard Single-Mode Fibre
StPS	Step per Span

Notations

Notation	Description
$\sigma_\ell(.)$	Neural network's activation function for layer ℓ
ω	Angular frequency
R_s	Baud rate
$\mathbf{b}_\ell$	Bias vector for layer ℓ
$\mathbf{c}^{(t)}$	Cell state of of the time-step t in LSTM and GRU
D	Chromatic dispersion parameter
q_x	Complex envelope of the signal in x-polarization propagating in the fiber as a function of time t and distance z
q_y	Complex envelope of the signal in y-polarization propagating in the fiber as a function of time t and distance z
$\mathbf{q}$	$[\mathbf{q}_x, \mathbf{q}_y]^T$
$\mathbf{F}$	Discrete Fourier transform matrix
M	Effective channel memory
A_{eff}	Effective cross-sectional area
n_{eff}	Effective refractive index
α	Fiber loss
L_{sp}	Fiber span length
$\mathbf{\Gamma}_f$	Forget gate weights of LSTM
v_g	Group velocity
β_2	Group velocity dispersion
$\mathbf{h}^{(t)}$	Hidden state of the time-step t in RNNs (vanilla LSTM GRU)
$\mathbf{x}^{(t)}$	Input to the time-step t of the RNN
$\mathbf{F}^\dagger$	Inverse discrete Fourier transform matrix
α_{lr}	Learning rate
L_{ker}	Length of the 1-D convolution kernel
n_l	Linear refractive index coefficient
$\hat{L}(W)$	Loss function (value) given the wights matrix W

Notation	Description
$W_\ell^{<t>}$	Neural network's weight matrix of layer ℓ at iteration t
n_{nl}	Nonlinear refractive index coefficient
γ	Nonlinearity coefficient
N_{ch}	Number of channels in the convolution layer
N_{stps}	Number of step per span
N_{taps}	Number of taps
$\mathfrak{L}$	Optical fiber length
$(\varphi_i)_{i=1}^n$	Orthonormal basis spanning a Hilbert space H^n
$\mathbf{\Gamma}_o$	Output gate weights of LSTM
v_p	Phase velocity
h	Planck constant
τ	PMD parameter
$\mathbf{\Gamma}_r$	Relevance gate weights of LSTM
$\mathcal{L}_2$	Set of all finite energy functions
$\text{MLP}(\sigma, d, N_\ell)$	Set of all MLPs with d-dimensional input, N_ℓ layers, and activation $\sigma(.)$
σ_{sig}	Sigmoid activation
c	Speed of light in vacuum
δ	Step size in SSFM
T_s	Symbol period
σ_{tanh}	Tanh activation
β_1^{Adam}	The exponential decay rate for the first moment estimates in Adam algorithm
β_2^{Adam}	The exponential decay rate for the second moment estimates in Adam algorithm
$s_x^{(i)}, \hat{s}_x^{(i)}$	Transmitted symbol at symbol time-step i modulated on x-polarization, and the corresponding recovered symbol at receiver
$s_y^{(i)}, \hat{s}_y^{(i)}$	Transmitted symbol at symbol time-step i modulated on y-polarization, and the corresponding recovered symbol at receiver
$\mathbf{\Gamma}_u$	Update gate weights of LSTM and GRU
λ	Vacuum wavelength
k_w	Wave number

List of publications

The following publications have resulted from the research carried out within the doctoral program:

I. Abtin Shahkarami, Mansoor Yousefi, and Yves Jaouën, "Complexity reduction over Bi-RNN-based nonlinearity mitigation in dual-pol fiber-optic communications via a CRNN-based approach (invited paper)," *Journal of Optical Fiber Technology*, DOI:10.1016/j.yofte.2022.103072.

II. Abtin Shahkarami, Mansoor Yousefi, and Yves Jaouën, "Efficient Deep Learning of Kerr Nonlinearity in Fiber-Optic Channels Using a Convolutional Recurrent Neural Network (invited contribution)," *Deep Learning Applications*, vol. 4, Springer Nature. [Accepted, in press]

III. Abtin Shahkarami, Mansoor Yousefi and Yves Jaouën, "Efficient Deep Learning of Nonlinear Fiber-Optic Communications Using a Convolutional Recurrent Neural Network," in *IEEE International Conference on Machine Learning and Applications (ICMLA)*, 2021, pp. 668-673, DOI: 10.1109/ICMLA52953.2021.00112. [Selected paper of the conference]

IV. Abtin Shahkarami, Mansoor Yousefi, and Yves Jaouën, "Attention-Based Neural Network Equalization in Fiber-Optic Communications," in *Asia Communications and Photonics Conference (ACP)*, Optical Society of America, 2021, p. M5H.3, DOI:10.1364/ACPC.2021.M5H.3.

V. Abtin Shahkarami, Hossein Bobarshad, and Nader Bagherzadeh, "A stream-sensitive distributed approach for configuring cascaded classifier topologies in real-time large-scale stream mining systems," *Springer Nature Applied Sciences*, DOI:10.1007/s42452-019-0565-6. [Based on master's thesis, published during doctoral program]

Chapter 1

Introduction

The ubiquitous usage of fiber-optic communications has ushered in a new era of data networking. The use of optical fibers in place of copper cables nowadays has made it possible for telecom connections to be established over longer distances and with lower loss in the transmission medium. Optical fibers have made it possible to transmit data at speeds that are several orders of magnitude greater than those possible with the electrical transmission.

These benefits have resulted in the extensive use of optical fiber communication systems for a variety of applications ranging from the backbone infrastructure of long-haul communication systems to Ethernet systems, broadband distribution systems, and general data networking. In particular, nowadays, more than 99% global Internet traffic pass through optical fibers [1]. Fig. 1.1 shows the map of submarine optical fiber cables connecting countries and continents by May 2022.

Consequently, with the exponential increase in the demand for data rates worldwide, there is escalating pressure imposed on fiber-optic networks. Based on the CISCO annual report, the Internet traffic has grown 3.7-fold from 2017 to 2022 (a compound annual growth rate of 30%), and IP traffic will reach 50 GB per capita in 2022, up from 16 GB per capita in 2017 [2]. In an estimate, there will be 5.3 billion total Internet users (66% of the global population) by 2023, up from 3.9 billion (51% of the global population) in 2018 [3], and the gigabyte equivalent of all movies ever made will cross global IP networks among these users every 1 minute [2].

Given this drastic growth in demand, the quality of fiber-optic systems needs to be

raised to meet the situation and tolerate the pressures. Concerning this matter, besides coherent detection combined with high-speed digital signal processing (DSP) circuits, the underlying algorithms to reverse the distortion effects of fiber-optic channels arising from their physical properties should be improved. These effects include attenuation, chromatic dispersion (CD), polarization-mode dispersion (PMD), and the Kerr effect (also called the quadratic electro-optic effect), which is caused by the change in the refractive index of the optical link in response to the applied electric field [4]–[6].

As opposed to CD and PMD, which are linear effects, the Kerr effect is proportional to the square of the transmitted signal amplitude and does not vary linearly with it. This nonlinear property limits the achievable information rate of the conventional transmission techniques. It also makes the equalization task at the receiver (RX) a complex nonlinear process requiring substantially higher computing resources than that for linear channels. Digital backpropagation (DBP) [7], [8], Volterra series transfer function [9], [10], maximum-likelihood sequence estimation (MLSE) [11], [12], and optical phase conjugation [13]–[15] are the main conventional nonlinear equalization solutions, in this regard, which suffer from high computational complexity hindering their real-time functioning in practical systems.

With the advent of the concept of the neural network in 1943 by Warren McCulloch and Walter Pitts [17], and following the progress of this field and its implementation in light of the advancement of computers, and the success of this field in many areas of research and industry [18]–[20], neural networks gradually found their ways into fiber-optic communications, and they were leveraged for equalization [21]–[29] to attain higher efficiency in terms of complexity-performance trade-off.

These networks, which are composed of concatenated linear and nonlinear operators, can be optimized via the backpropagation algorithm [30] and variants of stochastic Gradient Descent [31, Ch. 4] to capture the complex nonlinear input-output correlations and mathematical structures in the presented data, and thus they are used to learn the deterministic distortion patterns across the signal by observing multiple input-output samples of the fiber-optic channel.

Neural networks are hierarchically organized into groups of basic processing units or neurons, and their performance and efficiency are associated with this organization and the scheme that information propagates from the input layer to the output. According to this fact, different levels of performance and efficiency, in terms of performance-

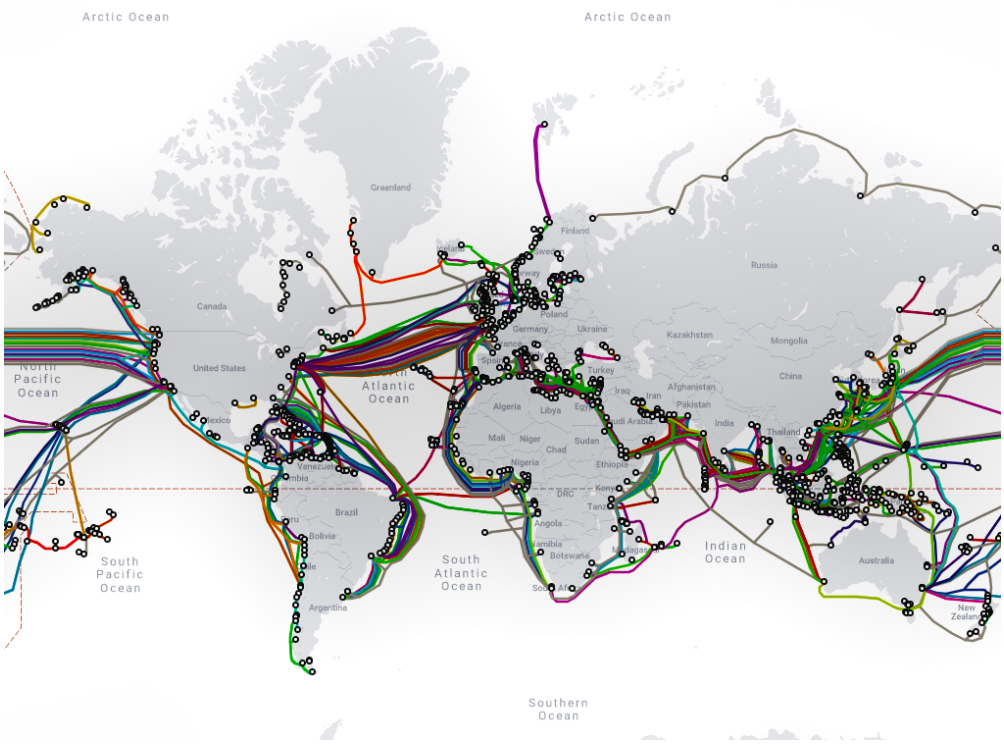


Figure 1.1: Operating submarine optical fiber cables by May 2022 as the infrastructure of the Internet, connecting continents and countries [16]

complexity trade-off, have been obtained via different adopted architectures for neural network-based equalization; and there is ongoing research to devise more efficient neural network models capable of attaining comparable or higher performance than other networks while incurring lower or comparable complexity, respectively.

This topic has been the subject of the doctoral program resulting in this manuscript. This doctoral program started in May 2019 within the FONTE-EID project, in the framework of the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 766115. During this period, we investigated a variety of state-of-the-art neural network-based equalization methods, with a focus on bidirectional recurrent neural network (bi-RNN) -based approaches considering their superior performance, with their implementation on a natively-developed long-haul fiber-optic transmission system model. In these studies, we analyzed the sources of computational inefficiency in the processing structures of the adopted networks; and subsequently, by addressing them, we proposed a more efficient architecture formed as a hybrid convolutional recurrent neural network (CRNN) comprising a

CNN-based encoder and a unidirectional many-to-one vanilla RNN working in tandem, each best capturing specific types of channel impairments while compensating for the shortcomings of the other.

We have demonstrated that for the nonlinearity and joint nonlinearity-PMD mitigation at RX in a 64 GBd dual-polarization 16-QAM optical transmission over 14×80 km standard single-mode fiber (SSMF), which is a typical system architecture for fiber-optic transmission systems, the proposed CRNN model achieves the comparable performance to the state-of-the-art bi-RNN, bi-GRU, bi-LSTM, and CNN+bi-LSTM -based equalization approaches [32]–[35] with greater than 50% fewer number of FLOPs, and it achieves superior performance over multi-layer perceptron (MLP) and convolutional neural network (CNN) -based approaches with lower complexity.

As the dissemination results of the carried out research during the doctoral program, we published two conference papers, one journal paper (*invited paper*), one peer-reviewed book chapter (*invited paper*), nine deliverables to European Union regarding FONTE-EID project, and the software project entitled "Fiber-Optic Transmission System Modeling" under the GNU General Public License v3.0. It is noteworthy to mention that among these publications, the conference paper "Efficient Deep Learning of Nonlinear Fiber-Optic Communications Using a Convolutional Recurrent Neural Network" was placed among the ten selected papers of the 2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA). Subsequently, we were invited to submit the extended version of the paper to be published as a book chapter in "Deep Learning Applications, vol. 4," published by Springer Nature.

The concepts in these publications in the elaborated form and their backgrounds aim to form the core of this doctoral manuscript. With respect to this, in this manuscript, following providing a succinct background and overview of fiber-optic transmission systems and neural networks, we elaborate on the methods and outcomes of the conducted research and investigations to provide a clear understanding of the research problem, achieved results, and the developed software libraries and resources within the doctoral program.

This doctoral manuscript is organized into five chapters to better convey the objective discussed. In the remainder of this section, we provide a brief introduction to each of these chapters.

Chapter 2 provides a concise overview of the principles of the dual-polarization fiber-

optic communication systems covering a variety of concepts about the fundamental technology in these systems, transmission schemes, fiber-optic channel model, different types of channel impairments and distortions, and required digital signal processing (DSP) at the receiver to compensate for them. The purpose of this chapter is two-fold. First, to provide the user with an understanding of the considered optical fiber communication system in this study and the corresponding developed software. Second, and more importantly, to familiarize the reader with the concept of channel impairments, especially nonlinearity, and the substantial processing overhead required to compensate for it.

Chapter 3 briefly discusses the underlying theory of the neural networks, different categories of the neural networks, including feed-forward neural networks (FNNs) and recurrent neural networks (RNNs), and the concept of feature extraction, dimensionality reduction, and latent space. This section aims to familiarize the reader with the concept of the architecture of the neural network and its decisive role in the performance and complexity of the neural model.

Chapter 4 splits in to two parts. The beginning part focuses on the state-of-the-art neural network-based equalization methods, including MLP, CNN, bi-RNN, bi-GRU, bi-LSTM, and CNN+bi-LSTM methods, and it reviews and debates them. Through critical analysis, this section tries to shed light on the sources of complexity and inefficiency in these networks, which forms the motivation for the devised CRNN model for addressing them. The second part of this chapter presents the proposed CRNN model in detail, elaborates on its process flow, and substantiates its underlying logic.

Chapter 5 dedicates to evaluating the devised CRNN model in comparison to the discussed state-of-the-art methods discussed in Chapter 4. This chapter, in the beginning, clearly explains the modeled fiber-optic communication systems, the methods and algorithms for the modeling, the considered system setup, channel parameters, neural network setups and optimization, and the equalization target. Following clarifying the methods and algorithms, the chapter discusses the numerical results in the considered configurations and comprehensively debates the complexity versus the gain of the methods.

The conclusions section finally briefly reviews and concludes the manuscript. The future studies section, following the conclusions, moreover, discusses the potential areas for further research and investigations.

Chapter 2

Optical fiber transmission systems: from modulation to nonlinearity mitigation

As telecommunications evolved, the demand for quicker and more efficient data transmission grew. Faced with this demand, with the advent of silica-based optical fibers, the deployment of optical fiber communications as a principal infrastructure for telecommunication systems, especially for long-haul communications, was primarily taken into account.

Optical fiber communications, or alternatively fiber-optic communications, refers to the transmission of information from one location to another using light pulses transmitted over an optical fiber. Any optical fiber communication system consists of four major components: transmitter, optical fiber cable, optical amplifier, and receiver.

Briefly saying, the transmitter modulates data on an eclectic field, converts it to the optical domain, and transmits it over the optical fiber. The optical fiber, which consists of a core surrounded by a lower refractive index layer called cladding, directs the optical pulse towards the receiver. The receiver converts the optical pulses into the equivalent electrical field and applies digital signal processing algorithms to it to compensate for the channel's deterministic effects; a process termed equalization. Also, as signal attenuation happens along the optical fiber link owing to fiber loss, amplifiers are installed at various points along the optical fiber connection to compensate for the

attenuation.

Depending on the application, the system components may change. Systems used for low-capacity lines, particularly for local area networks, leverage schemes and components slightly distinct from those utilized by network providers offering ultra-high data rates across large distances. Despite this, the fundamental concepts are the same regardless of the system.

In this chapter, we provide a summary of the principal concepts in the mentioned fiber-optic communication systems components, from the physical properties of the fiber to the required digital signal processing at receiver DSP.

2.1 Optical fiber as a data transmission medium

An optical fiber is a flexible, transparent medium made by drawing glass or silica, which is intended to direct the transmitted light pulses toward the other end. The pulse propagation in optical fibers operates based on a physical phenomenon called total internal reflection, which means the full reflection of the light beam within a medium from the surrounding boundaries back into the medium. This phenomenon is rooted in the difference in the refractive index of the two adjacent materials, and it happens when the refractive index of the surrounding material is lower than the reflective index of the medium.

Optical fibers have been established based on this phenomenon. These mediums consist of a core, around which there is another layer called cladding, which is made of similar material, i.e., silica, but with a slightly lower refractive index. This results in as the light pulse propagates down the core and hits the boundaries, it undergoes total internal reflection and is therefore confined inside the optical fiber's core. Outside of the cladding, there exists a plastic layer called coating, with the purpose of providing protection to the core and cladding. Fig. 2.1 shows this anatomy of the optical fiber.

Depending on the properties of the fiber's core, optical fibers are categorized into single-mode and multi-mode fibers. Multi-mode fibers have a core diameter of $\sim 50\mu m$, enabling the propagation of multiple light modes. However, in consequence, the propagation in these fibers incurs modal dispersion and experiences higher reflections, re-

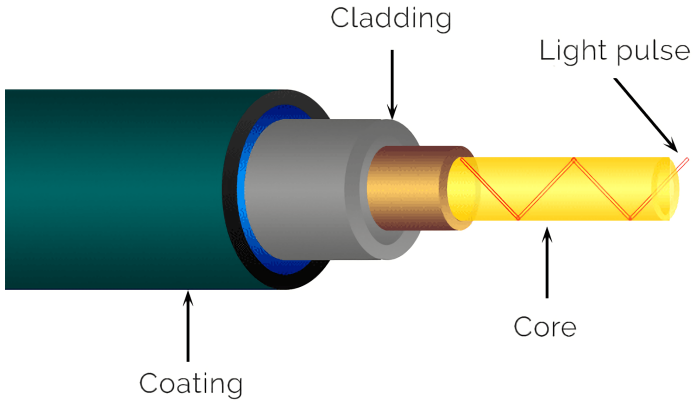


Figure 2.1: The general anatomy of an optical fiber, relied on the total internal reflection phenomenon [36].

sulting in lower bandwidth. [1] On the contrary, the core's diameter in single-mode fibers (SMFs) is reduced ($8 - 10.5\mu m$) such that it only allows for propagation in a single mode of light, resulting in higher bandwidths.

Propagation of signals in both types, however, is subject to [2] attenuation, chromatic dispersion (CD), polarization-mode dispersion (PMD), and nonlinear impairments [4]–[6] imposing distortions to the signal.

2.1.1 Attenuation

As the light pulse travels through the fiber, its intensity is attenuated. Attenuation in optical fibers is caused by two factors: absorption and scattering.

Absorption is induced by light absorption and conversion to heat by molecules of the fiber's glass. The principal absorbers are residual OH-ions and dopants that are employed to change the refractive index of the glass.

The main cause of attenuation, however, is scattering. Scattering happens when a wave and a particle interact in such a manner that the energy from the directed propagating

[1] These fibers are only efficient for short-distance transmissions ($<$ 2km) and have an easier and less costly manufacturing process.

[2] Not limited to

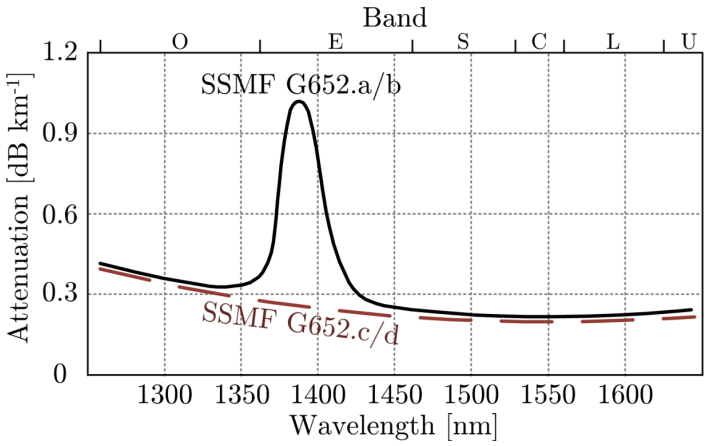


Figure 2.2: Attenuation rate of optical fiber as a function of the wavelength [37].

wave is removed, and it is transferred to other directions. In this case, the light is not absorbed but rather redirected.

Fig. 2.2 shows the loss rate in a standard SMF (SSMF) as a function of the wavelength. As this figure demonstrates, the minimum level of loss occurs at near the wavelength 1550nm with the rate of 0.2 dB/km. Note that the absorption peak at 1380nm due to OH-ions concentration is currently suppressed for the modern SSMFs (i.e., G.652 c/d fibers).

Due to the attenuation, the optical link is usually split into a number of so-called spans. At the end of each span, an amplification process via an amplifier compensates for the fiber loss. Section 2.3.2 describes this process.

2.1.2 Chromatic dispersion

Although laser sources are spectrally narrow in fiber-optic transmission systems, they are not monochromatic. This indicates that the input light pulse has several wavelength components. Moreover, modulation of the optical carrier intrinsically induces spectral broadening.

Chromatic dispersion refers to the phenomenon in which different spectral components of a light pulse have different velocities (traveling at different speeds) in the medium, depending on their frequency. This causes the pulse to spread, leading to intermixing

the slower wavelengths of one pulse with the faster wavelengths of the next pulse, resulting in inter-symbol interference.

Two factors contribute to chromatic dispersion: material dispersion, which is related to the dependence of the the refractive index of the fiber material to the wavelength, and waveguide dispersion, which is associated with the waveguide properties.

As mentioned, chromatic dispersion leads a short light pulse to be broadened through the propagation. This phenomenon is referred to as group velocity dispersion, and it is quantified as the derivative of the reciprocal of the group velocity with respect to the radian frequency

$$\beta_2 = \frac{\partial}{\partial\omega}\frac{1}{v_g} = \frac{\partial}{\partial\omega}\frac{\partial k_w}{\partial\omega} = \frac{\partial^2 k_w}{\partial\omega^2}, \tag{2.1.1}$$

where v_g is the group velocity (derivative of the phase velocity $v_p = \omega/k_w$), ω is the angular frequency, and k_w is the wavenumber. β_2 is called group velocity dispersion parameter.

In fiber-optic communications, however, there is a tendency to define dispersion as a derivative with respect to wavelength (rather than angular frequency). This is realized by

$$D = \frac{\partial}{\partial\lambda}\frac{1}{v_g} = -\frac{2\pi c}{\lambda^2}\frac{\partial^2 k_w}{\partial\omega^2} = -\frac{2\pi c}{\lambda^2}\beta_2, \tag{2.1.2}$$

where λ is the vacuum wavelength. D is also termed as chromatic dispersion parameter.

2.1.3 Polarization mode dispersion

In dual-polarization systems, the electrical field of the transmitted polarization-multiplexed signal can be decomposed into two orthogonal polarization [3], each carrying a different signal. In an ideal optical fiber, the two orthogonal polarizations travel at the same speed; however, in a realistic fiber, the two polarizations propagate at fractionally different velocities, owing to random imperfections and non-circularity of the

[3] polarization is typically defined in the context of the electric field.

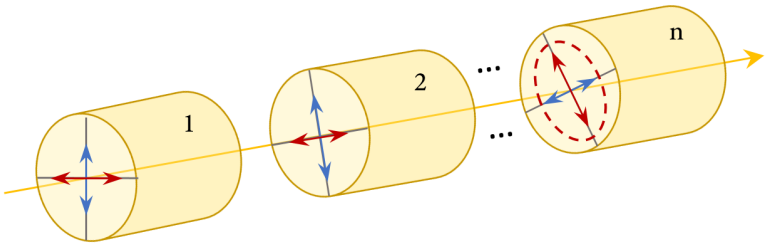


Figure 2.3: The realistic model of a cabled optical fiber as a chain of simple birefringent fiber segments concatenated at random rotational angles, resulting in PMD effects.

core, accumulating phase shift, and corresponding differential delay in proportion to distance [4] (termed as birefringence). This phenomenon is called polarization mode dispersion, abbreviated as PMD.

According to this phenomenon, fiber birefringence separates the input pulse into linear slow and fast polarization modes across short distances. This disparity in propagation time is referred to as the differential group delay (DGD). DGD has Maxwellian distribution, and its mean value grows as the square root of fiber length (The average DGD divided by the square root of fiber length is called the PMD coefficient).

Besides DGD, as a matter of fact, only quite short lengths of optical fiber exhibit uniform birefringences. A cabled optical fiber, however, is more accurately represented, in practice, as a chain of basic birefringent fiber segments connected at random rotational orientations (see Fig. 2.3). Each segment changes its input polarization state into another output state by modifying the relative phase of the local fast and slow waves. The output state is then forwarded to the subsequent segment, where the wave splits into local fast and slow waves, and the process starts over.

With regard to the discussed distortion effects in propagation through optical fibers, proper transmission, reception, and equalization techniques should be adopted for the transmitter and receiver in the fiber-optic transmission systems.

[4] A waveguide generating such a phase shift in termed as birefringent. There are two roots for fiber birefringence: non-circularity of the core, and some material imperfections in the fiber resulting in polarization-dependent refractive index

2.1.4 Nonlinearity

Nonlinear impairments are another sort of distortions experienced by a signal propagating through a fiber-optic channel, particularly in long-distance transmissions. For high bit-rate systems, Kerr effect is the more significant type of nonlinear impairments. Because of its nonlinear nature, the Kerr effect is one of the key limiting factors of achievable information rate (AIR) in optical fiber communications.

Kerr effect describes the change in the refractive index of a material in response to an applied electric field, which is proportional to the square of the electric field instead of having a linear relationship with it. This phenomenon, which is rooted in the anharmonic motion of bound electrons under the influence of the applied electromagnetic field [38, Ch. 1.3], prompts the notion of the effective refractive index

$$n_{eff}(z,t) = n_l + n_{nl}\frac{|E(z,t)^2|}{A_{eff}} \tag{2.1.3}$$

where n_l is the linear refractive index, and n_{nl} is the nonlinear refractive index coefficient. A_{eff} is the effective cross-sectional area defined as

$$A_{eff} = \frac{\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|E(x,y)|^2\,dxdy\right)^2}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|E(x,y)|^4\,dxdy} \tag{2.1.4}$$

where $E(x,y)$ is the optical mode field distribution [39]. The nonlinearity coefficient subsequently is defined as

$$\gamma = \frac{2\pi}{\lambda}\frac{n_{nl}}{A_{eff}} \tag{2.1.5}$$

A detailed and broad discussion on the concept of Kerr nonlinearity is presented in [40, pp. 275-297].

Kerr nonlinearity, however, is better understandable in the context of the nonlinear Schrödinder equation describing the propagation of the signal in the optical fiber. In Section 2.3, we elaborate on this topic.

2.2 Transmission in fiber-optic communications

In optical fiber communication systems, the aim of the transmitter is to load the information on light beams and transmits them to the optical fiber in such a way that the received signal can be efficiently recovered at the receiver, given the channel distortions. In dual-polarization optical fiber communication systems, this task involves (not limited to) modulation and polarization-division multiplexing (PDM). In this section, we briefly describe these concepts and discuss the adopted methods in this work for their implementation.

2.2.1 Modulation

Modulation is the process of converting information (digital bit stream or analog signal) to a form that can be physically transmitted over the communication medium. In this process, the information is called the baseband or modulating signal; and the periodic waveform that is used to physically transmit information by varying one or more properties of it based on the baseband is called the carrier signal.

This section briefly discusses the mathematical foundation of this concept and the processes comprising the modulation, including constellation mapping, pulse shaping, and power adjustment.

2.2.1.1 Mathematical foundation

The idea of modulation has been built upon the corollary that every finite-energy function f in a Hilbert space H^n can be decomposed into an orthonormal basis $\{\varphi_1, \varphi_2, \ldots, \varphi_n\}$ of the space. In this subsection, we mathematically discuss this concept and formulate the modulation process.

Definition 2.2.1 (Inner product)**.** An inner product on a vector space E over a field K is a function $< ., . >: E \times E \to K$ satisfying the following properties [1]:

(a) $\forall\, x, y \in E,\ \ < x, y >= \overline{< y, x >}$.

(b) $\forall\, x_1, x_2, y \in E,\ \forall\, \lambda_1, \lambda_2 \in K,\ \ < \lambda_1 x_1 + \lambda_2 x_2, y >= \lambda_1 < x_1, y > + \lambda_2 < x_2, y >$.

(c) $\forall\, x \in E,\ \ < x, x > \geq 0\ \&\ < x, x >= 0$ iff. $x = 0$.

Definition 2.2.2 (Inner product)**.** An inner product space is a vector space together with an inner product.

Definition 2.2.3. A Hilbert space H is a real or complex inner product space that is also a complete metric space with respect to the distance function induced by the inner product.

Theorem 2.2.1. *The set of all finite energy functions ($\mathcal{L}_2$ space) forms a Hilbert space over the field $\mathbb{C}$ with the following inner product:*

$$< f(t), g(t) >= \int_{-\infty}^{\infty} f(t)\overline{g(t)}dt. \tag{2.2.1}$$

Theorem 2.2.2. *Every Hilbert space has an orthonormal basis.*

According to Theoream 2.2.1 and Theoream 2.2.2 every finite-energy function f can be decomposed into an orthonormal basis $\{\varphi_1, \varphi_2, \dots\}$ of the space as the following

$$f(t) = \sum_k s_k \varphi_k(t) \tag{2.2.2}$$

where $s_k \in \mathbb{C}$.

This notion forms the underlying logic of the modulation process by taking s_ks as the information to be sent, and $\varphi_k(t)$s as the selected orthonormal basis of the space, referred to as pulse shapes. In other words, mathematically defining, modulation is the process of mapping symbols s_k (information) to the function $f(t)$ (modulated signal) by expanding the modulating signal (baseband) in the orthonormal basis of the space (pulse shapes), according to (2.2.2). By this approach, the received waveform at the receiver can be decoded by the projection of the received signal into the orthonormal basis to obtain the transmitted symbols, a process called demodulation.

To demonstrate this, let $X(t) = \sum_{k=1}^{n} s_k \varphi_k$ be the transmitted signal. In an identity channel where the received signal Y equals transmitted signal X, the symbol s_i can be simply demodulated as follows:

$$\hat{s}_i =< Y, \varphi_i >=< X, \varphi_i >=< \sum_{k=1}^{n} s_k \varphi_k, \varphi_i >= \sum_{k=1}^{n} s_k < \varphi_k, \varphi_i >= s_i. \tag{2.2.3}$$

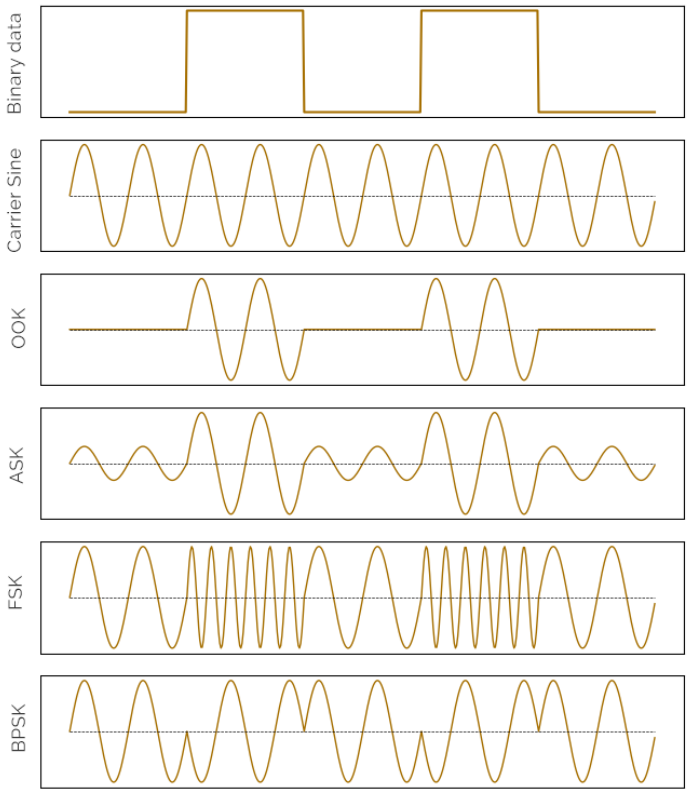


Figure 2.4: An illustration of four common modulation techniques: OOK, ASK, FSK, BPSK, to modulate a binary data stream over a sine carrier wave.

Based on the physical interpretation and assignments for s_ks and φ_ks in (2.2.2) different modulation techniques are developed. In the following section, we discuss the fundamental modulation techniques for digital communication systems.

2.2.1.2 Fundamental modulation techniques for digital communication systems

Fundamentally, there exist four main modulation techniques for digital communication systems, which are based on modifying the amplitude, frequency, phase, and combination of amplitude and phase to reach higher spectral efficiency, defined as the information rate that can be transmitted over a given bandwidth [41, Ch. 7].

On-off keying (OOK) and amplitude shift keying (ASK) are the two main types of amplitude modulation formats. In OOK, the binary signal (bit stream) toggles the

carrier amplitude on and off, resulting in the creation of OOK. On the other hand, ASK is produced by shifting the carrier amplitude between two different amplitude levels. Similarly, frequency-shift keying (FSK) is produced by shifting the carrier between two different frequencies named the mark and space frequencies. Fig. 2.4 presents an illustration of the OOK, ASK, and FSK modulation techniques.

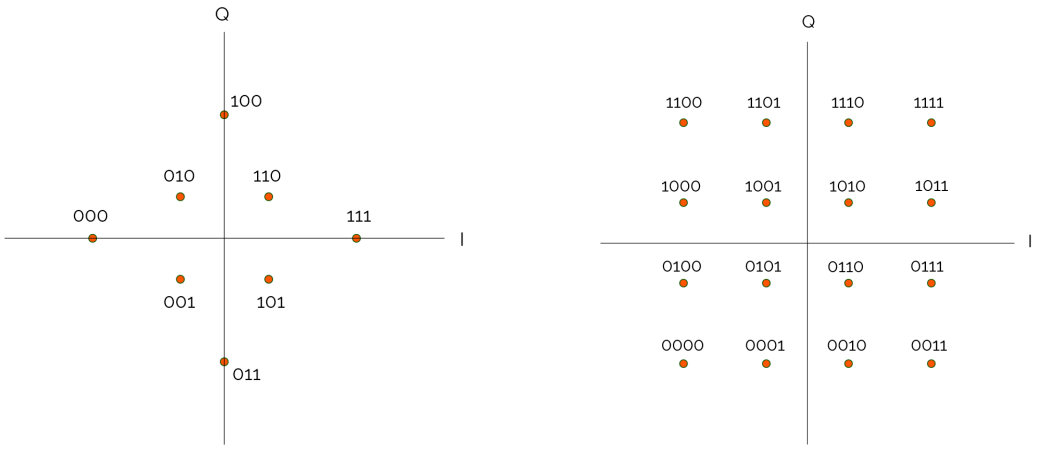
Binary phase-shift keying (BPSK) and quadrature phase-shift keying (QPSK) are two main techniques in the modulation based on the modification of the phase. In BPSK, for each change in the binary state, the carrier periodic waveform is shifted 180°. An example of BPSK modulation is also demonstrated in Fig. 2.4.

In QPSK, two periodic carrier waveforms are produced 90° apart. Then each phase is modulated by the binary data, resulting in four distinct sine signals that are shifted in phase by 45°. The final signal is produced by adding the two phases together. Carriers with distinct phases are generated for each unique pair of bits.

Modulation based on the modification of both amplitude and phase levels is called quadrature amplitude modulation (QAM). This modulation technique substantially improves spectral efficiency and allows for transmitting more information per symbol. In view of this, the QAM technique was adopted for the considered fiber-optic transmission system during the doctoral program.

QAM makes use of both amplitude and phase components to offer a form of modulation with higher spectral efficiency. In QAM, two carriers with 90° phase shift are modulated and merged to form a single carrier signal. The term quadrature in QAM refers to this 90° phase difference. In this technique, one of the two signals is called In-phase, and the other is called Quadrature. Both In-phase and Quadrature carriers can have amplitude variations.

According to this fact, as opposed to basic signals toggling between two levels, several points with different phase and amplitude levels (known as symbols) can be used in QAM. These points constitute a so-called constellation diagram (see. Fig. 2.5) in which each point (symbol) represents a specific phase and amplitude, and it is used for denoting specific pairs of bits. The number of these points in a constellation diagram typically is of power of 2, i.e., 16, 64, ..., which are referred to as 16-QAM, 64-QAM, ..., formats. By using more points in the constellation, the information rate that can be transmitted per symbol is increased, but by this action, the points get closer to each other, and they become prone to noise and other distortions, resulting in a higher



(a) 8-QAM (b) 16-QAM

Figure 2.5: Sample 8-QAM and 16-QAM constellations. QAM uses a mix of amplitudes and phases to achieve higher spectral efficiency. Each state denotes a symbol representing a sequence of bits. The number of bits carried per symbol in 8-QAM is 3 bits, and in 16-QAM is 4 bits.

bit-error-rate (BER). Fig. 2.6 demonstrates the BER performance of a communication system over an additive white Gaussian noise (AWGN) channel $Y = X(t) + N(t)$, where $N(t)$ is circularly distributed Gaussian noise, in transmission using different QAM formats. Fig. 2.7 also shows the corresponding obtained constellations at the receiver, which illustratively shows the signal degradation due to the random noise coming from the AWGN, by the rise of QAM order.

It should be mentioned that in the context of Section 2.2.1.1, the constellation points which account for s_ks in (2.2.2) are mathematically modeled by being drawn from $\mathbb{C}$. The orthonormal basis φ_ks are also chosen to be the set $\{p(t - i/R_s)\}_{i=-\infty}^{\infty}$ where $p(t)$ is the pulse shape and R_s is the baud rate. In the next part, we elaborate on the pulse shaping, in particular root-raised cosine pulse shapes, which were adopted for the considered communication system in our doctoral studies.

2.2.1.3 Pulse shaping

Mathematically expressing, pulse shaping is the process of adopting the pulse shape $p(t)$ with which $\{p(t - i/R_s)\}_{i=-\infty}^{\infty}$ forms an orthonormal basis for the space, so that the baseband can be modulated to the waveform $q(t, 0) = \sum_{i=-\infty}^{\infty} s_i p(t - i/R_s)$.

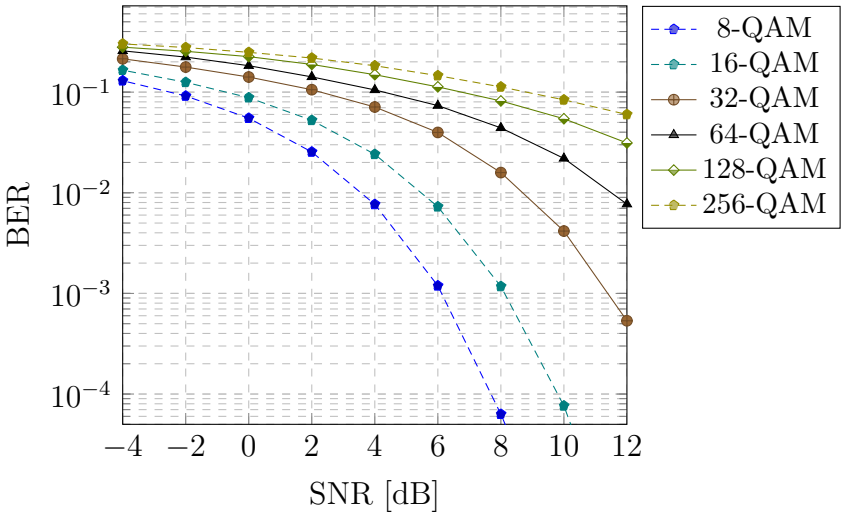
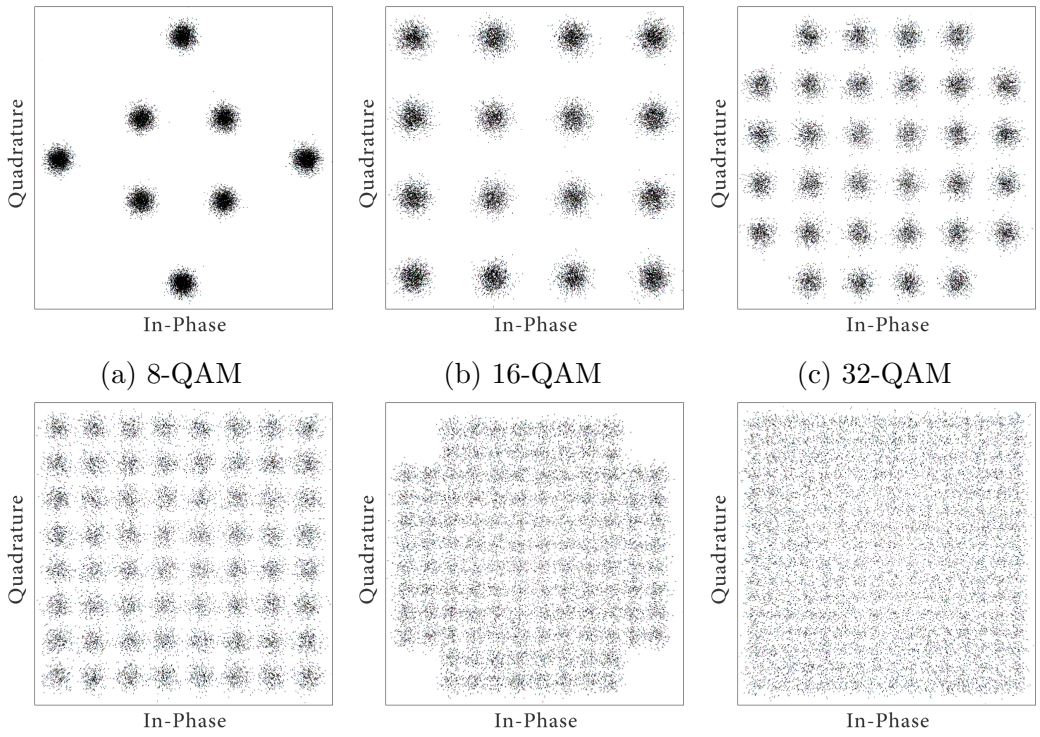


Figure 2.6: BER performance of a transmission system as a function of signal-to-noise ratio (SNR) over an AWGN channel using different QAM orders.



(a) 8-QAM (b) 16-QAM (c) 32-QAM

(d) 64-QAM (e) 128-QAM (f) 256-QAM

Figure 2.7: Obtained constellation at receiver in the considered transmission system over an AWGN channel at SNR 12 dB using different QAM formats.

In the context of physics, the term "pulse shaping" refers to modifying the waveform of transmitted pulses to make the transmitted signal more optimally suited to the communication channel. The employment of a pulse shaping filter is necessary for communications systems because of two essential criteria that must be satisfied by a communications channel: 1) generating band-limited signals, 2) eliminating interference from adjacent symbols, known as inter-symbol interference. An optimal pulse shaping filter that is applied to each symbol can fulfill both of these needs simultaneously.

Theoretically, an ideal pulse shaping filter is the sinc pulses (also called Boxcar filter). In the frequency domain, the sinc filter produces a rectangular shape, which perfectly limits the effective bandwidth. However, technically it is not feasible to implement the sinc filter accurately, considering that it is a non-causal filter (its output depends on future input as well) with relatively slowly decaying tails.

The common practical pulse shapes that are used in communication systems are the Gaussian filter, Raised-cosine (RC) filter, and Root raised-cosine (RCC) filter.

Gaussian pulse: A Gaussian filter has an impulse response in the form of a Gaussian function. The impulse response of the Gaussian filter is given by

$$h_{gs}(t) = \frac{\sqrt{\pi}}{\alpha_{gs}} \exp\left[-(\frac{\pi}{\alpha_{gs}}t)^2\right], \tag{2.2.4}$$

where α_{gs} is associated to 3-dB bandwidth-symbol time product of the Gaussian filter as follows

$$\alpha_{gs} = \frac{1}{BT_s}\sqrt{\frac{\log 2}{2}} \tag{2.2.5}$$

where B is bandwidth, and $T_s = 1/R_s$.

According to (2.2.4), by the rise of α_{gs}, the spectrum occupancy of the Gaussian filter is reduced, and the impulse response extends over neighboring symbols, which results in an increase in the inter-symbol interference.

Raised-cosine pulse: RC filter is a solution to address the problem of the sinc pulse, in order to eliminate infinitely long tails in either the time or frequency domain

(associated with rectangular pulse shape). The impulse response of an RC filter is as follows

$$h_{rc}(t) = \begin{cases} \dfrac{\pi}{4T_s}\,\mathrm{sinc}\left(\dfrac{1}{2\beta_{ro}}\right), & t = \pm\dfrac{T_s}{2\beta_{ro}} \\ \dfrac{1}{T_s}\,\mathrm{sinc}\left(\dfrac{t}{T_s}\right)\dfrac{\cos\left(\dfrac{\pi\beta_{ro}t}{T_s}\right)}{1-\left(\dfrac{2\beta_{ro}t}{T_s}\right)^2}, & \text{otherwise,} \end{cases} \tag{2.2.6}$$

where $0 \leq \beta_{ro} \leq 1$ is the roll-off factor.

Root raised-cosine pulse: RRC filter has a similar underlying motivation to RC filter, with a frequency response equal to the square root of the frequency response of RC, i.e $|H_{rrc}(f)| = \sqrt{|H_{rc}(f)|}$. The impulse response of an RRC filter is as follows

$$h_{rrc}(t) = \begin{cases} \dfrac{1}{T_s}\left(1+\beta_{ro}(\dfrac{4}{\pi}-1)\right), & t = 0 \\ \dfrac{\beta_{ro}}{T_s\sqrt{2}}\left[\left(1+\dfrac{2}{\pi}\right)\sin\left(\dfrac{\pi}{4\beta_{ro}}\right)+\left(1-\dfrac{2}{\pi}\right)\cos\left(\dfrac{\pi}{4\beta_{ro}}\right)\right], & t = \pm\dfrac{T_s}{4\beta_{ro}} \\ \dfrac{1}{T_s}\dfrac{\sin\left[\pi\dfrac{t}{T_s}(1-\beta_{ro})\right]+4\beta_{ro}\dfrac{t}{T_s}\cos\left[\pi\dfrac{t}{T_s}(1+\beta_{ro})\right]}{\pi\dfrac{t}{T_s}\left[1-\left(4\beta_{ro}\dfrac{t}{T_s}\right)^2\right]}, & \text{otherwise.} \end{cases} \tag{2.2.7}$$

The advantage of the RRC filter is that it is its own matched filter, thanks to being symmetric and real-valued.

2.2.1.4 Signal power

The area under the squared magnitude of a continuous-time signal $x(t)$ is defined as the energy E_s of the signal. That is to say

$$E_s = \int_{-\frac{T}{2}}^{\frac{T}{2}} |x(t)|^2 dt, \tag{2.2.8}$$

where T is the signal duration. The power of the signal $x(t)$ is defined as the energy divided by time, i.e., mathematically

$$P = \lim_{T\to\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} |x(t)|^2 dt. \tag{2.2.9}$$

Let $\mathcal{L}_B^2$ be the space of finite-energy functions that are band-limited to B Hz with the basis $(\varphi_k(t))_{k=1}^N$. Subsequently, if $x(t) \in \mathcal{L}_B^2$, $x(t) = \sum_k^N s_k \varphi_k(t)$ for some $s_k \in \mathbb{C}$. Using the basis $\varphi_k(t) = \text{sinc}(Bt - k/B)$ and the orthogonality property, the integral in the definition of power can be computed analytically so that

$$P = \frac{1}{N} \sum_{k=1}^{N} |s_k|^2, \tag{2.2.10}$$

given that $T = N/B$ and $\int \text{sinc}^2(Bx)dx = 1/B$.

When dealing with random processes $X(t)$, P becomes a random variable, and subsequently the term power connotates the average power. In the particular case that $x(t)$ is expanded in a basis as discussed, and s_ks are a sequence of independent and identically distributed (iid) random variables with variance σ^2, we have

$$P_a = \frac{1}{N} \sum_{k=1}^{N} \mathbb{E}[|s_k|^2] = \sigma^2. \tag{2.2.11}$$

In a broader definition, however, the average power for a general random process is defined as

$$\begin{aligned} P_a &= \lim_{T\to\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \mathbb{E}[|x(t)|^2] dt \\ &= \lim_{T\to\infty} \frac{1}{T} \int_{t_1=-\frac{T}{2}}^{\frac{T}{2}} \int_{t_2=-\frac{T}{2}}^{\frac{T}{2}} \mathbb{E}[X(t_1)X^*(t_2)]\delta(t_2 - t_1) dt_1 dt_2, \end{aligned} \tag{2.2.12}$$

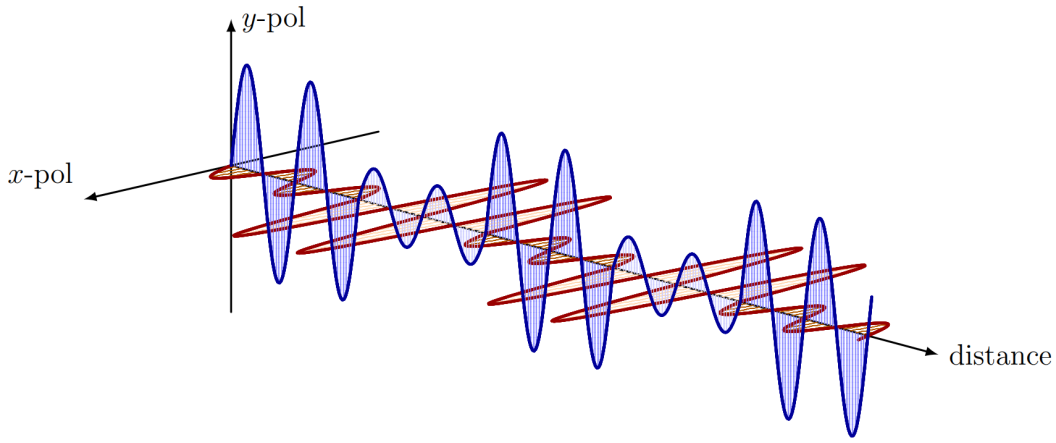


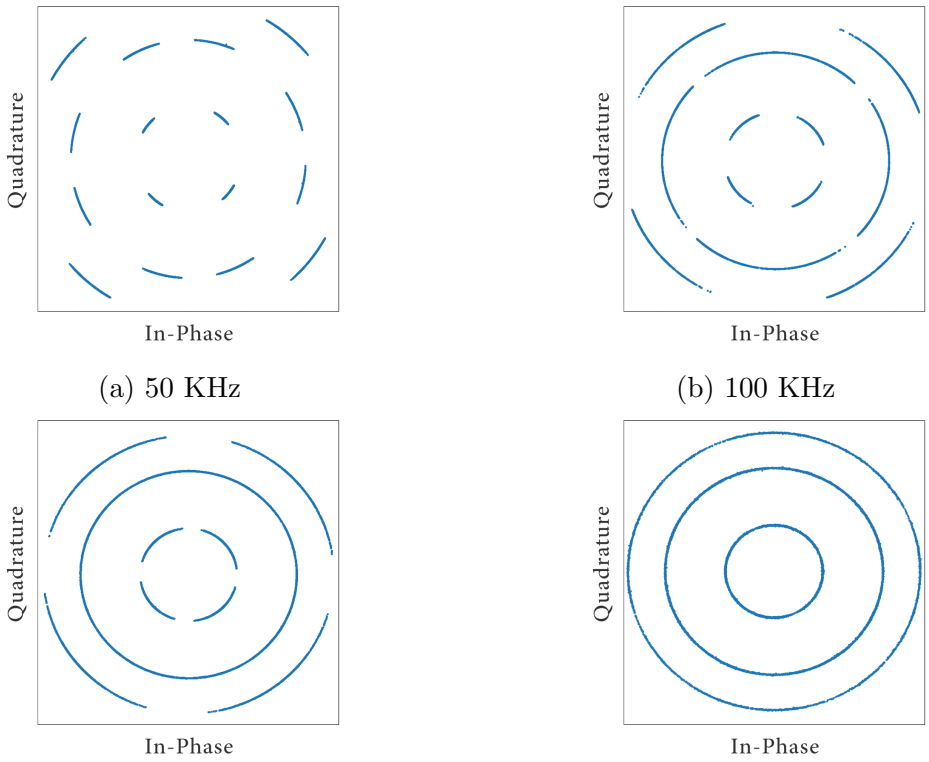
Figure 2.8: Propagation of two ASK signals in distance, in parallel, through polarization-division multiplexing approach.

where $\delta(.)$ is the Dirac-delta generalized function.

2.2.2 Polarization-division multiplexing

In communication systems based on electromagnetic waves, two information channels can be transmitted simultaneously on the same carrier frequency utilizing two orthogonal polarization states. This physical layer approach for multiplexing signals is called polarization-division multiplexing (PDM). In dual-polarization optical fiber communication systems, PDM is used to double the information rate by multiplexing two modulated signals and transmitting them in parallel using the separate left and right circularly polarized optical pulses through the channel. Fig. 2.8 illustrates an ideal dual-polarization propagation, where two modulated signals are propagating in distance simultaneously by being multiplexed on orthogonal polarization states of the space.

In optical fiber communication system, applying PDM also results in some side effects due to the fiber birefringence, termed polarization mode dispersion (PMD), introducing the notion of group differential delay and rapid changes of the state-of-polarization (SOP) over the entire Poincaré sphere, especially in long-haul systems. We discuss these effects more in detail in Section 2.3.



(a) 50 KHz (b) 100 KHz

(c) 500 KHz (d) 5000 KHz

Figure 2.9: Effect of the phase noise on the constellation based on different laser linewidths during 2560 ns observation in 64 GBd transmission. The phase offset level extends towards covering the entire rings by increasing the linewidth.

2.2.3 Laser phase noise

As a matter of fact, the light emitted by a single-frequency laser is not completely monochromatic [5], and it incorporates some phase noise, which can be expressed as the fluctuations in the optical phase. This phenomenon results in a finite linewidth [6] of the laser output.

Spontaneous emission and quantum noise are the contributors to the laser phase noise. Additionally, there is a potential for the presence of technical noise impacts, such as those resulting from vibrations in the resonator mirrors or temperature changes in the gain medium [6], [42, Ch. 10].

[5]Monochromatic light is defined as a light with just one optical frequency in its optical spectrum.
[6]A laser's linewidth is the width of the power spectral density of the emitted electric field

The laser phase noise is typically modeled by a Wiener process, where the phase noise rotation of s_p^i (the symbol at time-step i on the polarization p), denoted by θ_p^i, is realized by

$$\theta_p^i = \sum_{k=1}^{n} \Theta_k \tag{2.2.13}$$

where $\Theta_k \sim \mathcal{N}(\mu, \sigma_\Theta^2)$, in which

$$\sigma_\Theta^2 = 2\pi\Delta\nu T_e, \tag{2.2.14}$$

where $\Delta\nu$ denotes the laser linewidth and T_e is the sample period. Fig. 2.9 shows the effect of the laser phase noise on a 16-QAM signal using the constellation in Fig. 2.5b based on different laser linewidth. As it is clear in the figures, by the rise of laser linewidth, the phase offset spans toward covering the entire rings.

2.3 Channel model of dual-polarization optical fiber transmission systems

In dual-polarization optical fiber communication systems, following the PDM process, the multiplexed signal is transmitted to the optical fiber. As discussed, due to the attenuation, the optical fiber is split into a number of segments called spans, where at the end of each, an amplification process compensates for the fiber loss by optical amplifying the signal (see Fig. 2.10), which is also accompanied by imposing amplified spontaneous emission (ASE) noise to the signal.

2.3.1 Signal propagation in one span of single-mode optical fiber

Propagation of signals in two polarizations of the electric field over one span of SSMF can be modeled by the coupled nonlinear Schrödinger equation (CNLSE) [43], [44],

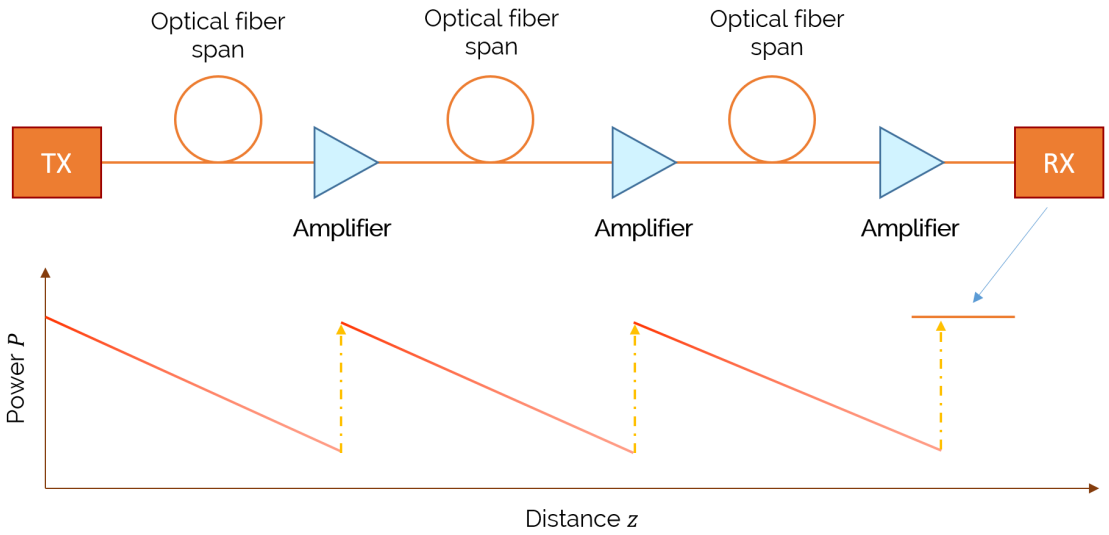


Figure 2.10: Schematic of an optical link. The optical link is divided into a number of spans. At the end of each span, an amplifier compensates for the fiber loss by amplifying the signal.

which describes the interactions between the two states of polarization along the fiber. The equation for the x-polarization is [6, Chap. 6.1]

$$\frac{\partial q_x(t,z)}{\partial z} = -\frac{\alpha}{2}q_x - \beta_{1x}\frac{\partial q_x}{\partial t} - \frac{j\beta_2}{2}\frac{\partial^2 q_x}{\partial t^2} + j\gamma\left(|q_x|^2 + \frac{2}{3}|q_y|^2\right)q_x. \tag{2.3.1}$$

Here, $q_x(t,z)$ is the complex envelope of the signal in the x polarization propagating in the fiber as a function of time t and distance z, α and β_2 are respectively the attenuation and CD coefficients, and γ is the nonlinearity parameter. The first-order dispersion coefficient β_{1x} depends on the polarization, giving rise to DGD and PMD effects. [4, Sec. VI.C]. The equation for the y-polarization is similar to (2.3.1) upon swapping x and y in (2.3.1).

Equation (2.3.1) is numerically solved using the split-step Fourier method (SSFM) [8] with distributed PMD as follows. The fiber span is divided into K segments of length δ. In each segment i, a linear, PMD, and nonlinear step is performed consecutively as follows.

1. **Linear step**

Loss and CD are applied in the frequency domain to the signal as

$$\hat{q}_x(\omega, z) \mapsto \exp\left(-\frac{\alpha}{2}\delta + j\frac{\beta_2}{2}\omega^2\delta\right)\hat{q}_x(\omega, z), \tag{2.3.2}$$

where $\hat{q}_x(\omega, z)$ is the Fourier transform of $q_x(t, z)$.

2. **PMD step**

To model the distributed PMD, a unitary matrix $J^{(i)}(\omega)$ is applied to the signal vector $\mathbf{q}(t, z) = [q_x(t, z), q_y(t, z)]^{\mathrm{T}}$ in the frequency domain as

$$\hat{\mathbf{q}}(\omega, z) \mapsto J^{(i)}(\omega)\hat{\mathbf{q}}(\omega, z), \tag{2.3.3}$$

where $\hat{\mathbf{q}}(\omega, z)$ is the Fourier transform of $\mathbf{q}(t, z)$, and

$$J^{(i)}(\omega) = R^{(i)}D^{(i)}(\omega), \tag{2.3.4}$$

in which

$$R^{(i)} = \begin{pmatrix} e^{j\frac{\phi_i}{2}}\cos(\theta_i) & e^{-j\frac{\phi_i}{2}}\sin(\theta_i) \\ -e^{j\frac{\phi_i}{2}}\sin(\theta_i) & e^{-j\frac{\phi_i}{2}}\cos(\theta_i) \end{pmatrix}, \tag{2.3.5}$$

is the rotation matrix, where $(\theta_i)_{i=1}^K$ and $(\phi_i)_{i=1}^K$ are sequences of independent and identically distributed (iid) random variables drawn from a uniform distribution on $[0, 2\pi)$. In consequence, the state of polarization (SOP) is uniformly distributed over the surface of the Poincaré sphere along the fiber. Further, $D^{(i)}(\omega)$ is the DGD operator

$$D^{(i)}(\omega) = \begin{pmatrix} e^{-j\omega\frac{\tau^{(i)}}{2}} & 0 \\ 0 & e^{j\omega\frac{\tau^{(i)}}{2}} \end{pmatrix}, \tag{2.3.6}$$

where $(\tau^{(i)})_{i=1}^K$ are DGD parameters, which we assume to be a sequence of iid random variables drawn from the probability distribution $\mathcal{N}(0, \tau\sqrt{\delta})$, where τ (measured in $\mathrm{ps}/\sqrt{\mathrm{km}}$) is the PMD parameter.

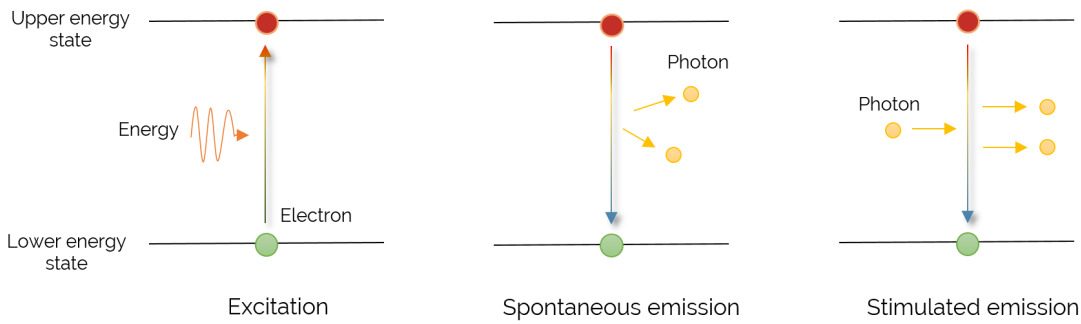


Figure 2.11: Spontaneous and Stimulated emission. In spontaneous emission, there is no interaction with other photons, while in stimulated emission, there is, and the phase and direction are copied.

3. **Nonlinear step**

Kerr nonlinearity effect for x-polarization is modeled in the time domain as

$$q_x(t,z) \mapsto \exp\left(j\gamma\delta\left(|q_x(t,z)|^2 + \frac{2}{3}|q_y(t,z)|^2\right)\right) q_x(t,z). \tag{2.3.7}$$

For the y-polarization, the process is similar by swapping x and y in (2.3.7).

2.3.2 Optical amplification

Following each optical fiber span, there is an optical amplification process to compensate for the attenuation. Optical amplification is commonly achieved by stimulated emission when the rare-earth-doped fiber [7] in the amplifier is optically pumped [8] with a light from a different source, such as a laser diode.

Once an optical gain medium is optically pumped, the electrons within the medium are excited from a state with a lower energy level to a state with a higher energy level. Because higher energy states are generally less stable than those of lower energy, radiated energy eventually relaxes the excited medium to a lower energy state. This

[7] A kind of glass fiber used in fiber amplifiers and laser, which is doped with laser-active rare earth ions

[8] Optically pumping entails injecting light into a medium in order to electronically excite it (or a group of its elements) into a higher energy state.

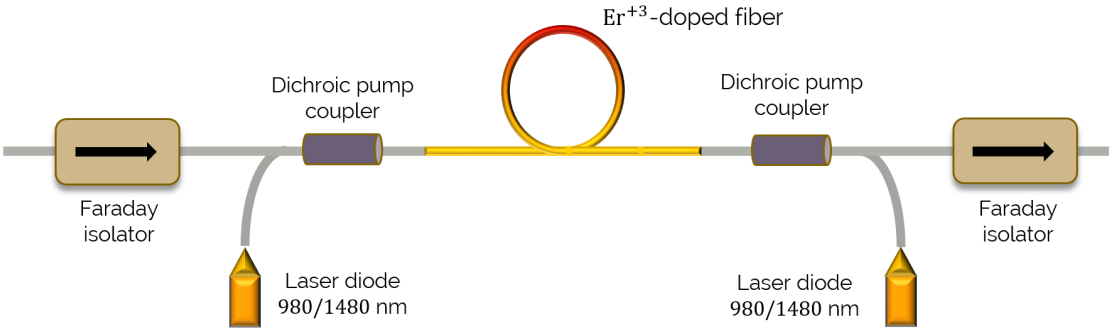


Figure 2.12: Schematic structure of a simple EDFA amplifier. The pump power is supplied by two laser diodes. The pump light is injected through dichroic fiber couplers. Faraday optical isolators are used at the two ends to reduce the back-reflection sensitivity.

energy radiation may take the form of photon emission. When this emission occurs without contact with other photons, it is referred to as spontaneous emission. The direction and phase of the released photons are random in this scenario. On the other hand, it is termed stimulated emission if the emission occurs when the excited electron interacts with another photon. In this particular instance, the direction, as well as the phase, are taken from the photons that have interacted [6, Ch. 3.1], [45], resulting in amplification of the optical signal. This conceptualization is shown graphically in Figure 2.11.

According to this mechanism, the commonly adopted type of amplifiers (EDFA) in fiber-optic links are Erbium-doped fiber amplifiers. The following section describes these amplifiers.

2.3.2.1 Erbium-doped fiber amplifiers

When it comes to long-haul optical fiber communications, EDFAs are by far the most adopted fiber amplifiers. EDFAs have the ability to effectively amplify light in the wavelength area of 1550 nm, where the optical fiber incurs the lowest loss.

In EDFA, amplification is conducted through an Erbium-doped optical fiber which is pumped with light from the laser diodes. A general setup for an EDFA is illustrated in Fig. 2.12. The pump light, which is typically 980 nm or 1480 nm in wavelength, is used to excite the Erbium ions (Er^{3+}) into an excitation state.

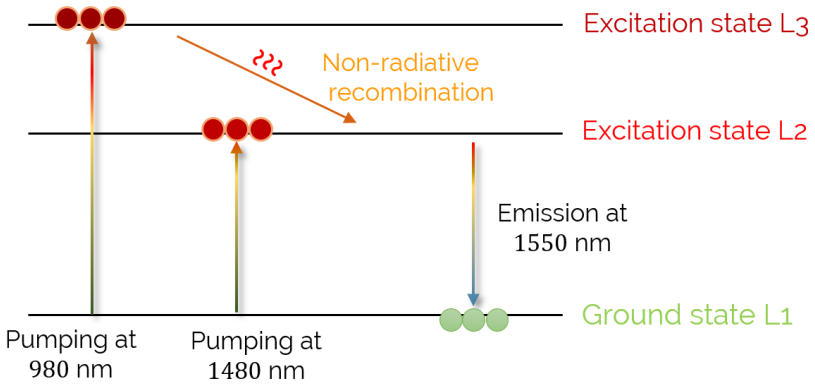


Figure 2.13: Erbium excitation diagram for stimulated emission

When pumped at 1480 nm, Er^{3+} absorbs the pump light and is excited from the ground level L1 into the excited state L2, as shown in Fig. 2.13. In this case, amplification by stimulated emission occurs at around 1550 nm when there is sufficient pump power supplied to the fiber and a population inversion [9] is generated between the ground state (L1) and the excited state L2.

In pumping at 980 nm, Er^{3+} is excited from the ground level L1 into the excited state L3, as it is shown in Fig. 2.13. Following this excitation, however, due to the short lifetime of the excited state L3, Er^{3+} is quasi-instantly relaxed to the excited state L2. Upon this relaxation, a population inversion is generated between the ground level and excited state L2, which results in the amplification occurring at 1550 nm.

EDFA amplification, however, is accompanied by introducing some noise to the signal, mainly originating from amplified spontaneous emission (ASE) noise, which consecutively diminishes the signal-to-noise ratio. The following subsection describes this process.

2.3.2.2 Amplified spontaneous emission noise effect

In amplification by EDFAs, and generally amplifiers, following pumping, spontaneous emission usually happens. When spontaneous emission coincides with stimulated emission, it is amplified through it. This phenomenon, which is called amplified spontaneous emission, introduces noise to the signal and has a destructive effect on the system

[9] Population inversion refers to the condition where in a system, such as a group of atoms, a greater proportion of the members are in higher, excited energy states than in lower, unexcited energy states

performance. This process is modeled by the addition of zero-mean white complex circularly symmetric Gaussian noise with the power spectral density per polarization

$$N_{ASE} = (G-1)h\nu n_{sp} \tag{2.3.8}$$

where $G = e^{\frac{\alpha}{2}L_{sp}}$ is the amplification gain (L_{sp} is the span length), h is Planck constant, ν is the carrier frequency, and n_{sp} is the spontaneous emission factor ($1 < n_{sp} < \sim 1.5$). The spontaneous emission factor, however, cannot be directly observed, and it is characterized within the reported noise figure (NF) of the amplifier, defined as

$$\mathrm{NF} = 2n_{sp}\frac{G-1}{G}. \tag{2.3.9}$$

NF in (2.3.9) is typically expressed in decibel as $\mathrm{NF}_{\mathrm{dB}} = 10\log_{10}(\mathrm{NF})$.

2.4 Optical receiver and digital signal processing

As the light pulses reach the end of the optical link, it is the role of the receiver to translate them into the corresponding transmitted bitstream. This process involves optical detection and a digital signal processing chain to reverse the deterministic effects and polarization-dependant effects of the channel, a process called equalization.

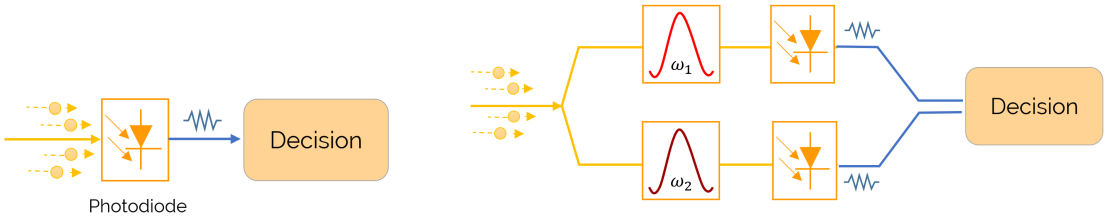
2.4.1 Optical detection

Optical detection can be in the form of non-coherent (direct) or coherent detection [10].

In non-coherent detection, the receiver computes decision variables based on the signal energy. An example of direct detection for ASK signals is shown in Fig. 2.14, where a photodiode [11] releases an electrical current proportional to the energy (intensity) of the optical signal. Then a threshold decision element determines the value of the

[10] Not limited to, e.g., differentially coherent detection and a hybrid of non-coherent and differentially coherent detection

[11] A photodiode is a semiconductor that loses electrons when hit by a photon of a certain wavelength. Electrons are released when the photodiode is hit by light energy.



(a) (b)

Figure 2.14: Non-coherent detection for : a) ASK signals, b) binary FSK

corresponding symbol according to the energy of the electrical current. For FSK signals, this follows the separation of frequency levels, as it is shown in Fig. 2.14b.

As a matter of fact, adopting non-coherent detection scheme results in two main flaws: a) Energy-based detection allows signals to encode just one degree of freedom per polarization per carrier, hence decreasing spectral and power efficiency; b) Loss of phase information during detection is an irreversible transition that precludes linear filters from fully equalizing linear channel impairments such as CD and PMD. Considering these issues, non-coherent detection is not suitable for QAM signals, and long-haul fiber-optic channels [46].

Coherent detection, on the other hand, is based on the recovery of the full electric field, which, when sampled at the Nyquist rate, can be forwarded to the digital signal processing block to compensate for channel impairments. As information can be stored in both amplitude and phase, or alternatively in both in-phase (I) and quadrature (Q) components of a carrier, coherent detection provides the most versatility in modulation formats [46].

In coherent detection, the received signal is mixed with a continuous waveform generated by a local oscillator (LO) oscillating at roughly the same frequency, serving as an absolute phase reference. The overall construction of a coherent optical receiver in dual-polarization systems is shown schematically in Fig. 2.15. In this structure, initially, the polarizations of both the incoming optical signal and the waveform produced by LO are split using polarization beam splitters (PBSs). To recover the in-phase and quadrature components of the signal, each polarization enters a 90° optical hybrid, allowing coherent mixing of the LO (adjusted at the carrier frequency) with the incident x- and y-polarization components of the optical signal. As Fig. 2.15 shows, the output of the optical hybrids is then linked to balanced photodiodes, which provide two

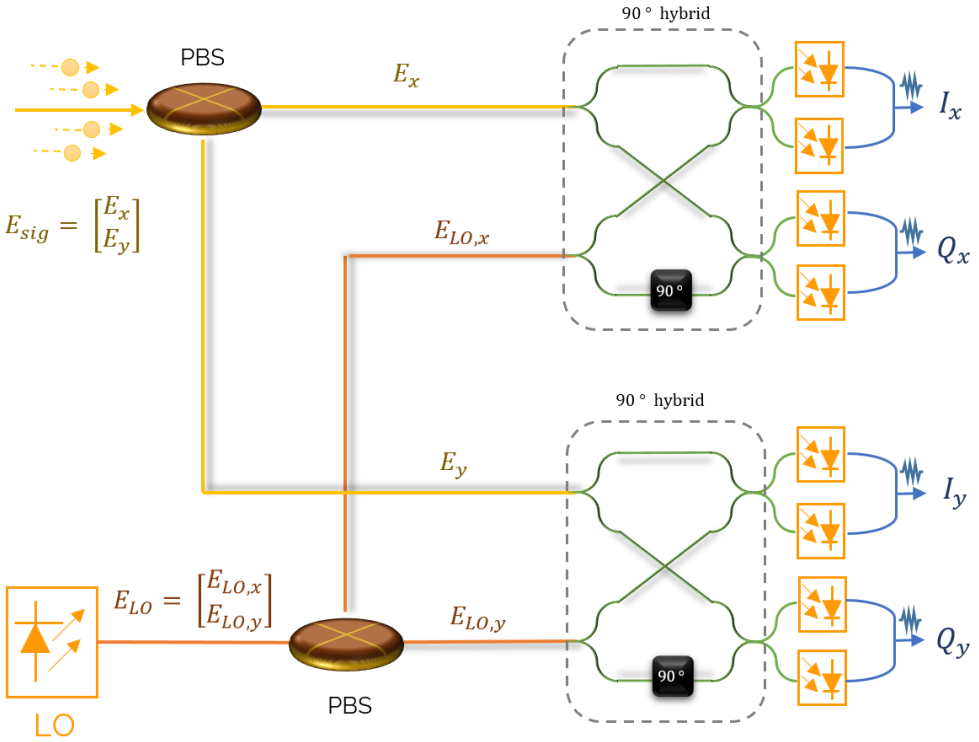


Figure 2.15: General structure of a dual-polarization detector. receiver

currents corresponding to the I and Q components of the optical signal corresponding polarization, given by

$$I_Q(t) = \tau_{PD}\sqrt{P_{sig}P_{LO}}\,cos(\theta_{sig}(t) - \theta_{LO}(t)) \tag{2.4.1}$$

$$I_I(t) = \tau_{PD}\sqrt{P_{sig}P_{LO}}\,sin(\theta_{sig}(t) - \theta_{LO}(t)) \tag{2.4.2}$$

where P_{sig}, P_{LO}, θ_{sig}, θ_{LO} are the powers and phases of the received signal and the local oscillator, respectively. τ_{PD} is the photodetection responsivity coefficient.

Consequently, the complex amplitude or the baseband signal is represented as

$$\begin{aligned} I_c(t) &= I_I(t) + jI_Q(t) \\ &= \tau_{PD}\sqrt{P_{sig}P_{LO}}\,e^{j(\theta_{sig}(t)-\theta_{LO}(t))}. \end{aligned} \tag{2.4.3}$$

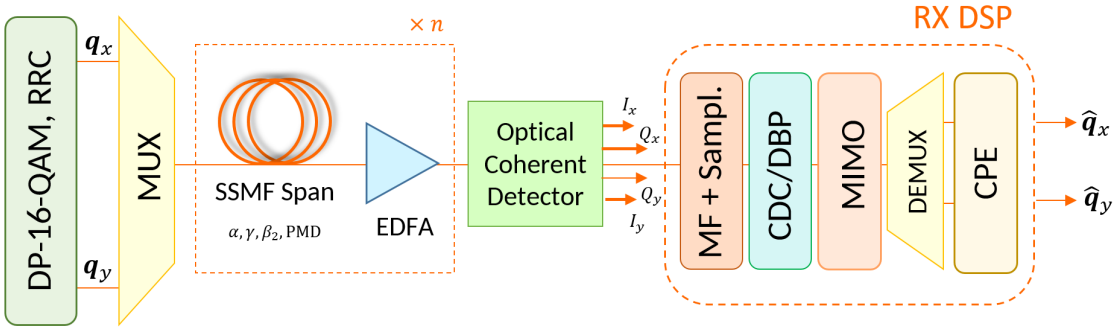


Figure 2.16: Conventional DSP chain in long-haul optical fiber communication system considered in this manuscript.

2.4.2 Digital signal processing in dual-polarization optical fiber communication

Following the full translation of the optical signal to the electric field, the signal is passed to a digital signal processing (DSP) chain, as Fig. 2.16 shows, to compensate for the channel impairments discussed in the previous section, a process referred to as equalization. The linear DSP incorporates [12] matched filtering, chromatic dispersion compensation, multiple-input multiple-output (MIMO) -based PMD compensation, and carrier phase estimation. The linear DSP, however, can only reverse the linear distortions, while nonlinear impairments require nonlinear equalization methods. Digital backpropagation is a primary method to be leveraged in the DSP chain in place of the CD compensation block to reverse the chromatic dispersion and nonlinear impairments jointly. In the following sections, we concisely describe these processing blocks.

2.4.2.1 Matched filtering

Following the coherent optical detection, matched filtering is applied to maximize the signal-to-noise ratio (SNR) by correlating the received waveform with a linear matched filter, u, as

$$y[n] = \sum_{k=-\infty}^{\infty} u[n-k]x[k], \tag{2.4.4}$$

[12] Not limited to.

(a) w/o matched filtering (b) With matched filtering

Figure 2.17: The effect of matched filtering in resulting constellation output by RX in an AWGN channel. The same is applicable to fiber-optic transmission systems.

where u is the time-reversed complex conjugate of the pulse shape [13], and $x[k]$ and $y[n]$, respectively, represent the input as a function of the independent variable k and the filtered output.

The intuition underlying an ideal matched filter is to correlate the received signal with a filter that is parallel with it, maximizing the signal, while being towards orthogonality to the noise, minimizing the noise effects. Fig. 2.17 shows an example of the final constellation output by RX in an AWGN communication system in the presence and absence of the matched filtering.

2.4.2.2 Chromatic dispersion compensation

Following matched filtering, in the linear DSP chains, there exists CD compensation block. This processing unit performs based on reversing the dispersion effect throughout the fiber, in a one-shot process as

$$\hat{q}_x(\omega, \mathfrak{L}) \mapsto \exp\left(-j\frac{\beta_2}{2}\omega^2\mathfrak{L}\right)\hat{q}_x(\omega, \mathfrak{L}), \tag{2.4.5}$$

where $\hat{q}(\omega, \mathfrak{L})$ is the Fourier transform of $q(t, \mathfrak{L})$, and $\mathfrak{L}$ is the fiber length.

[13]The RRC filter is its own matched filter.

2.4.2.3 MIMO-based equalization

Unlike CD, which can be regarded as a static effect, polarization-dependent effects are time-varying and thus require adaptive filters. As discussed, PMD effects results in a 2×2 frequency-dependent transfer function matrix. Therefore, following the CD (and nonlinearity) compensation, a complex-valued 2×2 MIMO finite impulse response (MIMO-FIR) filter is applied to reverse the PMD effects, such that the outputs are obtained as

$$\begin{aligned} q_x^{out}(t=k) &= \mathbf{h}_{xx}^H q_x^{in}(t=k', \ldots, t=k'+N_{\text{taps}}) + \mathbf{h}_{xy}^H q_y^{in}(t=k', \ldots, t=k'+N_{\text{taps}}) \\ q_y^{out}(t=k) &= \mathbf{h}_{yx}^H q_x^{in}(t=k', \ldots, t=k'+N_{\text{taps}}) + \mathbf{h}_{yy}^H q_y^{in}(t=k', \ldots, t=k'+N_{\text{taps}}), \end{aligned} \tag{2.4.6}$$

where q_p^{in} is the signal of polarization p input to the MIMO block, q_p^{out} is the corresponding output signal of MIMO for polarization p, N_{taps} is the number of taps (covering the PMD memory) and $\mathbf{h}_{xx}^H, \mathbf{h}_{xy}^H, \mathbf{h}_{yx}^H, \mathbf{h}_{yy}^H$ are the N_{taps}-length vectors representing the tap weights. $k' = k \times \text{SpS}$, where SpS denotes the number of sample per symbol (note that the output signal is at 1 SpS). $\mathbf{h}^H$ is the conjugate transpose of $\mathbf{h}$. Equation (2.4.6) for polarization p can be rewritten as

$$z_p(k) = \mathbf{h}_p^H \mathbf{q}^{in}(k), \tag{2.4.7}$$

where $z_p(k)$ denotes $q_p^{out}(t = k)$, $\mathbf{h}_p^H$ denotes the concatenation of corresponding $\mathbf{h}$ vectors for polarization p, and $\mathbf{q}^{in}(k)$ denotes the concatenation of corresponding $\mathbf{q}^{in}$ vectors.

Constant modulus algorithm (CMA) and radius-directed equalization (RDE) are two of the main conventional PMD equalization and demultiplexing techniques to determine the optimal weights for respectively PSK and QAM signals [47].

In CMA, the constant modulus criteria is used to determine the tap weights, which are determined by minimizing the following cost function

$$L_{\text{CMA}}(\mathbf{h}_p) = \mathbb{E}[L_{\text{CMA},k}(\mathbf{h}_p)], \tag{2.4.8}$$

where

$$L_{\mathrm{CMA},k}(\mathbf{h}_p) = (|z_p(k)|^2 - R)^2, \tag{2.4.9}$$

in which

$$R = \frac{\mathbb{E}[|s_p(k)]|^4}{\mathbb{E}[|s_p(k)]|^2}, \tag{2.4.10}$$

where s_p is the sequence of symbols in the polarization p.

CMA is typically implemented via a stochastic gradient descent approach, where the filter coefficients at each sample are updated iteratively as follows

$$\mathbf{h}_p^{<k+1>} = \mathbf{h}_p^{<k>} - \alpha_{lr} \nabla L_{\mathrm{CMA},k}(\mathbf{h}_p^{<k>}), \tag{2.4.11}$$

where α_{lr} is the learning rate, $\nabla L_{\mathrm{CMA},k}(\mathbf{h}_p^{<k>})$ is the gradient at time step k, and $\mathbf{h}_p^{<k>}$ is the tap weights at time step k. The gradient is realized to be

$$\nabla L_{\mathrm{CMA},k}(\mathbf{h}_p^{<k>}) = (|z_p(k)|^2 - R)\overline{z_p(k)}\mathbf{q}^{in}(k). \tag{2.4.12}$$

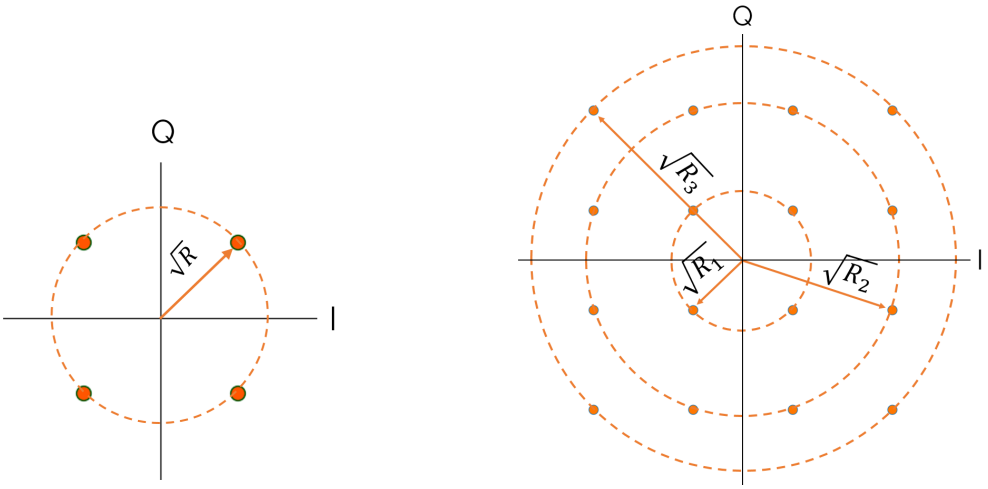
Although CMA is efficient for BPSK and QPSK, it is not optimal for QAM signals, such as 16-QAM, in the sense that the error does not go to zero when the equalizer has fully converged [48]. The RDE criterion is a generalization of the CMA algorithm for QAM signals, where the modulus of the constellation is not constant (see Fig. 2.18). In RDE, in particular, for 16-QAM signals [14], the cost function is considered as

$$L_{\mathrm{RDE}}(\mathbf{h}_p) = \mathbb{E}[L_{\mathrm{RDE},k}(\mathbf{h}_p)], \tag{2.4.13}$$

where

$$L_{\mathrm{RDE},k}(\mathbf{h}_p) = (|z_p(k)|^2 - R_0)^2, \tag{2.4.14}$$

[14] generalizable for higher order with a similar approach



(a) CMA for QPSK (b) RDE for 16-QAM

Figure 2.18: Radii of QPSK and 16-QAM constellations considered in CMA and RDE algorithm.

where if $|z_p(k)| < (\sqrt{R_1} + \sqrt{R_2})/2$ then $R_0 = R_1$, if $|z_p(k)| > (\sqrt{R_2} + \sqrt{R_3})/2$ then $R_0 = R_3$, otherwise, $R_0 = R_2$.

Realizing the tap weights using gradient descent in RDE is similar to what was discussed for CMA.

2.4.2.4 Carrier phase estimation

As discussed, the phase offset is another type of effect that the fiber-optic transmission systems deal with. Phase error and phase noise are the phase offset effects, respectively, resulting from channel deterministic effects and non-deterministic random effects from laser sources associated with the laser linewidth, as discussed in the previous sections. In order to compensate for these effects, a process termed carrier phase estimation (CPE) should be performed. CPE is formally defined as recovering the phase offset θ as follows

$$s(k) = e^{j\theta}\hat{s}(k) \tag{2.4.15}$$

where $\{s(k)\}_{k=1}^{N_s}$ are the transmitted symbols and $\{\hat{s}(k)\}_{k=1}^{N_s}$ is the resulting constellation after assuming ideally compensating all the channel impairments.

A variety of CPE algorithms have been proposed [49], [50]. However, in the case of non-blind recovery using pilot symbols, it suffices to calculate the angle between the transmitted symbols and the pilots over a sliding window and rotate the symbols of that window according to the mean angle between the pilots and transmitted symbols.

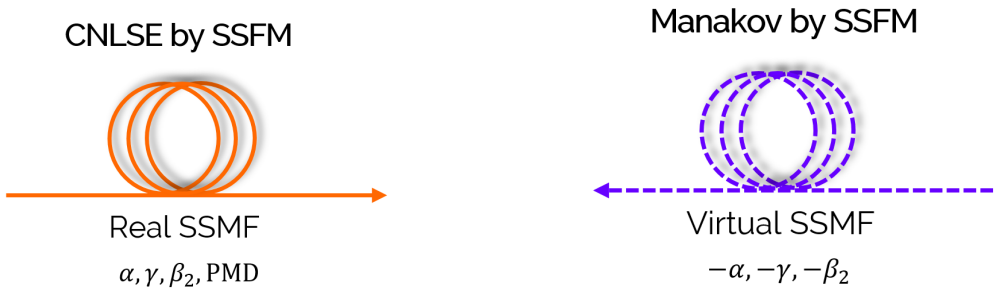
The output of CPE is then forwarded to the detection block, which is a simple block assigning the equalized symbol to the nearest symbol in the constellation, and subsequently de-maps the symbol to the bits.

2.4.2.5 Nonlinearity mitigation through digital backpropagation

As discussed, the linear DSP chain, which is composed of the processing units described in the previous subsections, is merely able to compensate for the linear channel effects.

Digital backpropagation (DBP) [7] is a major technique to jointly compensate the chromatic dispersion and nonlinear impairments, and it is implemented in the place of the CD compensation block in the DSP chain.

DBP reverses the deterministic effects of the channel by propagating the signal backward in distance using negated parameters in SSFM (see Fig. 2.19), without integrating PMD effects, thus only linear and nonlinear steps.



(a) Forward propagation modeling

(b) DBP

Figure 2.19: DBP's underlying approach. DBP uses SSFM with inverse parameters to reverse the deterministic effect of the channel

DBP, however, uses the Manakov equation as a reference for propagation, as in most cases, it considers the step sizes considerably larger than the one used for the forward propagation modeling. This is in view of the fact that, as discussed in 2.3, the birefringence variations change the SOP rapidly such that the field covers the entire Poincaré

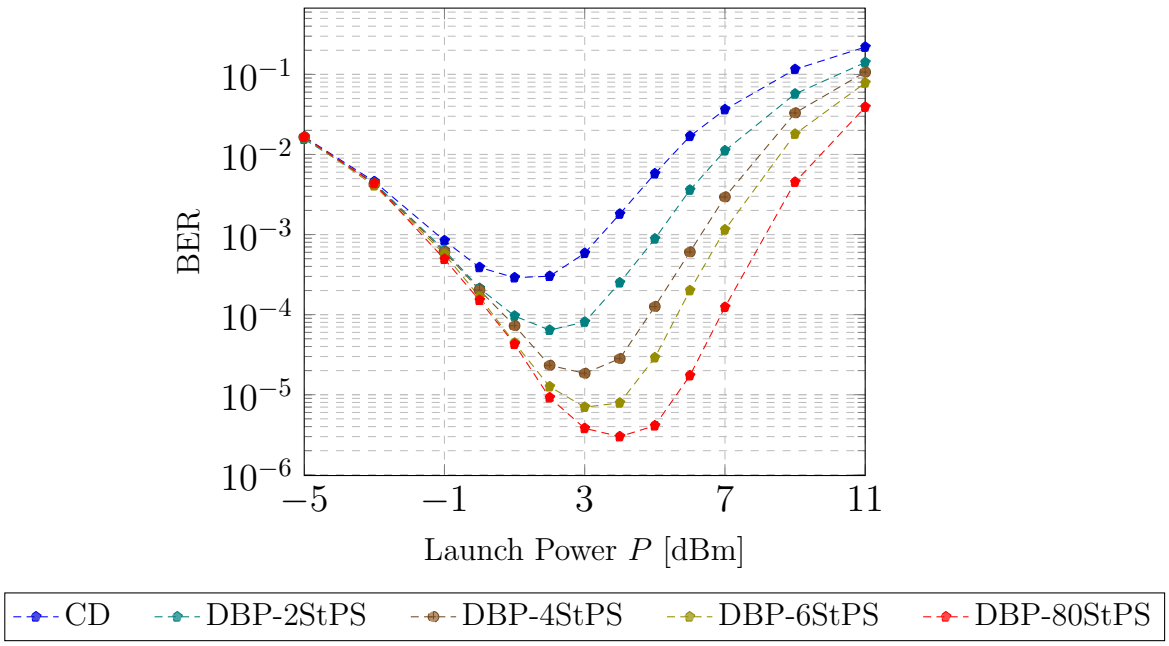


Figure 2.20: Corresponding BER performances of the receiver, based on using DBP with different StPS values, as a function of the total launch power.

sphere after a few kilometers. This results in the possibility of averaging the nonlinearity terms in CNLSE over the birefringence variations. The resulting propagation equation is called the Manakov-PMD equation [51]

$$\frac{\partial \mathbf{q}(t,z)}{\partial z} = -\frac{\alpha}{2}\mathbf{q} - \beta_1(z)\frac{\partial \mathbf{q}}{\partial t} - \frac{j\beta_2}{2}\frac{\partial^2 \mathbf{q}}{\partial t^2} + j\frac{8}{9}\gamma|\mathbf{q}|^2\mathbf{q}, \tag{2.4.16}$$

where $\mathbf{q}(t,z) = [q_x(t,z), q_y(t,z)]^{\mathrm{T}}$. Thus, in the nonlinear step, DBP reverses the nonlinearity effect at each nonlinear step of SSFM as

$$\mathbf{q}(t,z) \mapsto \exp\left(j\frac{8}{9}(-\gamma)\delta|\mathbf{q}(t,z)|^2\right)\mathbf{q}(t,z). \tag{2.4.17}$$

The linear step would be the same as (2.3.2) with negated β_2 parameter.

Depending on the step sizes in DBP or equivalently the number of step per span considered for it, DBP results in different performances. Fig. 2.20 shows the performance of different step sizes in DBP as well as the performance of CD compensation in a

considered dual-polarization fiber-optic communication system [15].

2.5 Conclusion

In this chapter, we discussed the principles of dual-polarization fiber-optic communication systems, including the physical properties of fiber-optic channels leading to different types of distortions, the attenuation phenomenon and the consequent optical amplification in optical links, mathematical modeling of the dual-polarization optical fiber channel with distributed PMD effects, common transmission schemes in dual-polarization communication systems and modulation formats, optical detection at receiver, and the conventional DSP chain to compensate the channel impairments in the electrical field.

We demonstrated that unlike linear effects such as CD and PMD effects which could be compensated through a linear DSP block at low complexity, nonlinear impairments is a major obstacle in optical fiber communications which requires a heavy computational process at the receiver to be partially compensated.

With regard to this matter, in recent years, neural networks have been widely integrated into the DSP chain to either optimize or replace the nonlinearity mitigation component in the receiver. In the following chapter, we elaborate on the principles of these structures, their underlying rationale enabling them to approximate complex functions, and a variety of neural network models within this field.

[15]The specification and fiber parameters of this system is described in Section. 5.1

Chapter 3

Learning by neural networks: principles and fundamental models

Artificial neural networks, commonly referred to as neural networks, are a sub-field of machine learning. In the field of machine learning, which is itself a sub-category of artificial intelligence, the attempts are towards finding a sort of algorithms that enables computer machines to learn and advance from past experiences without the need to be explicitly programmed [52]. These algorithms start with observations or data in a vector space E to finally output a determination or prediction in the same or another vector space K. Machine learning algorithms are commonly categorized as supervised or unsupervised.

In supervised learning, the vector space E of all possible inputs, the vector space K of all possible outputs, and the known ground truth input-output pairs (x_i, y_i) are taken first. Then the algorithm seeks the best function, $f : E \to K$, predicting the output $y \in K$ for an input $x \in E$. In order to evaluate the quality of obtained function f, a cost or loss function is defined, denoted by $L(Y, f(X))$, to penalize the prediction error, where $X \in E$ is a random variable, and $Y \in K$ is the corresponding ground truth. The loss function produces a criterion for choosing f, called expected prediction error (EPF). In the case of choosing the square error loss function, which is the most common loss function among the variety of adopted loss functions, EPF will form as follows [1] [53, Ch. 2]:

[1] Without loss of generality, X and Y are assumed be continuous over the fields.

$$\begin{aligned} \mathrm{EPF}(f) &= E[(Y - f(X))^2] \\ &= \int [y - f(x)]^2 Pr(x, y) dx dy. \end{aligned} \tag{3.0.1}$$

By conditioning on X, EPF can be written as

$$\mathrm{EPF}(f) = E_X E_{Y|X}([Y - f(X)]^2 \,|\, X). \tag{3.0.2}$$

Hence, it suffices to minimize EPF point-wise:

$$f(x) = \mathrm{argmin}_c E_{Y|X}([Y - c]^2 \,|\, X = x). \tag{3.0.3}$$

This yields the solution

$$f(x) = E(Y \,|\, X = x), \tag{3.0.4}$$

the conditional expectation (often referred to as the regression function [2]). Therefore, when the best evaluation metric is (mean) squared error, the conditional mean is the best prediction of Y at any point $X = x$.

Neural networks, which are inspired by the biological neural networks constituting the brain, are a branch of supervised learning which attempts to figure out (3.0.4), or any function leading to the minimum EPF, by a topology of basic processing units structured as a directed computational graph performing concatenated linear and nonlinear operators to the inputs.

The concept of neural networks was initially introduced in 1943 by neurophysiologist Warren McCulloch and mathematician Walter Pitts [17], and it has been in and out of favor for more than 70 years. However, following recent spectacular breakthroughs in this area, neural networks have abruptly attracted attention nowadays, and they have been extensively used in several fields of research to get superior outcomes.

Neural networks are organized into layers that are linked. Each layer is made up of a number of nodes known as hidden units or neurons, each of which has an activation

[2] In this manuscript, without loss of generality, we mainly focus on regression rather than classification, considering the study area of this thesis and the research carried out within the doctoral program.

function. Training samples are fed into the neural network through the input layer, where they are routed to the hidden layers, where they are processed using a weighted connection mechanism. Finally, the output layer delivers the answer.

A node's activation function, $\sigma(.)$, introduces nonlinearity to it. This is to address the so-called expressive power of the neural network [54]. As a result, the neural network functional form conforms to the hierarchical or Markovian structure. Most data of practical interest is generated by some form of these processes [55]. Table 3.1 presents a number of commonly used activation functions.

The majority of neural networks follow some form of learning algorithm that constantly corrects the weights of connections depending on the calculated final loss of the provided batch of training examples, a process named Gradient Descent [56]. After a cycle of forward activated flow of outputs and backward propagation of error, known as backpropagation [30], a correction is performed. Backpropagation generates the steepest descent through the error surface by performing a gradient over the vector space K towards the global/local minimum. Simply said, a neural network sets its weights at random at first, then compares its estimate to the ground truth depending on the input according to the loss function (a number of commonly used loss functions are presented in Table 3.2). The result is then used to make the correction through backward propagation of the calculated error. Mathematically saying, at each iteration t, following the backpropagation process, the weights are updated as follows

$$W_{\ell}^{<t+1>} = W_{\ell}^{<t>} - \alpha_{lr} \nabla \hat{L}(W_{\ell}^{<t>}), \tag{3.0.5}$$

where $W_{\ell}^{(t)}$ is the weight matrix of layer ℓ at iteration t, $\hat{L}$ is the resulting loss by the current weights, $\nabla \hat{L}$ is the gradient of $\hat{L}$, and $\alpha_{lr} > 0$ is the learning rate.

As noted, neural networks form a directed computational graph applying concatenated linear and nonlinear operators to the inputs. In a general categorization, depending on how information is propagated from the input layer to the output, neural networks are classified into two groups:

- **Feed-forward Neural Networks** (FNNs). Each neuron in an FNN is only connected to neurons in the following layer. As a result, information solely propagates forward from the input layer to the output layer, with no feedback loop.

Table 3.1: Common activation functions adopted in neural networks and their corresponding derivative (for backpropagation). The parameter α in ELU and Leaky ReLU is a hyper-parameter.

Name	**Function** $\sigma(x)$	**Derivative** $\sigma'(x)$
Sigmoid	$\frac{1}{1+e^{-x}}$	$\sigma(x)(1-\sigma(x))$
Gaussian	e^{-x^2}	$-2xe^{-x^2}$
Hyperbolic tangent (tanh)	$\frac{e^x-e^{-x}}{e^x+e^{-x}}$	$1-\sigma(x)^2$
Exponential linear unit (ELU)	$\begin{cases} \alpha\left(e^x-1\right) & \text{if } x \leq 0 \\ x & \text{if } x > 0 \end{cases}$	$\begin{cases} \alpha e^x & \text{if } x < 0 \\ 1 & \text{if } x > 0 \\ 1 & \text{if } x = 0 \text{ and } \alpha = 1 \end{cases}$
Rectified linear unit (ReLU)	$\begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{if } x > 0 \end{cases}$	$\begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x > 0 \\ \text{undefined} & \text{if } x = 0 \end{cases}$
Leaky rectified linear unit (Leaky ReLU)	$\begin{cases} \alpha x & \text{if } x \leq 0 \\ x & \text{if } x > 0 \end{cases}$	$\begin{cases} \alpha & \text{if } x < 0 \\ 1 & \text{if } x > 0 \\ \text{undefined} & \text{if } x = 0 \end{cases}$

Table 3.2: Common loss functions adopted in the machine learning tasks. In the formulations, y_i denotes the ground truth and $\hat{y}_i = f(x_i)$ denotes the predicted value

Loss function	**Formulation**
Mean Square Error	$\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2$
Mean Absolute Error	$\frac{1}{N}\sum_{i=1}^{N}\lvert y_i - \hat{y}_i\rvert$
Mean Bias Error	$\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)$
Log-Cosh	$\frac{1}{N}\sum_{i=1}^{N}\log(\frac{e^{(y_i-\hat{y}_i)} + e^{-(y_i-\hat{y}_i)}}{2})$
Huber Loss	$L_\delta(y_i, \hat{y}_i) = \begin{cases} \frac{1}{2}(y_i - \hat{y}_i)^2 & \text{for } \lvert y_i - \hat{y}_i\rvert \le \delta, \\ \delta \cdot \left(\lvert y_i - \hat{y}_i\rvert - \frac{1}{2}\delta\right), & \text{otherwise.} \end{cases}$
Binary Cross-entropy (for classification)	$-\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log \hat{y}_i + (1-y_i)\log(1-\hat{y}_i)\right]$

- **Recurrent Neural Networks** (RNNs). In contrast to FNNs, RNNs support feedback connections; thus, information can be forwarded to previous layers in the network through feedback links.

Among the well-established neural network models, multi-layer perceptrons (MLPs), also known as dense neural networks, and convolutional neural networks (CNNs) are the two primary neural network types within the category of the FNNs. Vanilla RNNs, long-short term memory (LSTM), and gated recurrent unit (GRU) networks, on another side, are the main neural network types in the category of RNNs. The topology of the cells in RNN networks can also form a variety of RNN formats, including many-to-many (or seq2seq), many-to-one (many inputs, one output), and one-to-many.

In the following sections of this chapter, we elaborate on these concepts and discuss the underlying structure of each of the mentioned neural networks type.

3.1 Multi-layer perceptrons and universal approximation theorem

A multi-layer perceptron, abbreviated as MLP, is a stack of interconnect layers (fully-connected) on top of each other over a system of weighted connections. In MLPs, connections between the nodes do not form a cycle. Each layer ℓ with d_ℓ hidden units performs the mapping $\Psi_\ell : \mathbb{R}^{d_{\ell-1}} \to \mathbb{R}^{d_\ell}$ on the output of the previous layer, $\mathbf{x_{\ell-1}}$, as follows:

$$\begin{aligned} \mathbf{x}_\ell &= \Psi_\ell(\mathbf{x_{\ell-1}}) \\ &= \sigma_\ell\left(W_\ell^T \mathbf{x_{\ell-1}} + \mathbf{b}_\ell\right), \end{aligned} \tag{3.1.1}$$

where $W_\ell \in \mathbb{R}^{d_{\ell-1} \times d_\ell}$ is the weight matrix, $\mathbf{b}_\ell \in \mathbb{R}^{d_\ell}$ is the bias vector, and $\sigma_\ell(.)$ is the activation function. Fig. 3.1 illustrates a schematic of an MLP with an input layer compromising 6 hidden units, 2 hidden layers, and an output layer having one unit outputting the answer.

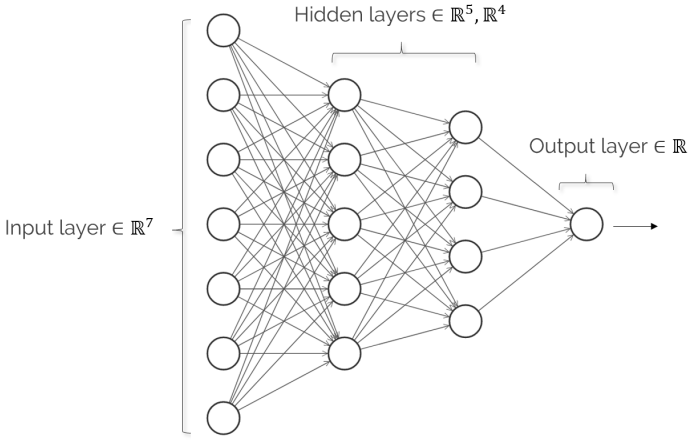


Figure 3.1: Schematic of an MLP comprising the input layer with 7 hidden units, 2 hidden layers having 5 and 4 hidden units, respectively, and the output layer with 1 hidden unit to output a single value.

The underlying logic of these networks is reinforced by the universal approximation theorem. Consider the the topological space $C(K)$, for $k \subset \mathbb{R}^d$, defined as

$$C(K) = \{f : K \to \mathbb{R} : f \text{ continuous}\} \tag{3.1.2}$$

with the uniform norm

$$||f||_\infty = \sup_{x \in K} |f(x)| \tag{3.1.3}$$

Assuming K to be compact, according to the Riesz representation theorem [57, Ch. 6], the topological dual space of $C(K)$ is

$$\mathcal{M} = \{\mu : \mu \text{ is a signed Borel measure on } K\}. \tag{3.1.4}$$

Subsequently, the term *universality* is defined as follows:

Definition 3.1.1. Let MLP(σ, d, N_ℓ) be the set of all MLPs with d-dimensional input, N_ℓ layers, and activation function $\sigma : \mathbb{R} \to \mathbb{R}$, where $d \in \mathbb{N}$, $N_\ell \in \mathbb{N}$ and σ is continuous. Let $K \subset \mathbb{R}^d$ be compact. MLP(σ, d, N_ℓ) is said to be *universal*, if it is dense in $C(K)$ [58].

Definition 3.1.2. Assume $d \in \mathbb{N}$ and $k \subset \mathbb{R}^d$ to be compact. A continuous function $f : \mathbb{R} \to \mathbb{R}$ is termed as *discriminatory*, iff. there exists a measure $\mu \in \mathcal{M}$ such that

$$\int_K f(ax - b)d\mu(x) = 0, \quad \text{for all } a \in \mathbb{R}^d, b \in \mathbb{R} \tag{3.1.5}$$

then $\mu = 0$ [58].

Given the definitions above, the universal approximation theorem is stated as follows:

Theorem 3.1.1 (Universal approximation theorem)**.** *If $d \in \mathbb{N}$, $K \subset \mathbb{R}^d$ be compact, and $\sigma : \mathbb{R} \to \mathbb{R}$ be discriminatory, then* $\mathrm{MLP}(\sigma, d, 2)$ *is universal* [59].

In simple (and approximate) words, the universal approximation theorem asserts that any continuous function $f : \mathbb{R}^n \to \mathbb{R}$ can be approximated arbitrarily well by an MLP with one hidden layer consisting of a finite number of hidden units with a discriminatory activation function [3].

However, despite the universal approximation theorem, it is unknown what the sufficiently large number of hidden units to be set for an MLP is to enable it to approximate a function. An enough required number might be infeasibly large for machines, and it also may lead to overfitting [4] during the training process. Considering these facts, the study of various models of neural networks has always been under consideration.

3.2 Convolutional neural networks

Convolutional neural networks, abbreviated as CNNs, are a category of feed-forward neural networks that are generally composed of convolution layers and sometimes pooling layers. These networks were originally proposed for efficient learning in 2D images, but following the success and underlying solid logic, they were further extended to other fields, including 1-D signal processing.

[3] As an instance, any bounded measurable Sigmoid function is discriminatory

[4] Overfitting refers to producing an analysis that matches too closely or perfectly to just the training data, and thus not well-generalizable to unseen data, leading to failure in reliably predicting future observations.

CNNs are historically inspired by the organization of the animal visual cortex, in which individual cortical neurons only react to stimuli in a narrow section of the visual field known as the receptive field, and different neurons' receptive fields partly overlap, covering the full visual field.

CNNs are mainly composed of a series of convolution layers followed by a number of fully-connected layers. The convolution layers use filters to perform convolution operations on input I, producing a feature-map fed to the next layer. A filter (kernel) is a weight matrix that is learned via iterations of gradient descent during training. For one-dimensional data with N_{ch} channels, the shape of a filter is $L_{ker} \times N_{ch}$, where L_{ker} is the filter size. After performing a convolution on data by applying this filter on it, the mth element of the feature-map in the layer ℓ would be as follows:

$$\mathbf{x}_\ell[m] = \sigma_\ell(\mathbf{z}_\ell[m] + \mathbf{b}_\ell[m]), \tag{3.2.1}$$

where

$$\mathbf{z}_\ell[m] = \sum_{k=1}^{N_c} \sum_{i=1}^{L_{ker}} W[i,k]\mathbf{x}_{\ell-1}[m - \frac{L_{ker}}{2} + i, k], \tag{3.2.2}$$

in which W is the filter weight matrix, and $W[i,k]$ is the entry in the row i and the channel k.

It is possible to stack several feature-maps created by different filters on top of each other to make an overall feature-map. Then in the next layer, each of them is considered as a channel of the data. To streamline the underlying computation in the convolution process, striding [60] could be leveraged with the consideration of the information structure to avoid loss of principal features and only to skip redundant features.

The pooling layer is a downsampling process, usually following a convolution layer. Unlike convolutional layers, pooling layers operate separately on each channel, performing some kind of spatial invariance [31, Ch. 9.3]. Max and average pooling are the most common form of pooling where the maximum and average values, respectively, are taken within the kernel range.

3.3 Recurrent neural networks

RNNs are a category of neural networks designed to operate on temporal sequences of data with correlated samples. These neural networks are composed of recurrent cells whose states are influenced not only by the current input but also by the past or even future (for bi-RNNs) time-steps. This scheme is enabled by the so-called memory of the RNN, which is emulated by the hidden state property. That is to say, the sequential information is maintained in the RNN's hidden state, which is utilized to positively affect the processing of each new time-step as the sequence steps forward.

3.3.1 Vanilla RNNs

A vanilla RNN layer is made up of a number of recurrent cells, governed by the equations

$$\begin{aligned} \mathbf{h}^{(t)} &= \sigma_1(W_h\mathbf{h}^{(t-1)} + W_x\mathbf{x}^{(t)} + \mathbf{b_h}), \\ \mathbf{y}^{(t)} &= \sigma_2(W_y\mathbf{h}^{(t)} + \mathbf{b_y}), \end{aligned} \tag{3.3.1}$$

where $\mathbf{x}^{(t)} \in \mathbb{R}^{n_i}$ and $\mathbf{h}^{(t)} \in \mathbb{R}^{n_h}$ are respectively the input and the hidden state of time-step t; $W_h \in \mathbb{R}^{n_h \times n_h}$, $W_x \in \mathbb{R}^{n_h \times n_i}$, $W_y \in \mathbb{R}^{n_y \times n_h}$ are the weight matrices, $\mathbf{b_h} \in \mathbb{R}^{n_h}$ and $\mathbf{b_y} \in \mathbb{R}^{n_y}$ are the bias vectors, and $\sigma_1(.)$ and $\sigma_2(.)$ are the activation functions.

Thanks to the memory property of RNNs, they are able to capture long-term dependencies among the temporal sequence. However, a major challenge with RNNs is the problem of vanishing gradients in the backpropagation, which limits the performance of RNNs for large numbers of time-steps. Several variants of RNNs have been proposed to mitigate this problem and enable the processing of longer sequences. LSTM and GRU networks are primary variants in this respect [61]. These networks incorporate gates in the RNN cells to regulate the flow of information. These gates learn which information should be kept or discarded in a time series, allowing the relevant information to be preserved throughout a long sequence of RNN cells. The following subsections describe LSTM and GRU networks in more detail.

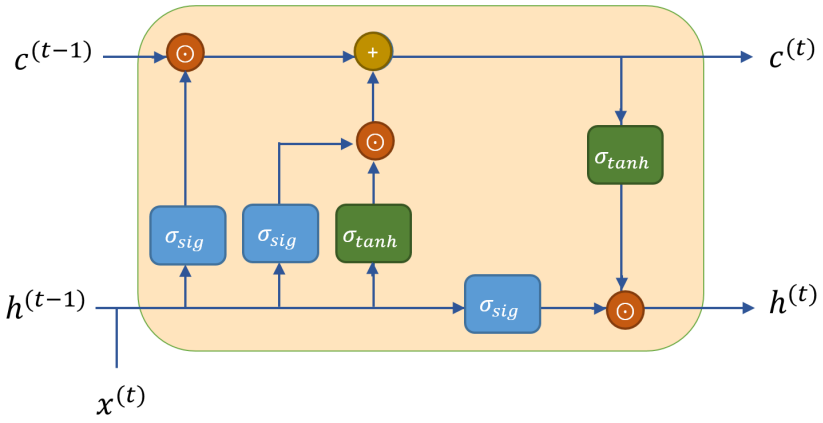


Figure 3.2: Schematic of the process flow in an LSTM cell.

3.3.2 Long-Short Term Memory Networks

LSTM networks incorporate three supplementary gates in addition to a cell state to facilitate handling large sequences: update gate Γ_u, forget gate Γ_f, and output gate Γ_o. The cell state acts as an additional supportive memory that keeps track of the relevant information throughout the process flow.

The forget gate determines what information should be discarded and what information must be retained from the previous cell state. It jointly processes the previous hidden state as well as the current input and passes the resulting feature through the sigmoid function resulting in the values in the interval (0,1). The closer this value is to 0, the more the corresponding information is forgotten, and vice versa. Similarly, the update and output gates learn, respectively, what values in the cell state should be updated and what information in the current cell state should be kept as the output of the hidden state. The equations describing an LSTM cell in the time-step t are as follows:

$$
\begin{aligned}
\tilde{\mathbf{c}}^{(t)} &= \sigma_{\text{tanh}}\left(W_{ch}\mathbf{h}^{(t-1)} + W_{cx}\mathbf{x}^{(t)} + \mathbf{b_c}\right) \\
\mathbf{\Gamma_u} &= \sigma_{\text{sig}}\left(W_{uh}\mathbf{h}^{(t-1)} + W_{ux}\mathbf{x}^{(t)} + \mathbf{b_u}\right) \\
\mathbf{\Gamma_f} &= \sigma_{\text{sig}}\left(W_{fh}\mathbf{h}^{(t-1)} + W_{fx}\mathbf{x}^{(t)} + \mathbf{b_f}\right) \\
\mathbf{\Gamma_o} &= \sigma_{\text{sig}}\left(W_{oh}\mathbf{h}^{(t-1)} + W_{ox}\mathbf{x}^{(t)} + \mathbf{b_o}\right) \\
\mathbf{c}^{(t)} &= \mathbf{\Gamma_u} \odot \tilde{\mathbf{c}}^{(t)} + \mathbf{\Gamma_f} \odot \mathbf{c}^{(t-1)} \\
\mathbf{h}^{(t)} &= \mathbf{\Gamma_o} \odot \sigma_{\text{tanh}}(\mathbf{c}^{(t)})
\end{aligned}
\tag{3.3.2}
$$

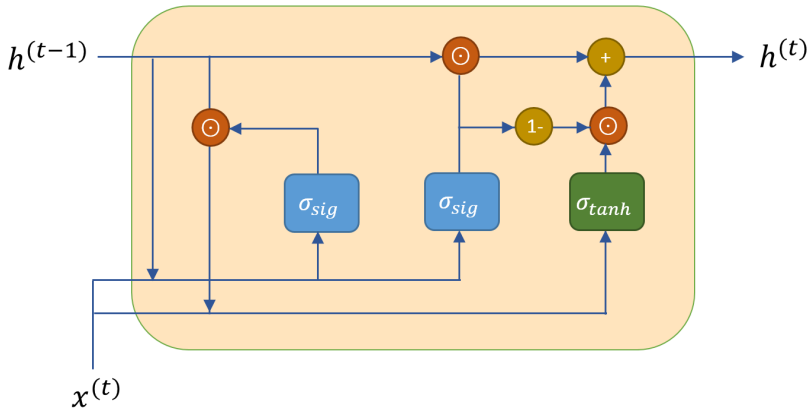


Figure 3.3: Schematic of the process flow in an GRU cell.

where $\mathbf{x}^{(t)} \in \mathbb{R}^{n_i}$ is input, $\mathbf{c}^{(t)} \in \mathbb{R}^{n_c}$ is cell state, $\mathbf{h}^{(t)} \in \mathbb{R}^{n_h}$ is hidden state, $W_{ch}, W_{uh}, W_{fh}, W_{oh} \in \mathbb{R}^{n_h \times n_h}$ and $W_{cx}, W_{ux}, W_{fx}, W_{ox} \in \mathbb{R}^{n_h \times n_i}$ are weight matrices, $\mathbf{b_c}, \mathbf{b_u}, \mathbf{b_f}, \mathbf{b_o} \in \mathbb{R}^{n_h}$ are biases, and σ_{sig} and σ_{tanh} are the sigmoid and tanh activations, respectively. $\odot$ is the Hadamard product. Schematic of an LSTM cell is depicted in Fig. 3.3.

3.3.3 Gated Recurrent Unit Networks

The GRU networks serve the same purpose as LSTM networks. In GRU units, there exist two gates, compared to three in LSTM, and the cell state equals the hidden state. Relevance (or reset) gate Γ_r and update Γ_u gate are the two gates working almost similar to forget and update gates in LSTM. The mathematical formulation of a GRU unit is as follows:

$$
\begin{aligned}
\tilde{\mathbf{c}}^{(t)} &= \sigma_{tanh}\left(W_{cc}(\mathbf{\Gamma_r} \odot \mathbf{c}^{(t-1)}) + W_{cx}\mathbf{x}^{(t)} + \mathbf{b_c}\right) \\
\mathbf{\Gamma}_u &= \sigma_{sig}\left(W_{uc}\mathbf{c}^{(t-1)} + W_{ux}\mathbf{x}^{(t)} + \mathbf{b_u}\right) \\
\mathbf{\Gamma}_r &= \sigma_{sig}\left(W_{rc}\mathbf{c}^{(t-1)} + W_{rx}\mathbf{x}^{(t)} + \mathbf{b_r}\right) \\
\mathbf{c}^{(t)} &= \mathbf{\Gamma_u} \odot \tilde{\mathbf{c}}^{(t)} + (1 - \mathbf{\Gamma_u}) \odot \mathbf{c}^{(t-1)} \\
\mathbf{h}^{(t)} &= \mathbf{c}^{(t)},
\end{aligned}
\tag{3.3.3}
$$

where $\mathbf{x}^{(t)}$, $\mathbf{c}^{(t)}$ and $\mathbf{h}^{(t)}$ have the same definition as in LSTM; $W_{cc}, W_{uc}, W_{rc} \in \mathbb{R}^{n_h \times n_h}$,

$W_{cx}, W_{ux}, W_{rx} \in \mathbb{R}^{n_h \times n_i}$, and $\mathbf{b_c}, \mathbf{b_u}, \mathbf{b_r} \in \mathbb{R}^{n_h}$. The schematic of a GRU cell is depicted in Fig. 3.3.

GRUs have been shown to be able to achieve comparable performance as LSTM in several applications, such as speech recognition, traffic load prediction, etc., with lower computational complexity [62]–[64].

3.3.4 Different topologies

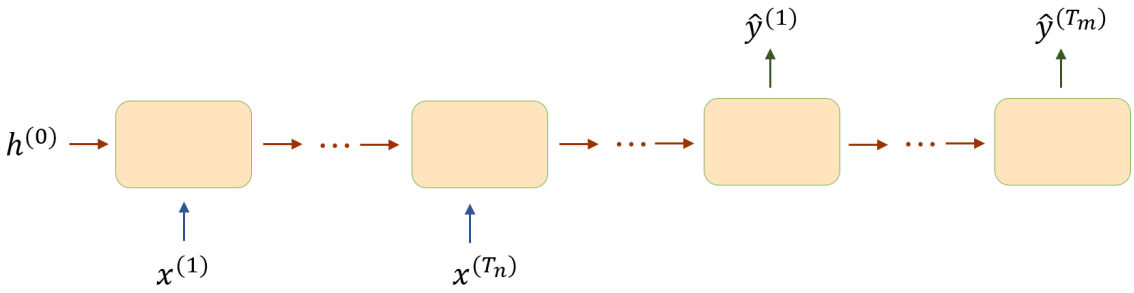
Based on the arrangement of the cells, recurrent structures can form different typologies. The common topologies for RNNs includes many-to-many (many inputs, many outputs), many-to-one (many inputs, one output), and one-to-many (one-input, many outputs) structures. Fig. 3.4, shows an schematic illustration of these topologies.

Depending on the application, each of these networks may achieve substantially different levels of efficiency in terms of performance and complexity. This is mostly due to the proportionality of the consequent number of generated time-steps, the type of extracted features, and the dimensionality of the resulting feature space to the actual unknown required complexity and feature space size for the domain of the learning task. The many-to-many topology in Fig. 1 has been demonstrated to outperform other forms of RNN in machine translation, and auto-encoder training [65], while the one-to-many design is suited for image captioning and music generation [66].

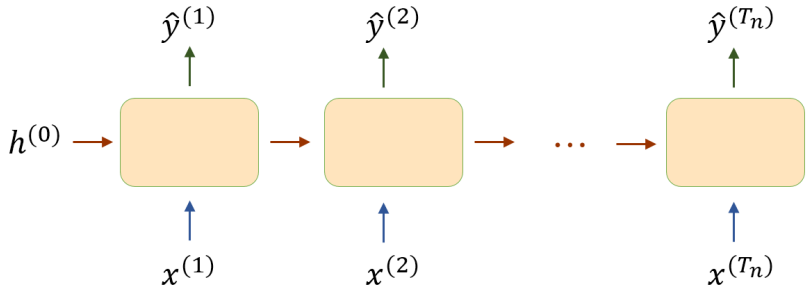
In the context of RNN-based equalization in fiber-optic communications, the many-to-many architecture in Fig. 3.4b has been adopted for equalization in the recent literature. We, however, in the following chapters, challenge the use of this topology and debate that via a specific scheme, a many-to-one topology, shown in Fig. 3.4d, can lead to the same performance but with substantially lower complexity.

3.4 Dimensionality reduction and latent space

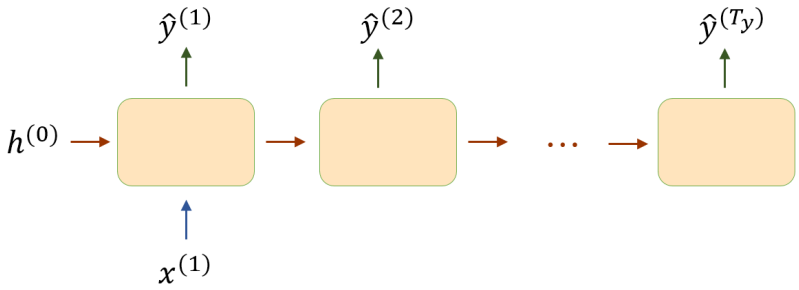
One of the obstacles that machine models deal with is the problem of dimensionality of the feature space (input data). In the context of neural network-based equalization, the concept of dimensionality is analogous to the memory that the neural network should cover times the number of samples per symbol (SPS). As the channel memory



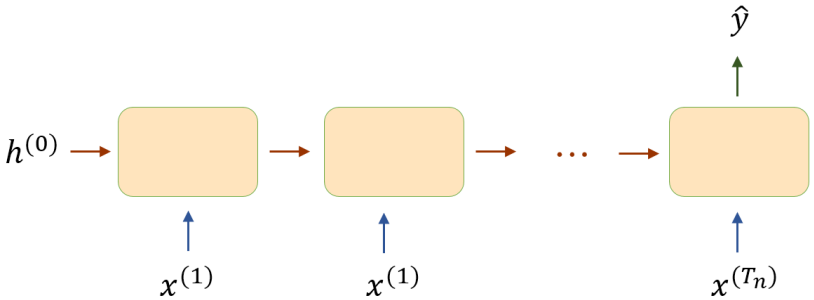
(a) Many-to-many (type 1)



(b) Many-to-many (type 2)

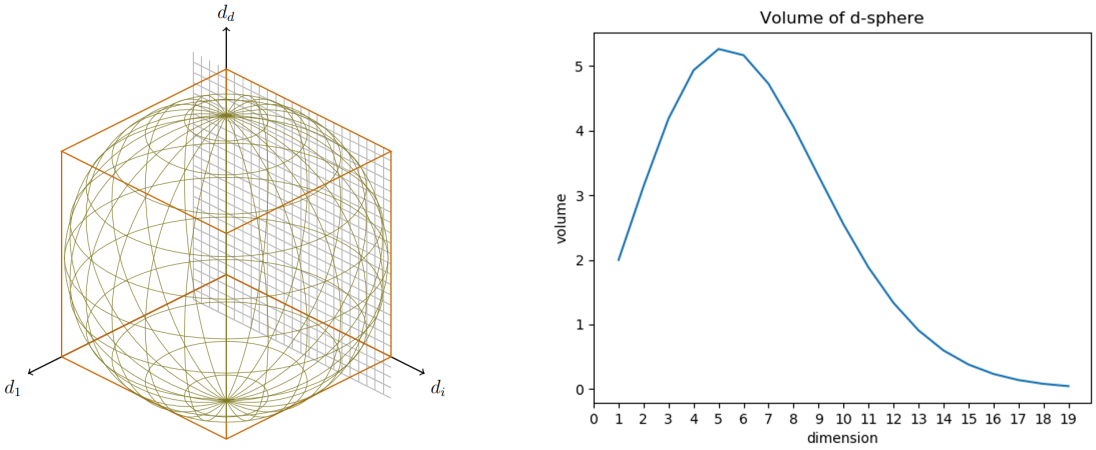


(c) One-to-many



(d) Many-to-one

Figure 3.4: Common different topologies for recurrent networks.



(a) (b)

Figure 3.5: a) Hypothetical schematic of d-dimensional unit sphere inscribed in a d-dimensional unit cube. b) Volume of the d-dimensional unit sphere as the dimension grows.

increase, the neural network faces a harder problem, and its performance is hindered more by bias and even variance. This issue is also applicable to other machine learning applications, and it is termed the curse of dimensionality.

The curse of dimensionality, in general, refers to the notion that problems can become highly more difficult to solve in high-dimensional spaces. This term has a variety of interpretations in different contexts. One interpretation refers to the growth of the complexity order of the methods as dimensionality grows. As an instance, minimum spanning tree can be solved by [67] in nearly linear time for n points when the dimensionality $d = 2$; but as d grows, the order goes to $O(n^{2-a(d)}(\log n)^{1-a(d)})$, where $a(d) = 2^{-(d+1)}$.

Another interpretation of the curse of dimensionality, which has a more emphasis in machine learning, refers to the exponential sparsity of samples in high dimensions. To provide an intuitive demonstration of this, consider the d-dimensional unit sphere, $\mathcal{S}$, inscribed in a d-dimensional unit cube, $\mathcal{Q}$. The volume of $\mathcal{S}$ is given by

$$vol(\mathcal{S}) = \frac{\pi^{d/2}}{\Gamma(\frac{d}{2}+1)} R^d, \tag{3.4.1}$$

where $\Gamma(.)$ is the gamma function. For any positive integer n, $\Gamma(n) = (n-1)!$. Ac-

cording to to the Stirling's approximation

$$n! \approx \sqrt{2\pi n}(\frac{n}{e})^n. \tag{3.4.2}$$

Thus it can easily be seen $\Gamma(\frac{d}{2}+1)$ grows highly faster than $\pi^{d/2}$, hence when $R=1$,

$$vol(\mathcal{S}) \to 0 \ \text{ as } \ d \to \infty.$$

In other words, the volume of the d-dimensional sphere with radius 1 quickly decreases to 0 as the dimension d moves to infinity. That is to say, a unit sphere in high dimensions has roughly no volume, while the unit cube's volume is always 1. Hence it is highly probable that there wasn't any close sample in the training dataset for a new unseen sample to the neural network.

In order to mitigate the challenges that arise from the curse of dimensionality, the concept of dimensionality reduction arose.

Dimensionality reduction is the process of transforming data from a high-dimensional space to a low-dimensional space such that the variance and the distance between the points are retained. Formally speaking, given the set of points $x_1, x_2, \ldots, x_n \in \mathbb{C}^d$, dimensionality reduction is describing $\mathfrak{P}$ in a lower dimension $k << d$, via a mapping f , such that $\forall x_i, x_j \in \mathfrak{P}$

$$|f(x_i) - f(x_j)|_2 \approx |x_i - x_j|_2$$

Interestingly, according to Johnson-Lindenstrauss Lemma [68], it is proved that for any $0 < \epsilon < 1$ and set of points $x_1, x_2, \ldots, x_n \in \mathbb{R}^d$, there exists a mapping $f : \mathbb{R}^d \to \mathbb{R}^k$, with $k = \Omega(\frac{logn}{\epsilon^2})$, such that

$$(1-\epsilon)|x_i - x_j|_2 \leq |f(x_i) - f(x_j)|_2 \leq (1+\epsilon)|x_i - x_j|_2. \tag{3.4.3}$$

Principal component analysis (PCA) [69] and autoencoders are effective dimensionality reduction techniques that have been established in a variety of fields of research. PCA is a linear approach that works based on changing the basis to the principal components of data and then ignoring the low-energy coordinates. The principal component of a collection of points in d dimensional space is a collection of d orthogonal vectors that

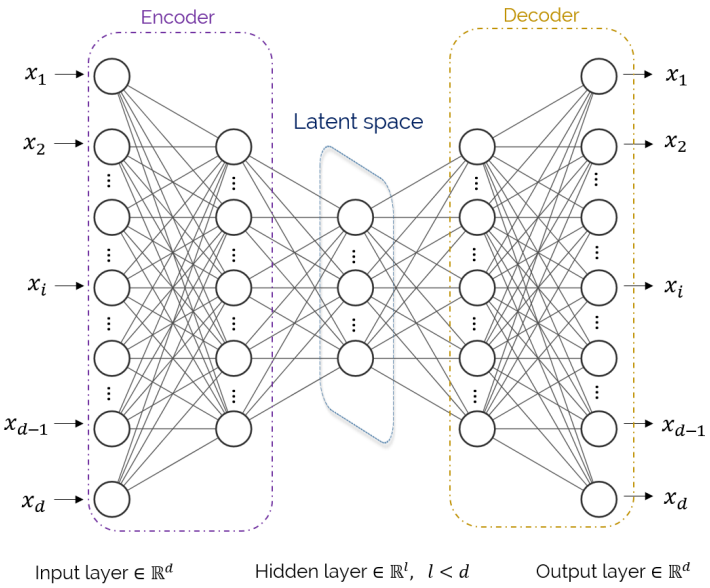


Figure 3.6: An ideal MLP autoencoder that leans how to encode an input into a lower-dimensional space, and the corresponding decoding function to retrieve the encoded input to the original value

best fits the data among the other orthogonal basis of the space. It turns out that the principal components are eigenvectors of the data's covariance matrix and they can be obtained by eigendecomposition or singular value decomposition of the data's covariance matrix [5].

On the other hand, the autoencoder is a nonlinear approach utilizing a neural network to learn a nonlinear function mapping the data to a lower-dimensional space (encoding) and an inverse function from the encoded data to the original representation (decoding). Fig. 3.6 shows a schematic of an ideal MLP autoencoder, which learns the function $f : \mathbb{R}^d \rightarrow \mathbb{R}^l$ $(l < d)$ and its inverse. The lower-dimensional space that the data is mapped to is called a latent space or embedding space [6].

We use this concept in our proposed approach for complexity reduction in bi-RNN-based equalization. In the following chapter, we explain how we use a convolutional

[5]Both the methods output the same result because of the symmetric positive semidefinite property of the covariance matrix

[6]In more accurate terms, a latent space is an embedding of a collection of elements inside a manifold in which the elements that are more similar to each other are positioned closer together. Thus, a latent space, theoretically, does not necessarily need to be in a lower dimension, although in most cases, it is.

encoder not only to reduce the dimensionality of the sampled waveform at the receiver toward 1 SpS, but also to enable the deployment of a unidirectional many-to-one vanilla RNN in place of bidirectional many-to-many recurrent structures.

3.5 Conclusion

In this chapter, we reviewed the fundamental neural network models and their underlying structures, including MLPs, CNNs, RNNs, LSTMs, GRUs, and different recurrent typologies. We also discussed the universality approximation theorem and debated that any continuous function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ can be approximated by an MLP with one hidden layer made up of a finite number of hidden units with a discriminatory activation function.

We also had a brief overview of the concept of dimensionality reduction and the application of neural network encoders in this field; which, as will be discussed in the next chapter, we exploit in devising a hybrid neural structure providing a higher computational efficiency compared to the state-of-the-art models in nonlinearity mitigation in optical fiber communications.

Chapter 4

Neural network-based equalization in optical fiber communication

As discussed in the previous chapters, in dual-polarization long-haul optical fiber communication systems, signal propagation is affected by chromatic dispersion (CD), polarization-mode dispersion (PMD), Kerr nonlinearity, and amplified spontaneous emission (ASE) noise [4]–[6]. In consequence, a cascade of digital signal processing algorithms is required to be performed at the receiver to reverse the deterministic channel effects; a process termed equalization. In this domain, CD and PMD can be compensated via linear equalization in a low-complexity regime; however, on the other hand, nonlinear distortions require nonlinear complex equalization solutions [47], [70]–[72].

A popular equalization method in optical fiber communication is digital backpropagation (DBP) [7]. DBP reverses the deterministic effects of the channel by propagating the signal backward in distance in the fiber modeled by the nonlinear Schrodinger (NLS) equation, using the split-step Fourier method (SSFM) [7], [8].

DBP, however, suffers from high computational complexity associated with a large number of spatial segments and processing bandwidth [73]. The alternative solutions such as Volterra series transfer function [9], [10], maximum-likelihood sequence (MLS) detection [11], [12], and optical phase conjugation [13]–[15] are also complex, hindering their real-time functioning in practice [74].

In recent years, data-driven solutions, prompted by neural networks, have shown potential to be satisfactory substitutes for the conventional algorithms [21]–[25], [27],

[29], [75]. The goal of these methods is to attain the performance of the conventional solutions with lower complexity and better generalizability over channel parameters. These solutions leverage neural networks made up of concatenated linear and nonlinear operators that can be optimized via the backpropagation algorithm to approximate a target function.

In this chapter, we review the recent state-of-the-art neural network-based equalization approaches in the literature, in particular model-agnostic methods, and analyze their sources of complexity. Subsequently, having the made analysis in hand, to avoid the investigated sources of complexity, we propose a convolutional recurrent neural network (CRNN) architecture equalizer model, comprising a CNN-based encoder and a unidirectional many-to-one vanilla RNN. In Chapter 5, we demonstrated that this proposed model achieves the same performance as the state-of-the-art bidirectional recurrent-based equalizers while having substantially lower complexity.

4.1 Neural network-based equalizers in the recent literature

According to the underlying rationale, the neural network-based equalization methods can be classified into two categories: model-driven approaches obtained via deep unfolding [76] of the fiber-optic channel model, and model-agnostic methods based on generic neural networks.

In model-driven methods, the neural network is constructed similar to the computation graph of the channel model, and the model parameters are then tuned using variants of the stochastic gradient descent (SGD), a process termed deep unfolding.

Learned DBP (LDBP) proposed by Häger et al. [77]–[79] is the main approach in this category. This approach uses the computation graph generated by SSFM as a blueprint for the neural network design, resulting in a convolutional neural network (CNN) with a trainable activation function. Fig. 4.1 depicts the schematic of this approach. The top processing branch in this figure corresponds to the computational graph generated by SSFM. The bottom processing branch is the computational graph of the proposed LDBP approach mimicking the computation graph of the SSFM with 2 step/span (StPS). Each network layer, i, is made up of two weight matrices $W_1^{(i)}, W_2^{(i)} \in \mathbb{C}^n$, where

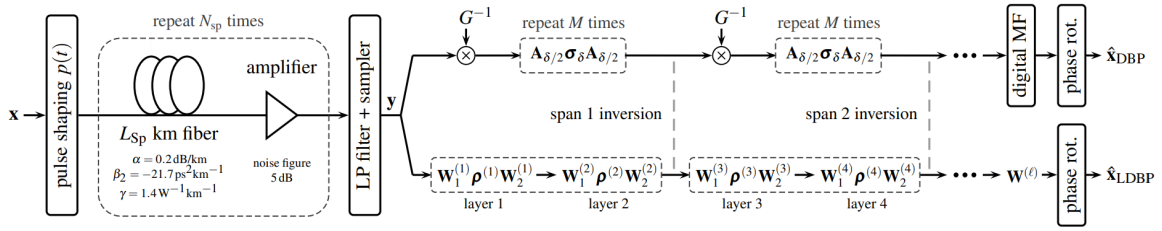


Figure 4.1: Block diagram in [77] illustrating the underlying rationale of the LDBP approach. The top processing branch at the receiver side corresponds to DBP with the computational graph generated by SSFM. The bottom processing branch presents the computational graph of the LDBP imitating the top processing branch.

n in the dimensionality of the network input $\mathbf{y} \in \mathbb{C}^n$ and the activation $\rho^{(i)} : \mathbb{C}^n \to \mathbb{C}^n$ where $\rho^{(i)}(x) = xe^{-j\alpha_i|x|^2}$ in which α_i is a trainable parameter. The weight matrices, however, are restricted to an equivalent circular convolution with a symmetric filter of length $2K_{\mathtt{m}} + 1$ ($K_{\mathtt{m}}$ is a hyper-parameter), i.e. the matrix rows are circularly shifted versions of $(u_{K_{\mathtt{m}}}, ..., u_1, u_0, u_1, ..., u_{K_{\mathtt{m}}}, 0, ..., 0)$, where $h_i \in \mathbb{C}$ and $u_{-i} = u_i$. The weight matrices are initialized by using the appropriately zeroed versions of $A_{\delta/2}$, multiplied by $e^{-\frac{\alpha}{2}L_{sp}}$ (α is the fiber loss and L_{sp} is the span length) for $W_1^{(i)}$ in the first layer of each span. Here, $A_\delta = \mathbf{F}^\dagger \text{diag}(H_1, \ldots, H_n)\mathbf{F}$ in which $\mathbf{F}$ and $\mathbf{F}^\dagger$ are respectively discrete Fourier transform (DFT) and inverse DFT (IDFT) matrices, and $H_i = \exp\left(-\frac{\alpha}{2}\delta + j\frac{1}{2}\beta_2\omega_i^2\delta\right)$ where ω_i is i-th DFT angular frequency, and δ is the step length in SSFM. There is also a final layer $W^{(\ell)} \in \mathbb{C}^{n\times m}$ accounting for the matched filtering layer, where m in the dimensionality of the input signal $\mathbf{x} \in \mathbb{C}^m$.

That being said, despite the fact that LDBP results in an excellent performance and substantially reduces complexity compared to DBP, it still results in high complexity. The computation graph of LDBP with N_{sp} spans and N_{stps} StPS is a CNN with $\ell = 3 \times N_{sp} \times N_{stps}$ successive linear and non-linear layers. For $N_{sp} = 32$ and $N_{stps} = 3$ considered in [77], [78], this results in $\ell = 288$ layers, which could be intolerably high. This issue is also pertinent to other similar approaches in this category, such as in [80] by Sidelnikov et al., where a roughly similar approach to LDBP is adopted to simulate DBP in dual-polarization wavelength-division-multiplexing (WDM) systems by customizing the nonlinear activation function to account for a different number of neighboring symbols from adjacent spectral channels; or as in [81] where Jiang et al. propose a so-called physics-informed MLP which is capable of solving the NLS equation in fiber-optic communications with lower complexity than SSFM.

Model-agnostic neural network equalizers, on the other hand, aim to learn the distortion patterns using a low-complexity shallow neural network that does not incorporate the channel model into the neural network design. Deep learning is performed here by pairing the sampled waveform at RX (or, typically, the waveform after CD compensation or linear equalization) with the corresponding transmitted symbols at the transmitter (TX).

Several model-agnostic neural networks have been proposed for mitigating nonlinear distortions in optical fiber transmission [21], [22], [24], [34].

As universal approximators, MLPs are among the first adopted models in this field. A number of MLP-based adopted models for equalization are discussed in [81], [82], [84], [85]. As an instance, Fig. 4.2 shows a schematic of the proposed MLP model by Catanese et al. [82]. This model, from the pre-processed sampled waveform, takes the target symbol at time-step i together with a window of its M neighbouring symbols (effective channel memory) at x- and $y-$ polarization, and it outputs the real and imaginary part of the equalized symbol. This input-output model, which is referred to as windowing, is common among the model-agnostic methods. As MLPs, however, attempt to capture the correlation among each pair of the samples in data at each layer using fully-connected layers, they are prone to over-fitting due to the large number of trainable parameters, and they could also incur a large number of floating-point operations (FLOPs).

Motivated by this problem, hybrid of the convolutional neural network and MLP (CNN+MLP) -based models have been investigated [83], [86], [87]. Using convolution layers, these models attempt to capture short-term dependencies among neighboring symbols. Following the convolution layers, they leverage a series of fully-connected layers to capture long-term dependencies. Fig. 4.3 shows a CNN+MLP model proposed by Chuang et al. [83] with 5 hidden layers, outperforming the Volterra nonlinear equalizers in a 128 Gbps PAM-4 optical transmission over 40 km standard SMF. That said, owing to the presence of fully connected layers introducing a large number of trainable parameters and floating-point operations (FLOPs), making the model prone to overfitting, CNN+MLPs might lead to inefficiency as well. With respect to this matter, efforts to design models with a minimal number of dense layers have recently grown.

Bidirectional recurrent neural network (bi-RNN) -based equalizers have recently piqued

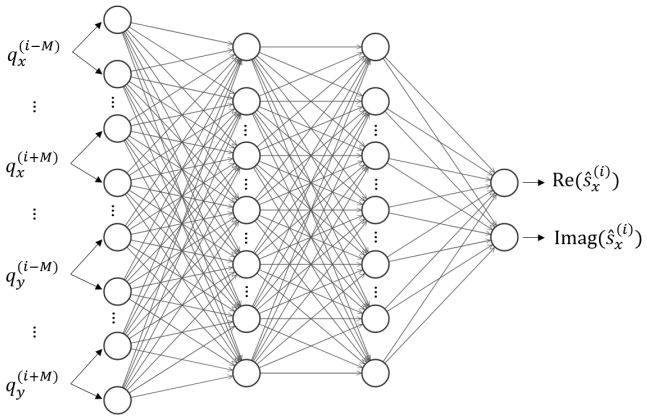


Figure 4.2: Schematic of the proposed model by Catanese et al. [82]. The neural network receives the window of samples $q_x(i-M,\ldots,i+M)$ and $q_y(i-M,\ldots,i+M)$, passes it from two fully connected hidden layers, and outputs the real and imaginary value of the equalized symbol at time-step i in x- polarization (similarly also for y-polarization with another network).

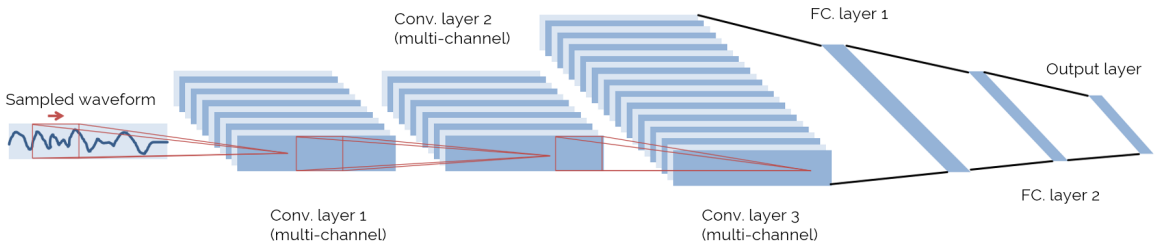


Figure 4.3: CNN+MLP model by Chuang et al. [83]. In this model, the window of samples is passed to the neural network, and the output layer indicates the equalized symbol.

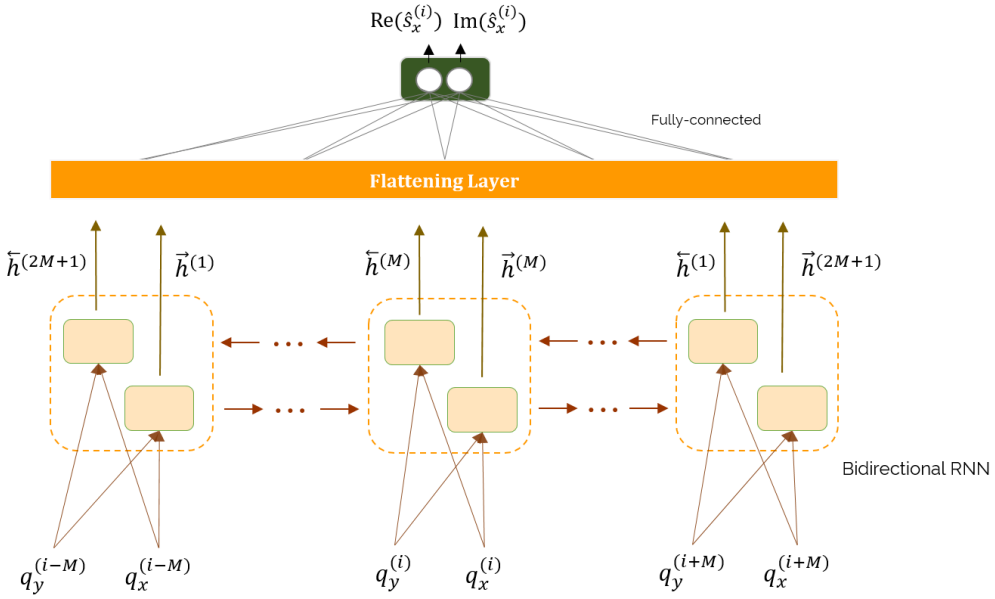


Figure 4.4: General structure of bi-RNN based equalizers, where the bi-directional layer receives the window of samples $q_x(i-M,\ldots,i+M)$ and $q_y(i-M,\ldots,i+M)$ and processes them. The output of the RNN cells then are flattened and, via a linear fully-connected layer, are processed. The network ultimately outputs the real and imaginary value of the equalized symbol at time-step i.

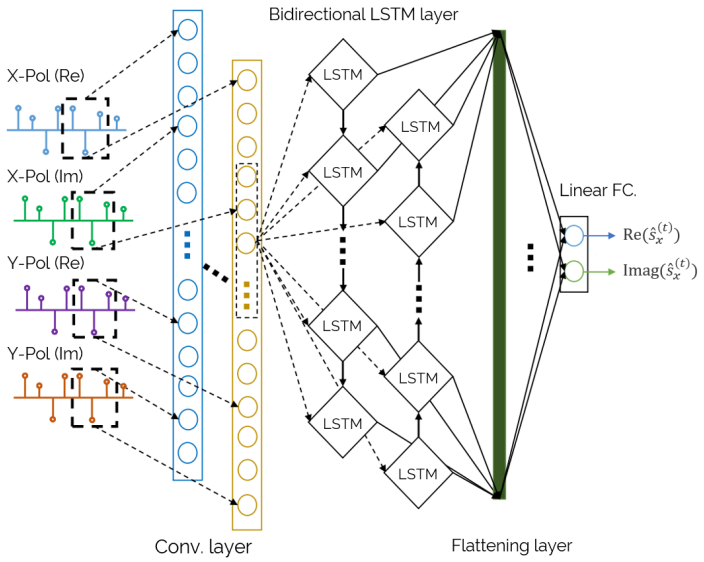


Figure 4.5: CNN+bi-LSTM model by Freire et al. [24], [34] where a convolution layer pre-processes the signal before the bi-LSTM layer.

attention in this regard [23], [32]–[35], [88]–[90]. RNNs are efficient in capturing long-term dependencies owing to their ability to keep track of effective information over a large sequence, resulting in the potential to span a wide channel memory for equalization.

It is shown by Liu et al. [33] that bi-directional gated recurrent units (bi-GRU) could achieve the same bit error rate (BER) in nonlinearity mitigation as bi-directional long-short-term memory (LSTM) models [35] with a lower number of FLOPs. Furthermore, Deligiannidis et al. [32] shows that a vanilla bi-RNN could reach the performance of bi-LSTM and bi-GRU models in some setups, but with a higher number of hidden units in the recurrent cells. The general structure of the bi-RNN -based approaches is depicted in Fig. 4.4, where using a bidirectional recurrent layer, a serializer, and a linear fully-connected layer, the model attempt to capture dependencies within a signal. A study by Freire et al. [24], [34] also shows that having a convolution layer prior to a bi-LSTM model could lead to superior performance compared to MLP, CNN+MLP, and bi-LSTM models in a 34.4 GBd optical transmission over 9×50 km TWC fibers. Fig. 4.5 shows the schematic of this model.

In this doctoral thesis, however, we note that bi-RNN-based equalizers could still be far from being implemented in realistic systems due to their high computing resource requirements.

We note three issues concerning the computational complexity of these neural network equalizers. First, a bi-RNN is not an optimal tool for capturing short-term dependencies. Second, in the gated RNNs, particularly LSTM and GRU structures, although the supplementary gates strengthen the neural network in the training mode, their complexity and overhead stay with the model in the inference mode as well. Third, having one RNN cell per time-step in the signal, as in [24], [32], [34], [35], results in a high number of RNN cells, which brings about a hefty computational load. This complexity is doubled furthermore by the bidirectional processing flow. Besides, in the case of joint nonlinearity-PMD compensation, the computational burden on the recurrent layer increases by the growth in the dimensionality of the signal.

To address these concerns, in this doctoral thesis, we propose a hybrid architecture comprising a CNN-based encoder and a unidirectional many-to-one simple (vanilla) RNN working in tandem. The CNN component captures short-range dependencies using convolutional kernels. Furthermore, in parallel, by taking the data into a latent

space using unbalanced multi-channeling and striding, it reduces the dimensionality of the signal towards 1 SpS. The structure also groups the neighbour samples whose dependencies have been captured into distinct blocks to be passed as the time-steps to the RNN. The RNN, then, detects the long-range dependencies within the reduced number of input dependencies to the layer. As the number of time-steps is substantially shrunk and the short-term dependencies have been extracted, a unidirectional many-to-one vanilla RNN is sufficient for this task. In the following section, we elaborate on this model in more depth.

4.2 Complexity reduction using latent space of CNNs

There often exist long- and short-term dependencies in the received signal, depending on distance and bandwidth, that need to be extracted for equalization. We note that although bidirectional recurrent networks are appropriate tools to capture long-distance dependencies, using them to learn short-term dependencies brings about a high computational inefficiency. This inefficiency is mainly rooted in the high number of time-steps, the presence of recurrent gates, and the bidirectional processing flow. Besides, in the case of joint nonlinearity-PMD compensation, the dimensionality of the signal passed to the neural network is usually 2 SpS or higher. This oversampling imposes an additional computational burden on the recurrent layer.

To address these problems, in this study, we aim to place a pre-processing block prior to the RNN to eliminate the mentioned sources of complexity. To this end, we leverage a CNN structure because of its twofold advantages. Firstly, CNNs are established to be efficient tools to capture short-term dependencies, which minimizes the overlap between short-term and long-term dependencies extraction. Secondly, by the multi-channeling and striding properties, they can take the data into a latent space where not only the neighbour samples whose dependencies have been captured are grouped in the depth-axis of the feature-map, but also using appropriate hyper-parameters, the dimensionality of data is reduced towards 1 SpS (the latter is only applicable to joint PMD-nonlinearity compensation).

This enables the RNN to take the vectors in the depth-axis of the feature-map provided by the CNN encoder, as the time-steps. This highly limits the number of time-steps

in the RNN, facilitating the replacement of gated recurrent cells with vanilla RNN cells. In addition, there is no need for a high number of hidden units in the RNN cells since there is no need for further processing to capture dependencies within the input vector as they have already been extracted. These together enable us to replace the bidirectional recurrent layer with a unidirectional many-to-one vanilla RNN layer. This is owing to the fact that the effective information (RNN memory) can be maintained along the layer while it is updated at each time-step in a cascade of quite limited number of RNN cells (3 to 6). Note that in the absence of a CNN block in the model, besides the burden of capturing both types of dependencies on the uni-directional RNN layer, the RNN memory should have been maintained throughout the long cascade of RNN cells, making the neural network inefficient for real-time operation.

In other words:

The first motivation underlying our approach is to minimize the overlap between capturing short-range and long-range nonlinear impairment patterns in the received waveform, in order to improve efficiency. With respect to this, we aim to leverage CNNs for capturing short-temporal dependencies, which are established tools for this purpose, and utilize recurrent structures exclusively for capturing long-range dependencies so that fewer number of units is used in each recurrent cell.

Considering that bi-directional RNN-based methods have relatively high complexity in general, the second motivation of our approach is to use a unidirectional structure for the recurrent layer. This necessitates a substantial reduction in the number of input time-steps to the recurrent layer so that the memory of the RNN can maintain throughout the pipeline. This requires the neighbour samples whose dependencies are captured are ground into individual blocks as time-steps for the recurrent layer.

The computational burden on the recurrent layer goes high as the signal dimensionality increases. The third stimulus of our approach is to reduce the dimensionality of the signal towards 1 SpS prior to the recurrent layer by a dimensionality reduction process. This also contributes to the requirement of fewer units in the recurrent cells.

The final ambition in our approach is to enable leveraging vanilla RNNs instead of LSTM or GRU units in the recurrent layer, given that, as debated, the overhead of the additional gates in the gated recurrent structures stay with the model in the inference mode, adding up to the model complexity. Having a limited number of time-steps and hidden units in the recurrent cells could facilitate this idea.

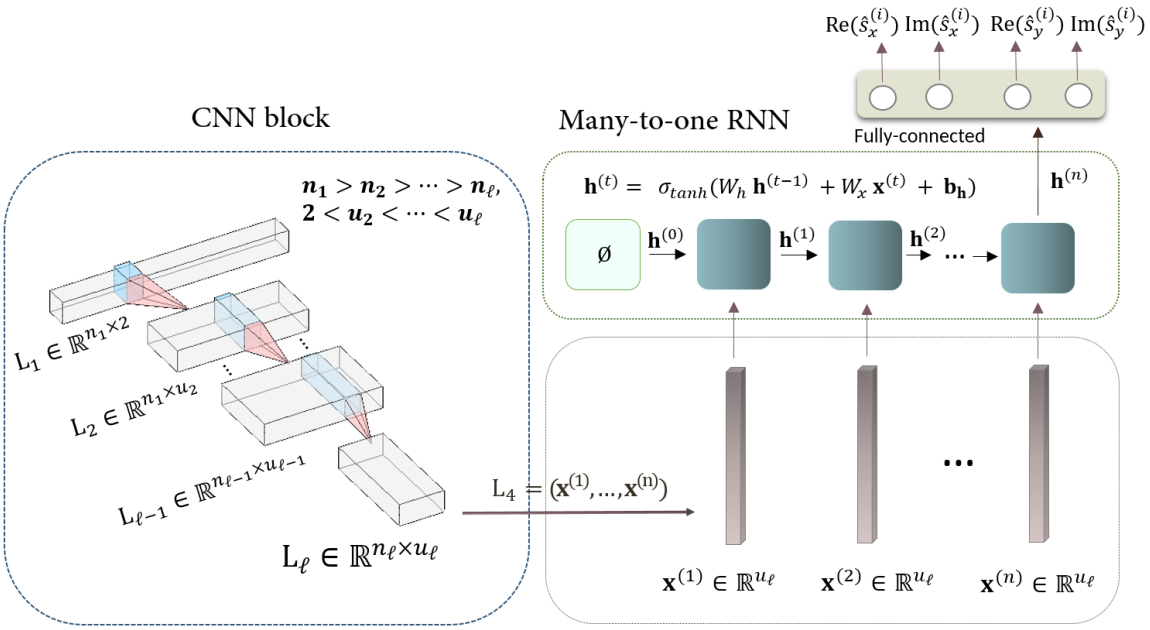


Figure 4.6: Schematic illustration of the process flow in the proposed CRNN model where a CNN-based encoder and a unidirectional many-to-one vanilla RNN work in tandem, each best capturing one set of channel impairments while compensating for the shortcomings of the other. The CNN block captures the short-range dependencies using multi-channel strided layers. The width of the feature-map gets narrower at each layer while its depth increases. The vectors in the depth-axis of the output feature-map by the CNN block are considered as the time-step for the unidirectional RNN layer.

To fulfill these objectives, we exploited a CNN-based encoder composed of multi-channel strided convolution layers prior to the recurrent layer. Multi-channeling enables the possibility of applying an ample number of filters to signal to extract the short-term dependencies. Moreover, combined with striding, it grants us a tool for an efficient dimensionality reduction via applying a disproportionate number of filters in each layer compared to the stride set for the layer. This dimensionality reduction process not only notably lessens the computational load on the recurrent layer but also diminishes the load on the final convolution layer, where a small-size filter can span a large area. Furthermore, at the output of the CNN block, the neighbor symbols whose dependencies have been captured are automatically grouped together as the individual vectors in the depth-axis of the output 2-D feature-map.

Based on this rationale, Fig. 4.6 illustrates the schematic of the process flow in the pro-

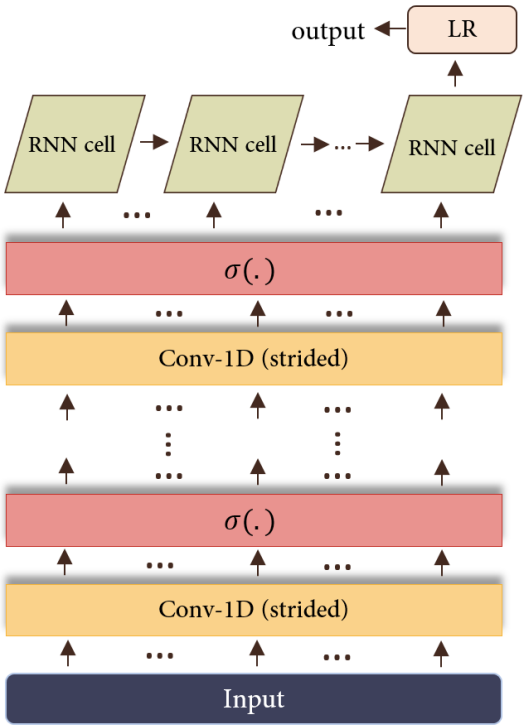


Figure 4.7: The architecture of the proposed CRNN model. The number of convolution layers is a hyper-parameter that is optimized.

posed CRNN model. This model, similar to the bidirectional recurrent-based methods, follows the windowing input-output format, and thus receives a data matrix containing windows $\mathbf{u}_x(t = i - M \times \text{SpS}, \ldots, t = i + M \times \text{SpS})$ and $\mathbf{u}_y(t = i - M \times \text{SpS}, \ldots, t = i + M \times \text{SpS})$ output by the previous block in the DSP chain for the x- and y- polarization, where M is the assumed channel memory and i is the index of the symbol that is going to be equalized [1]. This matrix is initially processed by a CNN. In the CNN, at each layer, the width of the data matrix (feature-map) is reduced (via striding), and the depth is increased (via applying filters). For joint nonlinearity-PMD compensation, narrowing the width and increasing the depth should be in a disproportionate manner to reduce the dimensionality (narrowing the width must be further). Following each layer, an activation function introduces nonlinearity to the output. The feature-map output by the CNN is then passed to the recurrent layer, where the vectors in the depth-axis are considered as the time-steps.

As via striding in the CNN block, the width of the feature-map has markedly reduced,

[1] A detailed description of the input-output model of the neural network is discussed in Sec. 5.1.3.

there exist a sharply declined number of time-steps for the RNN. Furthermore, as the dependencies within each time-step vector have already been captured, the required number of hidden units in the recurrent cells shrinks noticeably. Having a quite limited number of time-steps and a small-size hidden state, makes it possible to maintain the state throughout the layer while it is updated at each time-step, by using a unidirectional many-to-one vanilla RNN layer. The output of the RNN is consequently the hidden state output by the last recurrent cell. This output is then linked to 4 processing units via a linear fully-connected layer, which forms a linear regression model for each of the 4 units. These 4 units are supposed to output the real and imaginary parts of the symbols in the target position i, at x- and y- polarizations. Fig. 4.7 depicts the block diagram of the proposed model.

4.3 Conclusion

In this chapter, we reviewed a variety of state-of-the-art neural network-based equalizers for nonlinearity mitigation in optical fiber communications. We debated that bidirectional recurrent neural network-based methods have recently piqued big deals of attention owing to their superior performance; however, their computational complexity goes relatively high.

We analyzed the sources of computational inefficiency in these approaches and subsequently proposed a hybrid neural structure comprised of a CNN-based encoder and a unidirectional vanilla RNN working in tandem. We discussed that the CNN-based encoder not only captures the short-term dependencies and reduces the dimensionality of the signal, but it is also a key structure for enabling the practical implementation of a unidirectional RNN for capturing long-range dependencies.

In the next chapter, we demonstrate that for a 64 GBd dual-polarization 16-QAM optical transmission over 14×80 km standard SMF, the proposed CRNN model achieves a comparable performance to the bi-RNN, bi-GRU, bi-LSTM, and CNN+bi-LSTM -based equalizers adopted in [32]–[35], with greater than 50% fewer number of FLOPs.

Chapter 5

Implementation and numerical results

In this chapter, we discuss the implemented fiber-optic communication system model, the fiber parameters, system setup, and the developed software. Moreover, we discuss the applied training and evaluation methodology of the implemented neural network-based equalizers, and we elaborate on the resulting performance versus the complexity of the models.

5.1 Implemented polarization-multiplexed fiber-optic transmission system model

We modeled a dual-polarization 16-QAM 64 GBd point-to-point fiber-optic transmission system over 14×80 km SSMF optical-link, according to the system model illustrated in Fig. 5.1, with the following parameters: fiber loss $a_{\mathrm{dB}} = 0.2$ dB/km, chromatic dispersion D=17 ps/nm/km, nonlinearity parameter $\gamma = 1.4$ $\mathrm{W}^{-1}\mathrm{km}^{-1}$, PMD value 0.05 ps/$\sqrt{\mathrm{km}}$, EDFA noise figure NF = 5 dB, and the laser linewidth 100 kHz. For pulse shaping, root-raised cosine (RRC) filters with a roll-off of 0.25 are employed. Forward propagation was simulated using SSFM with 8 SpS and 80 step/span (increasing either value did not affect the results). The sampling rate at RX was set to 2 SpS.

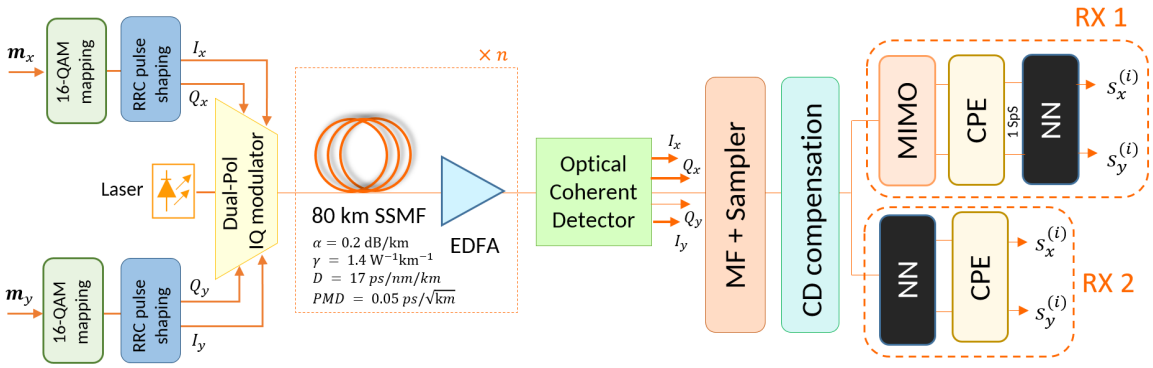


Figure 5.1: Block diagram of the considered fiber-optic transmission system. The top and bottom branches of the receiver represent the RX model 1 and 2, respectively.

The mentioned setup was adopted based on the common consensus in the community [32], [33], [77], [79], [82] on the modeling parameters, according to the current available fiber-optic technologies [91]. The length of the considered optical link also ensures a long-haul transmission and further resembles C-LION1 submarine cable [92] between Rostock, Germany and Helsinki, Finland.

The mentioned fiber optic transmission system was implemented as follows.

5.1.1 Transmitter model

At the transmitter, a bit stream $\mathbf{m_x} = (m_x^{(1)}, m_x^{(2)}, \ldots, m_x^{(N_b)})$, $m_x^{(i)} \in \{0, 1\}$, is mapped to a sequences of symbols $\mathbf{s_x} = (s_x^{(1)}, s_x^{(2)}, .., s_x^{(N_s)})$, where $s_x^{(i)}$ are drawn from a 16-QAM constellation. The sequence of symbols $\mathbf{s_x}$ is then modulated to a digital waveform $q_x(t, 0) = \sum_{i=1}^{N_s} s_x^{(i)} p(t - i/R_s)$, where $p(t)$ is the root-raised-cosine (RRC) pulse shape and R_s is the baud rate. This is similarly done for the bit stream $\mathbf{m_y}$ which is modulated to $q_y(t, 0)$. The waveforms $q_x(t, 0)$ and $q_y(t, 0)$ are then multiplexed into an electric field with a dual-polarization Mach-Zehnder in-phase (I) and quadrature (Q) modulator, that is transmitted over optical fiber. The modulator is driven by a laser with linewidth $\Delta\nu$ that introduces phase noise modeled by a Wiener process and Lorentzian power spectral density [6, Chap. 3.5]. The same realization of the phase noise is applied to both polarizations.

5.1.2 Channel model

The channel was implemented according to the explanation provided in Section 2.3.

5.1.3 Receiver model

At the receiver, an optical coherent detector translates the optical signal to four electrical signals, corresponding to the I and Q components of each polarization. It is assumed that the lasers at TX and RX operate at the same frequency, i.e., the carrier frequency offset (CFO) is zero. Consecutively, the sampled waveforms are forwarded to the digital signal processing (DSP) chain to compensate for the channel impairments.

The first step in the DSP is CD compensation, reversing the dispersion effect throughout the fiber in one-shot as

$$\hat{q}_x(\omega, \mathfrak{L}) \mapsto \exp\left(-j\frac{\beta_2}{2}\omega^2\mathfrak{L}\right)\hat{q}_x(\omega, \mathfrak{L}), \tag{5.1.1}$$

where $\mathfrak{L}$ is the fiber length. Next, depending on the task of the neural network, two receiver configurations, labeled as RX 1 and RX 2 in Fig. 5.1, are considered.

RX model 1. The neural network is placed after the linear equalization, with the aim of mitigating nonlinear channel impairments. The linear DSP consists of a cascade of CD compensation, a radius-directed-equalization (RDE) [93] -based multiple-input-multiple-output (MIMO) algorithm to compensate for the PMD, a demultiplexer to separate the two polarizations, and a two-stage carrier phase estimation (CPE) algorithm to compensate for the phase offset. The CFO compensation is not required, since the CFO is assumed to be zero.

RX model 2. The neural network is placed after the CD compensation, and its purpose is to jointly mitigate the nonlinearity and PMD. There is a CPE block after the neural network in this RX architecture to compensate for the laser phase noise, as the neural network is not able to compensate for the phase noise efficiently due to its randomness.

The input-output model of the neural networks in both RX models is shown in Fig. 5.2. To equalize the symbols $s_x^{(i)}$ and $s_y^{(i)}$ at time-step i, the neural network processes two

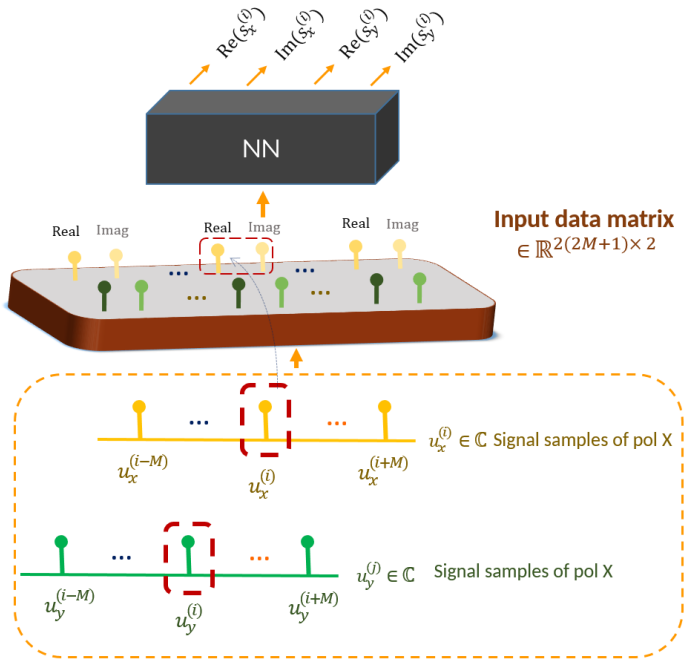


Figure 5.2: Input-output model of the neural networks in both RX models.

vectors $\mathbf{u}_x^{(i)}$ and $\mathbf{u}_y^{(i)}$ containing a window of the time-domain signal samples from each polarization output by the previous block in the DSP chain, *e.g.*,

$$\mathbf{u}_x^{(i)} = \left(u_x^{(i-M)}, \cdots, u_x^{(i)}, \cdots, u_x^{(i+M)}\right)^T. \tag{5.1.2}$$

The input vectors are centered at time-step i, and span $2M$ neighbor samples (M left, M right). If the effective channel memory in number of symbols is $\bar{M}$, then $M = \bar{M}\text{SpS} + (\text{SpS} - 1)/2$. The real and imaginary parts of the samples in $\mathbf{u}_x^{(i)}$ and $\mathbf{u}_y^{(i)}$ are then split and placed in every other position in a corresponding vector, *e.g.*,

$$\tilde{\mathbf{u}}_x^{(i)} = \left(\Re(u_x^{(i-M)}), \cdots, \Re(u_x^{(i)}), \Im(u_x^{(i)}), \cdots, \Im(u_x^{(i+M)})\right)^T. \tag{5.1.3}$$

Consequently, the vectors $\tilde{\mathbf{u}}_x^{(i)}$ and $\tilde{\mathbf{u}}_y^{(i)}$ are stacked together, and the matrix $U^{(i)} = (\tilde{\mathbf{u}}_x^{(i)}, \tilde{\mathbf{u}}_y^{(i)}) \in \mathbb{R}^{2(2M+1)\times 2}$ is passed to the neural network as input. For RX 1, the SpS at the input of the neural network is 1. After processing, the neural network outputs the real and imaginary parts of the equalized symbol in the x- and y- polarization at time-step i (4 outputs).

The neural networks are trained with $U^{(i)}$ as the input data matrix, and the corresponding correct transmitted symbols $s_x^{(i)}$ and $s_y^{(i)}$ in the standard 16-QAM constellation as

the ground truth. The performance of the receivers is measured in the test mode by the Q-factor

$$\text{Q-factor} = 20\log_{10}[\sqrt{2}\text{erfc}^{-1}(2\text{BER})], \tag{5.1.4}$$

where BER is the bit-error-rate, and erfc is the complementary error function.

5.2 Developed software

According to the modeling procedure discussed in the previous chapter, a polarization-multiplexed fiber-optic transmission system model was developed in Python 3.x, consisting of optical fiber channel modeling and DSP chain implementation at the transmitter and receiver. Certain modules of this system were developed under the GNU General Public License v3.0 (available at [94]), including QAM modulation, RRC Pulse shaping, signal propagation governed by CNLSE (realized by SSFM with distributed PMD integration), ASE noise modeling, laser phase noise modeling, sampling, matched filtering, chromatic dispersion compensation, digital back-propagation, RDE-based MIMO equalization, polarization separation and synchronization, demodulation, nearest neighbour detection, and Q-factor calculation.

5.3 Training methodology

MLP, CNN+MLP, bi-RNN, bi-GRU, bi-LSTM, and CNN+bi-LSTM neural structures with respectively similar architectures to [82], [83], [32] [33], [35], [34], as well as, the proposed CRNN model were considered for equalization, in the context of the two RX models in Fig. 5.1.

For both the RX models, the neural networks were trained with $U^{(i)}$ (discussed in Section 5.1.3) as the input data matrix, and the corresponding correct transmitted symbols $s_x^{(i)}$ and $s_y^{(i)}$ in the standard 16-QAM constellation as the ground truth.

The following subsections discuss the training procedure and hyper-parameters used for the neural networks in each of the RX models.

Table 5.1: Details of the hidden layers of the implemented neural networks for RX 1. The output layer for all the models is a linear fully-connected layer with 4 units.

Model	**Layer**	**Type**	**Details**
MLP [82]	layer 1	Dropout	dropout rate: 0.4
	layer 2,3	FC	#units: 1536, activ.: tanh
	layer 4	Dropout	dropout rate: 0.3
	layer 5,6	FC	#units: 1536, activ.: tanh
CNN+MLP [83]	layer 1	Conv-1D	L_{ker}: 49, strd: 1, #ch.: 4, activ.: relu
	layer 2	Conv-1D	L_{ker}: 49, strd: 1, #ch.: 6, activ.: relu
	layer 3	Conv-1D	L_{ker}: 49, strd: 1, #ch.: 8, activ.: relu
	layer 4,5	FC	#units: 768, activ.: tanh
bi-LSTM [35]	layer 1	bi-LSTM	#units: 144; activ.: tanh, sigmoid
	layer 2	Flattening	data format: channel last
bi-GRU [33]	layer 1	bi-GRU	#units: 144; activ.: tanh, sigmoid
	layer 2	Flattening	data format: channel last
bi-RNN [32]	layer 1	bi-RNN	#units: 240; activ.: tanh
	layer 2	Flattening	data format: channel last
CNN+bi-LSTM [34]	layer 1	Conv-1D	L_{ker}: 49, strd: 1, #ch.: 2, activ.: leaky relu
	layer 2	bi-LSTM	#units: 136, activ.: tanh, sigmoid
	layer 3	Flattening	data format: channel last
Proposed CRNN	layer 1	Conv-1D	L_{ker}: 49, strd: 1, #ch.: 4, activ: relu
	layer 2	Conv-1D	L_{ker}: 9, strd: 9, #ch.: 36, activ: relu
	layer 3	Conv 1-D	L_{ker}: 5, strd: 5, #ch.: 180, activ: relu
	layer 4	uni-RNN	#units: 454, activ.: tanh

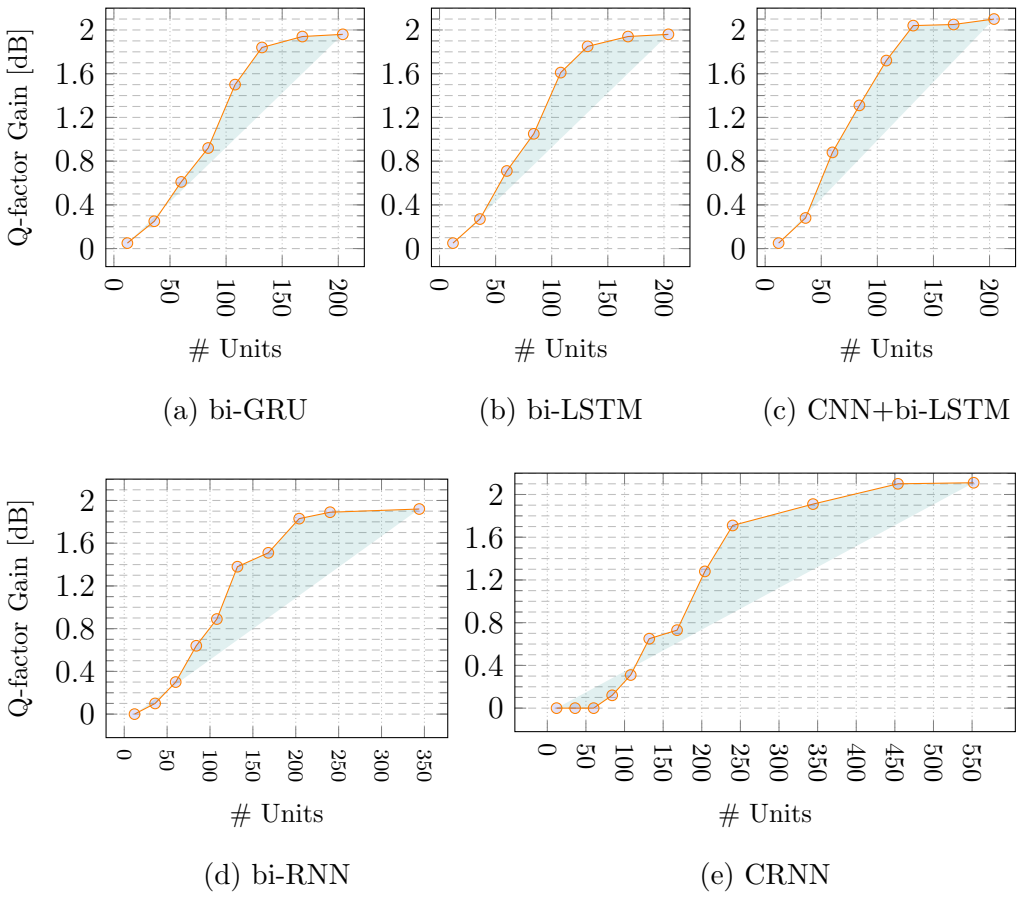


Figure 5.3: Q-factor gain over linear equalization as a function of the number of hidden units in the recurrent cells of the recurrent-based approaches for RX 1. Shaded area illuminates the nonlinear region on the #units-gain curve and the distance of the points to the linear relation.

5.3.1 RX model 1

As noted, the neural network in RX 1 is placed after CD compensation, RDE-based MIMO equalizer, demultiplexer, and CPE; and in consequence, it operates on 1 SpS signals with the input-output scheme illustrated in Fig. 5.2.

The details of the implemented neural networks for RX 1 are mentioned in Table. 5.1. The hyper-parameters in each case, especially the number of hidden units in the recurrent layers, were optimized using K-fold cross-validation, considering the performance-complexity trade-off. Fig. 5.3 illustrates the Q-factor gain over linear equalization as a function of the number of hidden units in each cell in the recurrent layers. The region

Table 5.2: Details of the hidden layers of the implemented neural networks for RX 2. The output layer for all the models is a linear fully-connected layer with 4 units.

Model	**Layer**	**Type**	**Details**
MLP [82]	layer 1	Dropout	dropout rate: 0.5
	layer 2	FC	#units: 1536, activ.: tanh
	layer 3,5	FC	#units: 1152, activ.: tanh
	layer 4	Dropout	dropout rate: 0.4
	layer 6	FC	#units: 768, activ.: tanh
CNN+MLP [83]	layer 1	Conv-1D	L_{ker}: 99, strd: 1, #ch.: 3, activ.: relu
	layer 2	Conv-1D	L_{ker}: 99, strd: 3, #ch.: 7, activ.: relu
	layer 3	Conv-1D	L_{ker}: 51, strd: 3, #ch.: 20, activ.: relu
	layer 5	Dropout	dropout rate: 0.3
	layer 4,6	FC	#units: 1152, activ.: tanh
bi-LSTM [35]	layer 1	bi-LSTM	#units: 158; activ.: tanh, sigmoid
	layer 2	Flattening	data format: channel last
bi-GRU [33]	layer 1	bi-GRU	#units: 158; activ.: tanh, sigmoid
	layer 2	Flattening	data format: channel last
bi-RNN [32]	layer 1	bi-RNN	#units: 270; activ.: tanh
	layer 2	Flattening	data format: channel last
CNN+bi-LSTM [34]	layer 1	Conv-1D	L_{ker}: 49, strd: 1, #ch.: 2, activ.: leaky relu
	layer 2	bi-LSTM	#units: 136, activ.: tanh, sigmoid
	layer 3	Flattening	data format: channel last
Proposed CRNN	layer 1	Conv-1D	L_{ker}: 99, strd: 1, #ch.: 4, activ: relu
	layer 2	Conv-1D	L_{ker}: 11, strd: 11, #ch.: 30, activ: relu
	layer 3	Conv 1-D	L_{ker}: 9, strd: 9, #ch.: 125, activ: relu
	layer 4	uni-RNN	#units: 494, activ.: tanh

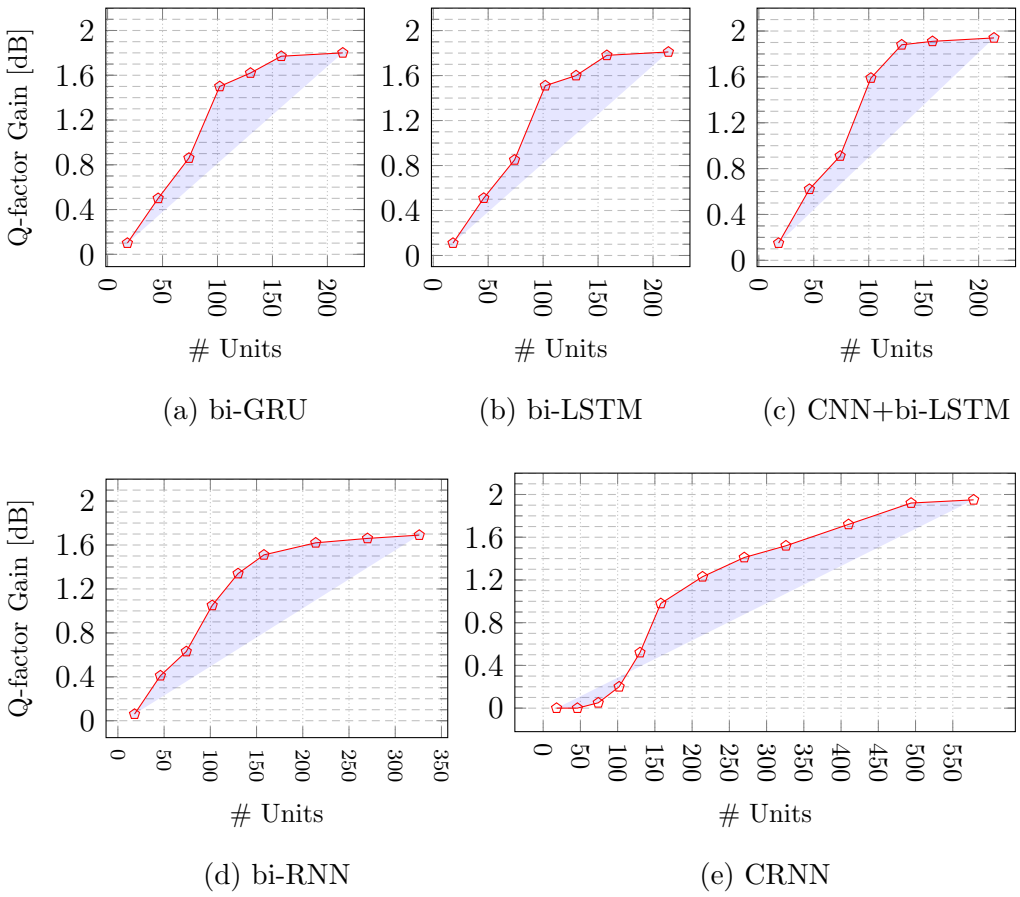


Figure 5.4: Q-factor gain over linear equalization as a function of the number of hidden units in the recurrent cells of the recurrent-based approaches for RX 2. Shaded area illuminates the nonlinear region on the #units-gain curve and the distance of the points to the linear relation.

of interest is the shaded area between the gain curve and the linear relation, where a linear increase in the number of hidden units yields a higher than linear payoff. Note that for bi-directional layers, the presented numbers in the figure and table are the sum of hidden units for both directions, half for each.

All models were trained using the Adam algorithm in Tensorflow 2.6 with mean square error (MSE) loss function, 2^{18} normalized input output training vectors (with the structure discussed in Section 5.1.3), batch size of 16, learning rate of 5×10^{-4}, and the decay rates $\beta_1^{Adam} = 8.5 \times 10^{-1}$, $\beta_2^{Adam} = 9.99 \times 10^{-1}$, on a Linux Fedora release 35 system with 96 CPUs AMD EPYC 7F72 24-Core Processor, 2 threads per core, and 259 GB RAM. The number of epochs was set to 120. The value of loss on the

validation set was calculated at each epoch, and the model with the lowest validation error over epochs was selected.

5.3.2 RX model 2

As discussed, the neural network in RX 2 is responsible for joint nonlinearity-PMD compensation upon receiving the sampled waveform after CD compensation with 2 SpS. As the signal is subject to the random phase noise effect, a CPE block is required outside the neural network to compensate for the phase noise. Despite this, the samples used for training the neural network are the pairs of normalized transmitted symbols at time-step i, and the corresponding window of the normalized signal output by the CD compensation block.

Table. 5.2 presents the details of the implemented neural networks for RX 2. The networks were tuned (Fig. 5.4) and trained using the same scheme and number of samples as for the neural networks in RX 1.

5.4 Complexity measurement

We use the number of incurred FLOPs as a metric for measuring the complexity of the models. In this section, we elaborate on the FLOPs that each of the MLP, CNN, RNN, GRU, and LSTM models incur.

For a fully-connected layer, according to (3.1.1),

$$\text{FLOPs}^{\{\text{fully-connected}\}} = n_i n_h + \eta n_h + n_h, \tag{5.4.1}$$

where n_i is the number of input features, n_h is the number of hidden units, and η is the number of FLOPs the nonlinear activation function takes. For a convolutional layer, also, the number of FLOPs is calculated using the following formula

$$\text{FLOPs}^{\{\text{conv}\}} = n_{ker} \times (2 \times e_{ker} - 1) \times L_{out} + \eta e_{out}, \tag{5.4.2}$$

where n_{ker} is the number of kernels, e_{ker} is the kernel shape, and e_{out} is the shape of the output feature-map. L_{out}, which is the length of the output feature-map, is obtained by

$$L_{out} = \lfloor \frac{L_{in} + 2 \times pad - (dil \times L_{ker} - 1) - 1}{strd} + 1 \rfloor, \tag{5.4.3}$$

where pad, dil and $strd$, respectively, signify padding, dilation [1], and stride. L_{ker} is the kernel length.

According to (3.3.1), the number of FLOPs performed in a vanilla RNN cell in one time-step is

$$\text{FLOPs}^{\{\text{RNN}\}} = n_i n_h + n_h^2 + 2n_h + \eta n_h, \tag{5.4.4}$$

From (3.3.2) the number of FLOPs incurred by an LSTM cell is derived as follows

$$\text{FLOPs}^{\{\text{LSTM}\}} = 4 \times [n_i n_h + n_h^2 + 3n_h] + 5\eta n_h, \tag{5.4.5}$$

In this calculation, with an approximation, it is assumed that tanh and sigmoid activations both impose the same number of FLOPs.

Based on (3.3.3), the number of FLOPs incurred by a GRU unit is calculated as follows

$$\text{FLOPs}^{\{\text{GRU}\}} = 3 \times [n_i n_h + n_h^2 + 2n_h + \eta n_h] + 5n_h, \tag{5.4.6}$$

In (5.4.6) similar to (5.4.5) it is assumed that sigmoid and tanh activations incur the same number of FLOPs.

Note that the number of FLOPs incurred by a recurrent layer equals the number of FLOPs incurred by one recurrent cell multiplied by the number of time-steps. For bidirectional layers, this value is doubled.

[1] dilation is off throughout this study

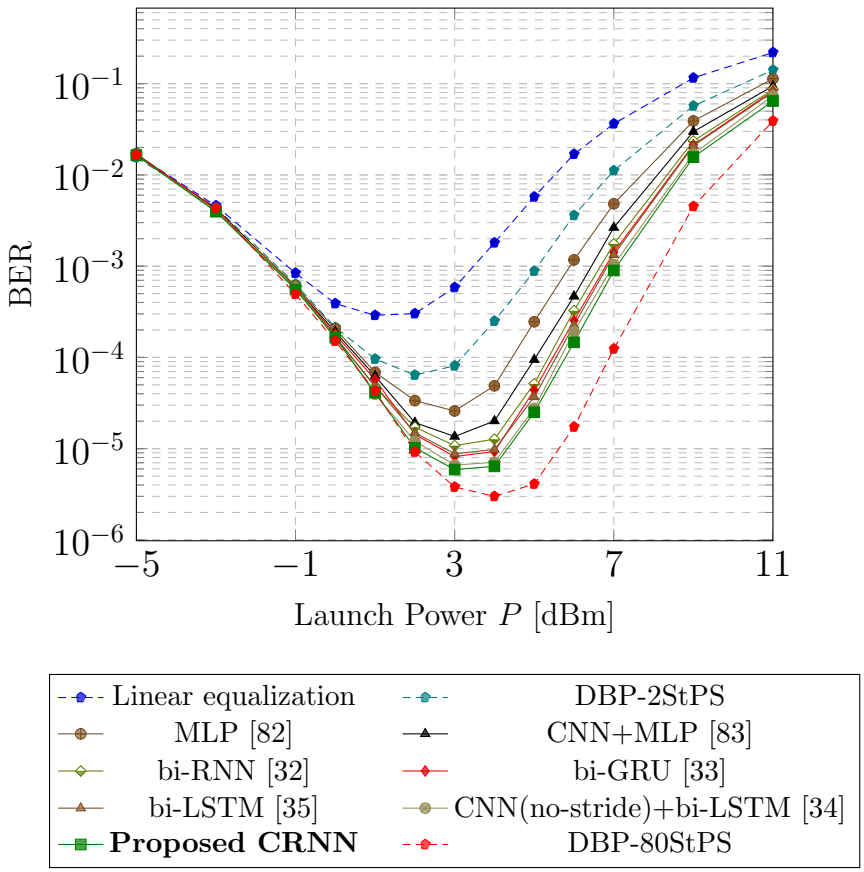


Figure 5.5: BER of the neural networks for RX 1 in the test mode as a function of the total launch power.

5.5 Evaluation

2^{24} unseen online-generated bits were used for the evaluation of the neural network models in both the RX types.

5.5.1 RX model 1

The resulting Q-factor and BER performance plots of the implemented approaches are demonstrated in Fig. 5.5, and Fig. 5.6, respectively. According to these figures, the recurrent-based equalizers provide slightly superior performance over MLP and CNN+MLP models with $0.42 - 0.69$ dB and $0.11 - 0.38$ dB Q-factor gain, respectively, at the optimal launch power, and in general $1.83 - 2.1$ dB gain over linear

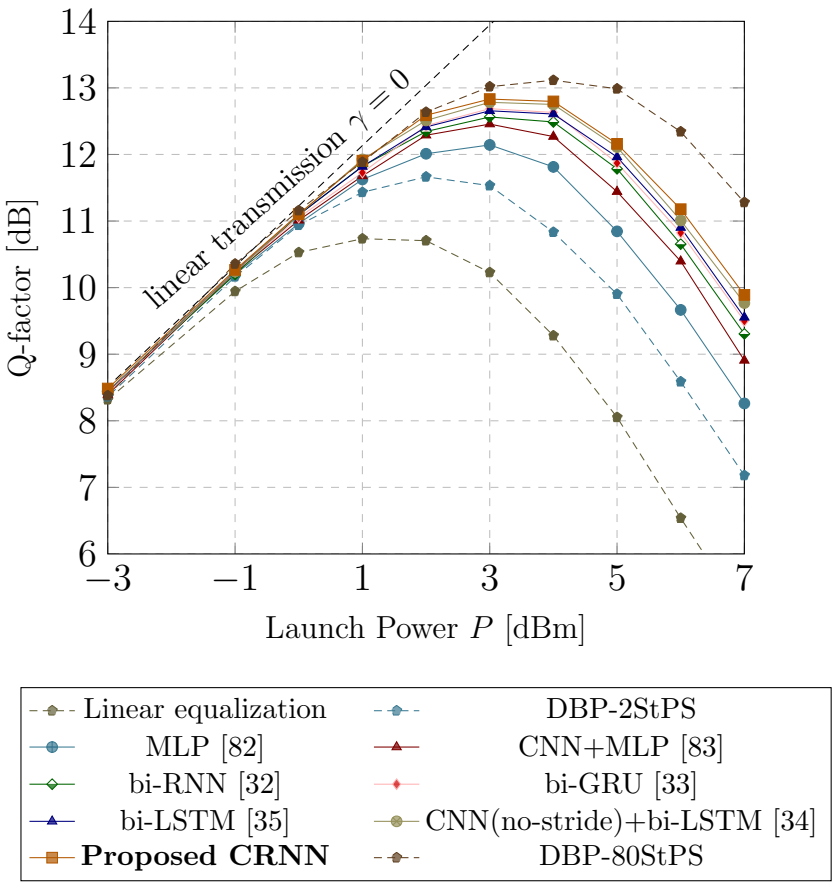


Figure 5.6: Q-factor of the neural networks for RX 1 in the test mode as a function of the total launch power.

equalization at optimal launch power. As previously noted, the inferior performance of MLP and CNN+MLP models is attributed to their susceptibility to over-fitting owing to the excessive amount of parameters associated with fully-connected hidden layers. Nonetheless, all the neural models achieve relatively close performance to each other with < 0.8 dB Q-factor difference. It is analyzed that this is in light of the fact that the number of parameters (weights) in the models are relatively limited ($< 10^7$) because of the limited size of the input. This results in mitigation of the over-fitting risk via an effective training strategy and escalating the likelihood of converging to the same low bias level via a suitable network complexity, which on the contrary, was investigated to be totally different for different models.

Table. 5.3 presents the complexity of the implemented neural network models measured by the number of FLOPs they incur in the inference mode per output symbol. These

Table 5.3: Neural networks' number of parameters and FLOPs per symbol, in RX 1.

Model	**#Params**	**#FLOPs/symbol**
MLP [82]	5,316,100	$\sim 26.90 \times 10^5$
CNN+MLP [83]	3,940,966	$\sim 21.79 \times 10^5$
bi-LSTM [35]	154,948	$\sim 54.61 \times 10^5$
bi-GRU [33]	144,292	$\sim 39.81 \times 10^5$
bi-RNN [32]	275,044	$\sim 33.40 \times 10^5$
CNN+bi-LSTM [27]	139.066	$\sim 47.39 \times 10^5$
Proposed CRNN	324,418	$\sim \mathbf{16.47} \times 10^5$

values are obtained analytically according to the formulas discussed in Section 5.4; however, they are also checked with their proportionality to the number of CPU cycles in the training mode obtained by the Linux kernel performance monitoring tool *perf* library. The full report of the CPU performance counters and trace-points recorded for each neural network is available at [95].

As Table. 5.3 demonstrates, although the proposed CRNN achieves a comparable performance to the bi-RNN based models, it has $> 50\%$ lower complexity compared to them, namely $\sim$50.6%, $\sim$58.6%, $\sim$69.8%, and $\sim$65.2% fewer number of FLOPs than bi-RNN, bi-GRU, bi-LSTM, and CNN+bi-LSTM models; which is thanks to the model's efficiency in minimizing the overlap between short-range and long-range dependencies extraction, having quite limited number of time-steps for the recurrent layer allowing for leveraging uni-directional many-to-one recurrent layer instead of bi-directional and using vanilla RNN in place of gated recurrent cells.

5.5.2 RX model 2

Fig. 5.7 and Fig. 5.8 demonstrate the resulting BER and Q-factor performance of the models [2]. As it is noticeable in comparison with the corresponding plots for RX 1, the

[2]The graphs are based on the output of the CPE block with the implementation of the neural network model prior to that.

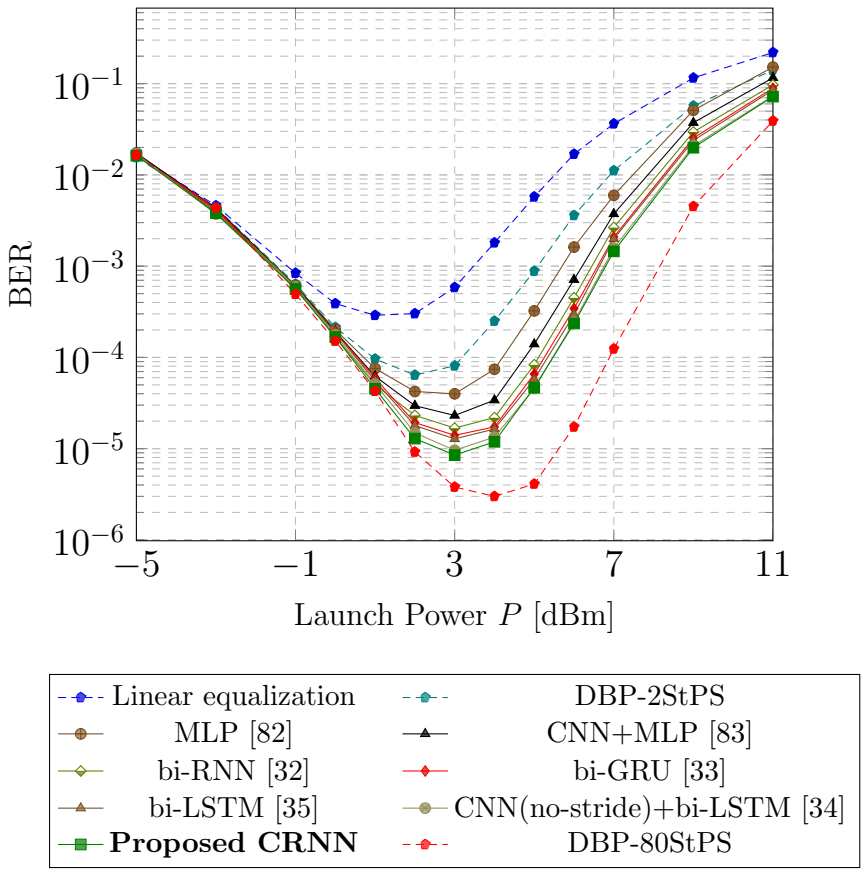


Figure 5.7: BER of the neural networks for RX 2 in the test mode as a function of the total launch power.

performance in RX 2 is diminished (averagely ∼0.2 dB), which is due to the destructive effect of unmitigated random phase noise in the training process.

In RX 2, similarly, although the neural networks achieve roughly comparable performance as demonstrated, they incur substantially different complexities. Table. 5.4 demonstrates the number of FLOPs incurred by each neural model in RX 2. As this table reports, the proposed CRNN incurs ∼52.1%, ∼58.1%, ∼69.7%, ∼65.4% lower complexity than bi-RNN, bi-GRU, bi-LSTM, and CNN+bi-LSTM models, respectively (roughly escalated ratios compared to that calculated for RX 1), which is in light of similar underlying reasons as discussed for RX 1, plus the dimensionality reduction towards 1 SpS prior to the recurrent layer.

Note that, as discussed in Section 5.1.3, in RX 2, the neural networks need to be retrained online frequently due to the polarization-dependant time-varying effects. The

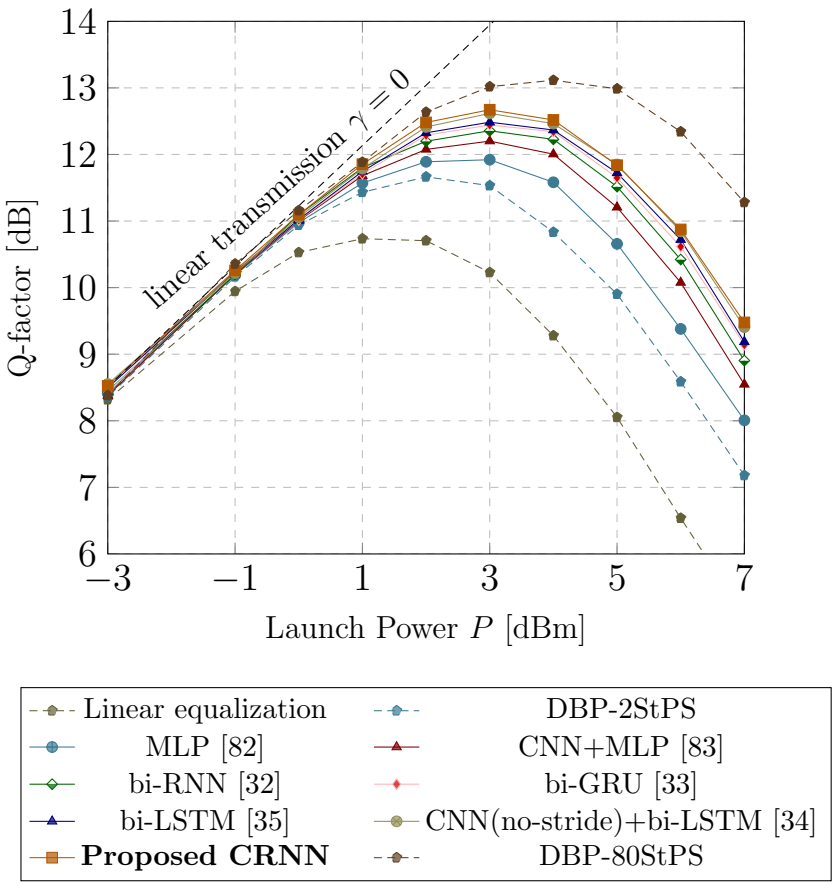


Figure 5.8: Q-factor of the neural networks for RX 2 in the test mode as a function of the total launch power.

neural networks need $\sim 10^5$ SGD iterations to be retrained. By assuming the number of FLOPs for one SGD iteration to be 4-6 times that of one forward propagation step, and the required retraining frequency to be 10^3 retrain/s, an overhead of $8 - 12\%$ on top of the complexities discussed in Table. 5.4, should be considered for the models, in addition to the high-speed memory requirement.

Table 5.4: Neural networks' number of parameters and FLOPs per symbol, in RX 2.

Model	#Params	#FLOPs/symbol
MLP [82]	6,348,292	$\sim 32.06 \times 10^5$
CNN+MLP [83]	2,680,183	$\sim 26.87 \times 10^5$
bi-LSTM [35]	176,964	$\sim 66.53 \times 10^5$
bi-GRU [33]	163,534	$\sim 48.68 \times 10^5$
bi-RNN [32]	324,814	$\sim 42.50 \times 10^5$
CNN+bi-LSTM [27]	166,002	$\sim 63.99 \times 10^5$
Proposed CRNN	344,281	$\sim \mathbf{20.36} \times 10^5$

5.5.3 RX 1 versus RX 2

The neural networks in RX 1 result in an average 0.2 dB superior Q-factor performance at optimal launch power compared to the neural networks in RX 2, as it is demonstrated in Fig. 5.9. We also investigated the case where the neural networks are trained using signals not affected by phase noise (which is not a realistic assumption), but tested with phase noise. In this scenario, RX 2 demonstrates a negligible performance improvement over RX 1 in the test mode. In short, while not having a higher performance than RX 1, the neural networks in RX 2 incur substantially higher complexities than RX 1, in addition to a frequent retraining requirement overhead, which is disproportionate to the number of FLOPs they pull out by eliminating the MIMO equalizer in the DSP chain. In view of this matter, it is believed that RX 1 is a better model to be adopted for deployment in fiber-optic transmission systems.

We also note that RX 1 and RX 2 result in different constellation diagrams.

As Fig. 5.10 shows, the constellation in RX 1 forms a square grid or "jail window" shape. This effect, which is commonly observed in other work [96]–[100], occurs when the neural network equalizer is trained based on regression using the MSE loss function. This, however, differs from the constellation obtained in RX 2. In RX 2, because of the random phase noise effect, persistent patterns do not exist in the input to be learned for detection. However, some forms of detection in the context of phase

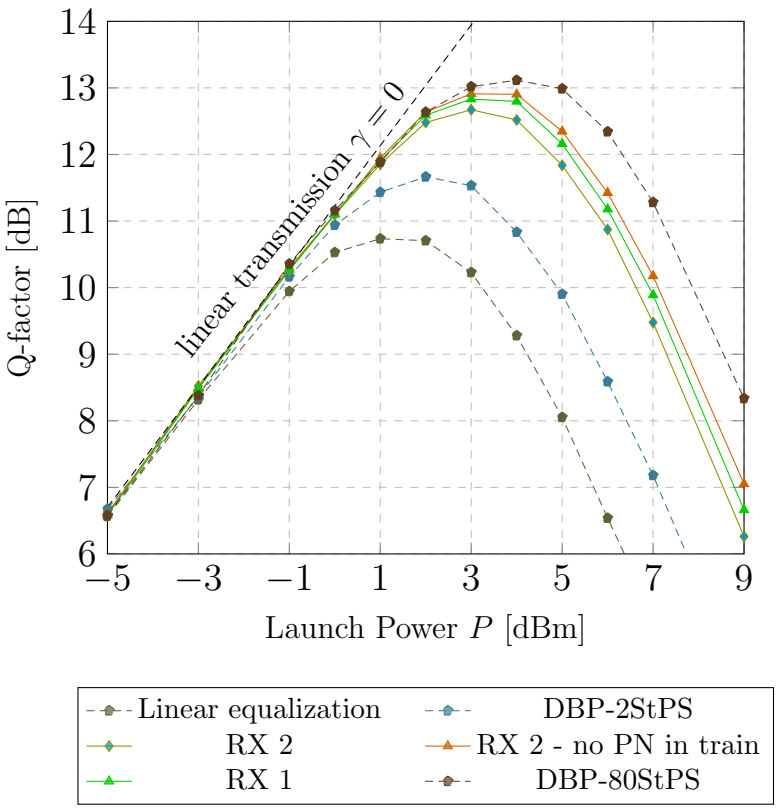


Figure 5.9: Comparison of the Q-factor performance of RX 1 and RX 2 in the test mode.

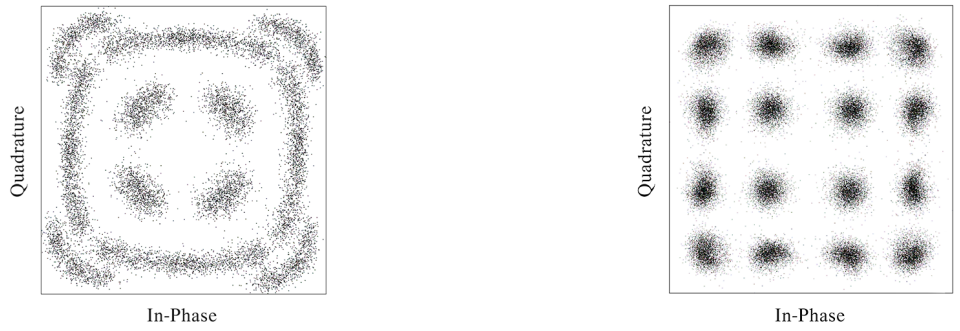
noise affected constellation happens, which results in non-uniform noise around the transmitted symbols. Fig. 5.11 illustrates a sample constellation after the neural network and after the CPE block in RX 2.

5.6 Conclusion

In this chapter, we provided a detailed description of the conducted investigations towards evaluating the proposed CRNN model in comparison with the state-of-the-art methods in neural network-based nonlinearity mitigation. We elaborated on the system architecture and parameters in detail, described the training and evaluation methodology precisely, and presented a comprehensive comparison of the performance versus complexity of the different adopted models.

(a) Before neural network (b) After neural network

Figure 5.10: Sample obtained constellation in RX 1 (a) after linear equalization (before the neural network), and (b) after the neural network.



(a) After neural network (b) After CPE

Figure 5.11: Sample obtained constellation in RX 2 (a) after the neural network, and (b) after the CPE block.

For 64 GBd dual-polarization 16-QAM optical transmission over 1120 km of standard single-mode fiber with 14 spans, we demonstrated that the proposed CRNN-based equalizer outperforms state-of-the-art bidirectional recurrent-based techniques while having $> 50\%$ less computational complexity measured by the number of FLOPs.

We also debated that having a neural network-based nonlinearity mitigation approach at the end of linear equalization could result in higher computational efficiency in terms of performance and complexity compared to a joint neural network-based nonlinearity-PMD compensation approach after chromatic dispersion compensation.

Conclusions

In this dissertation, we presented a concise review of the fundamentals of dual-polarization fiber-optic communication systems, its constituent components, and the different types of distortions that the signal is affected by through the propagation in the optical fiber. Considering the reviewed channel effects, we elaborated on the transmission techniques and digital signal processing carried out in these systems to mitigate the channel impairments and achieve higher information rates. We indicated that implementation of nonlinear equalization methods, which are required for equalizing the nonlinear channel effects, brings about a substantial computational overhead to the DSP unit compared to the linear DSP chain, especially with conventional solutions, in particular digital backpropagation.

In view of this matter, we introduced the concept of neural network-based equalization, which has recently piqued big deals of attention for obtaining comparable performance to conventional solutions while requiring lower computing resources. Towards this, we concisely reviewed the principles of neural networks, the underlying rationale, and the fundamental neural network structures for deep learning, utilizing which several neural network-based equalizers have recently been proposed in the community. We debated the state-of-the-art neural network-based equalization methods and made an analysis of their computational efficiency towards figuring out the stratagems within the models that could lead to computational inefficiency.

With regard to the analysis made, we proposed a hybrid neural network equalizer comprising a CNN-based encoder and a unidirectional many-to-one vanilla RNN working in tandem to mitigate nonlinear channel impairments in long-haul fiber-optic communications. We showed that the suggested CNN-based encoder not only minimizes the overlap between short-term and long-term dependencies extraction but also, by taking data into a latent space through multi-channeling and striding properties, it reduces

the number of significant features so that the subsequent layer can be implemented efficiently in the form of a unidirectional vanilla recurrent layer in lieu of a bidirectional gated recurrent layer.

We demonstrated that for 64 GBd dual-polarization 16-QAM optical transmission over 14×80 km SSMF, the suggested CRNN-based equalizer reaches a comparable performance to the state-of-the-art bidirectional recurrent-based approaches while having $> 50\%$ lower computational complexity compared to them; thanks to the efficiency of the model in minimizing the overlap between short-range and long-range dependencies extraction, leveraging unidirectional recurrent layer instead of bidirectional, using vanilla RNN in place of gated recurrent cells, and the dimensionality reduction prior to the recurrent layer.

We also investigated that when phase noise is present, a neural network-based approach to only compensate for the nonlinear impairments at the end of the linear DSP unit could result in greater efficiency in terms of performance and complexity compared to a combined nonlinearity-PMD mitigation solution via a neural network following the CD compensation block in the DSP unit.

Future studies

In this section, we discuss two of the compelling areas for future studies in the continuation of this doctoral thesis.

Attention-based pruning and quantization

One of the intriguing areas for future studies is to use an attention mechanism, as in [23], for punning and quantization of the connections in the proposed CRNN model. This process may lead to an ample complexity reduction in terms of the number of bit-wise operations in the interest of field-programmable gate array (FPGA) programming and high-performance computing (HPC) applied to the equalization process.

In a side study, presented in [23], we investigated that using an attention mechanism for learning and removing unnecessary connections in the final fully-connected layer of bidirectional recurrent-based models in block-to-block based equalization (as in [78]) could moderately reduce the number of FLOPs in the inference mode. The schematic of that approach is depicted in Fig. 5.12, where there exists one attention unit on top of each RNN cell to determine the intensity of the impact of that RNN time-step on each of the equalized symbol in the output vector. If it is turned out that the impact level of a connection is below a threshold, the connection can be dropped. Furthermore, in FPGA programming and circuit design, the number of bits dedicated to each connection can be assigned according to the determined impact.

In a future study, one could adopt this approach for determining the impact level of each/block of the connections in the unidirectional many-to-one recurrent layer of the proposed CRNN model by putting one attention unit per connection or a group of connections, depending on the available computing resources and the system setup.

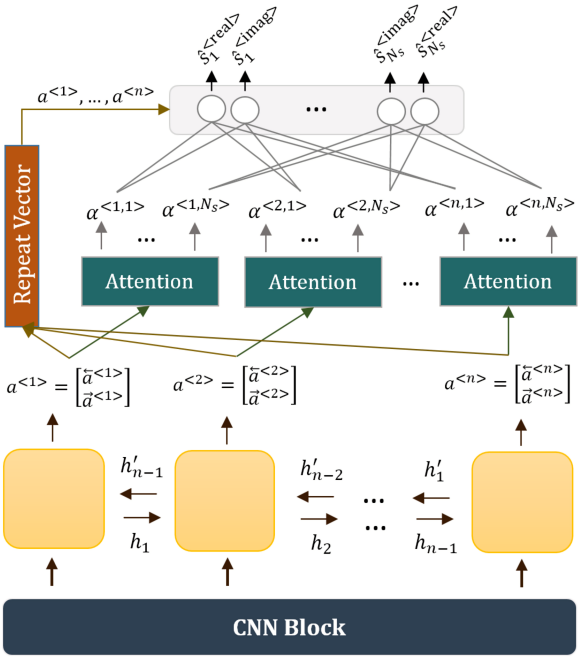


Figure 5.12: Presence of one attention unit on top of each RNN time-step j, in a CNN+bi-RNN -based equalizer, to determine the impact level of the RNN time-step on the equalized value for each of the symbol s_i.

Deep reinforcement learning-based joint transmitter and receiver optimization

One of the other interesting areas for future research is the concept of joint transmitter and receiver optimization using a hybrid of deep learning and reinforcement learning (RL), where the receiver is implemented via a neural network-based approach (the same as what is proposed in this manuscript), and the transmitter is implemented as an RL agent. The neural network is directly updated via variants of SGD at each iteration, and the RL-based transmitter gets trained according to an RL algorithm.

RL is about an agent interacting with the environment to learn an optimal policy for sequential decision-making problems, by trial and error. Trial-and-error learning is associated with so-called long-term rewards. In RL, the algorithm (agent) analyses the current situation (state), makes a decision, and receives feedback (reward) from the environment. Positive feedback denotes a reward, and negative feedback denotes a punishment due to the wrong decision. The main objective of RL is to determine the

best sequence of decisions that will enable an agent to solve a problem while optimizing long-term rewards [101, Ch. 1]. This is carried out via the following process.

At each time step t, the agent receives a state s_t in the state space $\mathcal{S}$ and selects an action a_t from the action space $\mathcal{A}$, according to the policy π. A policy is a mapping from states to probabilities of selecting each possible action, together with their associated rewards. In the event of an episodic problem, this cycle proceeds until the agent enters the terminal state, where it restarts. The return (accumulated reward) of a policy for the state s_t is calculated as

$$R_t = \sum_{k=0}^{T} \xi^k r_{t+k}, \tag{5.6.1}$$

where T is the time of termination, and r_t is the reward of the action at time step t, based on the policy. $\xi \in [0, 1)$ is the discount factor. The goal of the agent is to maximize the expectation of the return from each state. This is performed by estimating the value function. The value function measures the expected return for a state under a policy. Therefore, mathematically, the value of a state s under a policy π, denoted by $v_\pi(s)$, is defined as

$$v_\pi(s) = \mathbb{E}_\pi[R_t|s_t = s] = \mathbb{E}_\pi\left[\sum_{k=0}^{T} \xi^k r_{t+k}|s_t = s\right]. \tag{5.6.2}$$

Similarly, the value of taking action a at state s under policy π, is defined by the action-value function for policy π as

$$q_\pi(s, a) = \mathbb{E}_\pi[R_t|s_t = s, a_t = a] = \mathbb{E}_\pi\left[\sum_{k=0}^{T} \xi^k r_{t+k}|s_t = s, a_t = a\right]. \tag{5.6.3}$$

A foundational property of the value functions in RL is that for any policy π and any state s, the following consistency condition holds between the value of s and the value of its possible successor states [101, Ch. 3]

$$
\begin{aligned}
v_\pi(s) &= \mathbb{E}_\pi[R_t|s_t = s] \\
&= \mathbb{E}_\pi[r_{t+1} + \xi R_{t+1}|s_t = s] \\
&= \sum_a \pi(a|s) \sum_{s'} \sum_r p(s', r|s, a) \left[r + \xi \mathbb{E}_\pi[R_{t+1}|s_{t+1} = s']\right] \\
&= \sum_a \pi(a|s) \sum_{s',r} p(s', r|s, a) \left[r + \xi v_\pi(s')\right],
\end{aligned} \tag{5.6.4}
$$

where $\pi(a|s)$ denotes the probability that $a_t = a$ if $s_t = s$, under policy π.

Solving a RL task means finding the optimal policy π_*, whose expected return for all the states is higher than or equal to all the other policies. Optimal policies have the same state-value function, called the optimal state-value function. This function is defined as $v_*(s) = \max_\pi v_\pi(s)$ for all $s \in \mathcal{S}$. Optimal policies also have the same optimal action-value function $q_*(s, a) = \max_\pi q_\pi(s, a)$ for all $s \in \mathcal{S}$ and $a \in \mathcal{A}$. $\mathcal{A}(s)$ is the set of all available actions for the state s. Therefore,

$$
q_*(s, a) = \mathbb{E}_\pi[R_{t+1} + \xi v_*(s_{t+1})|s_t = s, a_t = a]. \tag{5.6.5}
$$

According to the Bellman optimality equation [102] and by (5.6.5),

$$
\begin{aligned}
v_*(s) &= \max_{a \in \mathcal{A}(s)} q_*(s, a) \\
&= \max_a \mathbb{E}_\pi[R_{t+1} + \xi v_*(s_{t+1})|s_t = s, a_t = a] \\
&= \max_a \sum_{s',r} p(s', r|s, a) \left[r + \xi v_*(s')\right].
\end{aligned} \tag{5.6.6}
$$

Various solutions have been proposed to achieve π_* through iterations of policy improvement, the prominent of which are Q-learning[103], policy gradient [104] , DQN [105], and DDPG [106]. The policy improvement process finishes when v_* is reached [101, Ch. 4], which is the case when

$$
\forall\, s \in \mathcal{S},\ \nexists\, a \in \mathcal{A}(s) \left[(a \neq \pi(s)) \wedge (q_\pi(s, a) > v_\pi(s))\right]. \tag{5.6.7}
$$

As discussed at the beginning of this section, one promising research field for future investigations is to extend the current CRNN -based receiver to a transceiver optimization approach, where the CRNN is trained using variants of SGD, based on the

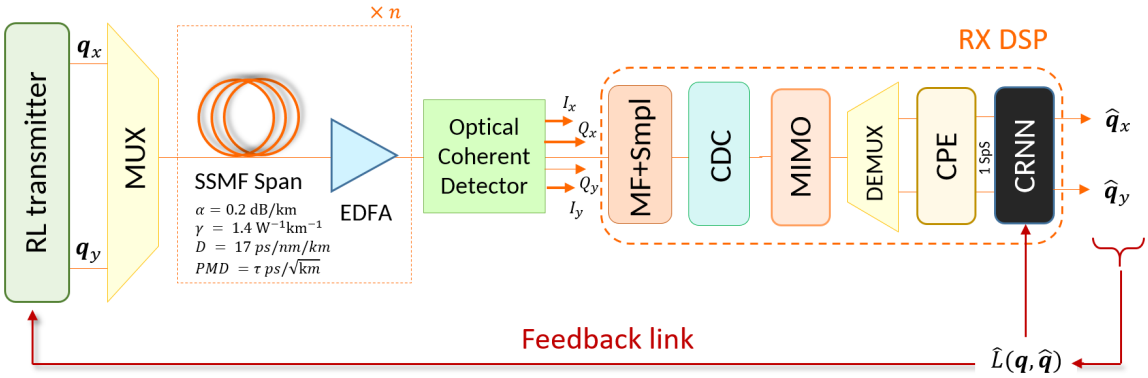


Figure 5.13: The general schematic of the potential deep reinforcement learning-based approach towards jointly optimizing transmitter and receiver, where the CRNN model is trained in a supervised learning manner via SGD, and the RL-based transmitter learns the optimal policy by updating its policy iteratively according to the feedback it receives.

calculated loss, and the transmitter is implemented as an RL agent which is optimized through the process of receiving rewards via a feedback link connected to the actions (an approach likewise [107], [108]). The domain of actions could be the adopted constellation shape, the pulse shape, or even the channel coding format. A schematic of this approach in the general form is illustrated in Fig. 5.13 where, roughly saying, the calculated loss is considered as the loss for the CRNN-based equalizer, and as the feedback for the transmitter.

This approach is appealing from the perspective that it does not require the full implementation of the channel's computational graph and prior knowledge about all the channel parameters as in the end-to-end deep learning-based methods [28], [109], [110]; thus, in case of success, it is considered to be more practical for realistic fiber-optic communication systems.

Besides the discussed research field for the future studies, leveraging the proposed CRNN model in post-processing of the nonlinear spectrum at RX in nonlinear Fourier transform (NFT)-based optical transmission systems [72], [111] is one the other noteworthy areas that could provide additional gains to these systems.

References

[1] M. Du, "Modern subsea cable systems: Towards ultimate cable capacity," in *Opt. Fiber Commun. Conf. Expo. (OFC)*, Mar. 2022, pp. 1–39. DOI: 10.1364/OFC.2022.M1F.1.

[2] I. Cisco Systems, *Cisco global 2022 forecast highlights*, https://www.cisco.com/c/dam/m/en_us/solutions/service-provider/vni-forecast-highlights/pdf/Global_2022_Forecast_Highlights.pdf, (accessed 20 5 2022), 2022.

[3] I. Cisco Systems, *Cisco annual internet report (2018–2023) white paper*, https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html, (accessed 20 5 2022), 2022.

[4] K. Kikuchi, "Fundamentals of coherent optical fiber communications," *IEEE J. Lightw. Technol.*, vol. 34, no. 1, pp. 157–179, Aug. 2015. DOI: 10.1109/JLT.2015.2463719.

[5] R.-J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *IEEE J. Lightw. Technol.*, vol. 28, no. 4, pp. 662–701, Feb. 2010. DOI: 10.1109/JLT.2009.2039464.

[6] G. P. Agrawal, *Fiber-optic communication systems*. John Wiley & Sons, 2012.

[7] E. Ip and J. M. Kahn, "Compensation of dispersion and nonlinear impairments using digital backpropagation," *IEEE J. Lightw. Technol.*, vol. 26, no. 20, pp. 3416–3425, Oct. 2008. DOI: 10.1109/JLT.2008.927791.

[8] G. Kramer, M. I. Yousefi, and F. R. Kschischang, "Upper bound on the capacity of a cascade of nonlinear and noisy channels," in *IEEE Inf. Theory Workshop (ITW)*, Jun. 2015, pp. 1–4. DOI: 10.1109/ITW.2015.7133167.

[9] F. P. Guiomar, J. D. Reis, A. L. Teixeira, and A. N. Pinto, "Mitigation of intra-channel nonlinearities using a frequency-domain volterra series equalizer," *Opt. Express*, vol. 20, no. 2, pp. 1360–1369, Sep. 2012. DOI: 10.1364/OE.20.001360.

[10] Y. Wang, L. Tao, X. Huang, J. Shi, and N. Chi, "Enhanced performance of a high-speed WDM CAP64 VLC system employing Volterra series-based nonlinear equalizer," *IEEE Photonics J.*, vol. 7, no. 3, pp. 1–7, Jun. 2015. DOI: 10.1109/JPHOT.2015.2436911.

[11] O. E. Agazzi, M. R. Hueda, H. S. Carrer, and D. E. Crivelli, "Maximum-likelihood sequence estimation in dispersive optical channels," *IEEE J. Lightw. Technol.*, vol. 23, no. 2, pp. 749–763, 2005.

[12] N. Alić, G. C. Papen, R. E. Saperstein, L. B. Milstein, and Y. Fainman, "Signal statistics and maximum likelihood sequence estimation in intensity modulated fiber optic links containing a single optical preamplifier," *Opt. Express*, vol. 13, no. 12, pp. 4568–4579, 2005.

[13] R. A. Fisher, *Optical phase conjugation*. Academic press, 2012.

[14] S. Jansen, D. Van Den Borne, P. Krummrich, S. Spalter, G.-D. Khoe, and H. De Waardt, "Long-haul DWDM transmission systems employing optical phase conjugation," *IEEE J. Sel. Top. Quantum Electron.*, vol. 12, no. 4, pp. 505–520, Aug. 2006. DOI: 10.1109/JSTQE.2006.876621.

[15] G. S. He, "Optical phase conjugation: Principles, techniques, and applications," *Prog. Quantum Electron.*, vol. 26, no. 3, pp. 131–191, May 2002. DOI: 10.1016/S0079-6727(02)00004-6.

[16] TeleGeography, *Map of global submarine cables*, https://www.submarinecablemap.com/, (accessed 20 5 2022), 2022.

[17] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *Bull. Math. Biophys.*, vol. 5, no. 4, pp. 115–133, Dec. 1943. DOI: 10.1007/BF02478259.

[18] L. Deng and D. Yu, "Deep learning: Methods and applications," *Found. Trends Signal Process.*, vol. 7, no. 3–4, pp. 197–387, 2014. DOI: 10.1561/2000000039.

[19] S. Pouyanfar, S. Sadiq, Y. Yan, *et al.*, "A survey on deep learning: Algorithms, techniques, and applications," *ACM Comput. Surv.*, vol. 51, no. 5, pp. 1–36, 2018. DOI: 10.1145/3234150.

[20] J. Ahmad, H. Farman, and Z. Jan, "Deep learning methods and applications," in *Deep learning: convergence to big data analytics*, Springer, 2019, pp. 31–42.

[21] F. Musumeci, C. Rottondi, A. Nag, *et al.*, "An overview on application of machine learning techniques in optical networks," *IEEE Commun. Surv. Tutor.*, vol. 21, no. 2, pp. 1383–1408, Nov. 2019. DOI: 10.1109/COMST.2018.2880039.

[22] A. Shahkarami, M. I. Yousefi, and Y. Jaouen, "Efficient deep learning of nonlinear fiber-optic communications using a convolutional recurrent neural network," in *IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Dec. 2021, pp. 668–673. DOI: 10.1109/ICMLA52953.2021.00112.

[23] A. Shahkarami, M. I. Yousefi, and Y. Jaouen, "Attention-based neural network equalization in fiber-optic communications," in *Asia Commun. Photonics Conf. (ACP)*, Dec. 2021, M5H–3. DOI: 10.1364/ACPC.2021.M5H.3.

[24] P. J. Freire, Y. Osadchuk, B. Spinnler, *et al.*, "Experimental study of deep neural network equalizers performance in optical links," in *Opt. Fiber Commun. Conf. Expo. (OFC)*, Jun. 2021, M3H–2. DOI: 10.1364/OFC.2021.M3H.2.

[25] C. Catanese, A. Triki, E. Pincemin, and Y. Jaouën, "A survey of neural network applications in fiber nonlinearity mitigation," in *Int. Conf. Transparent Opt. Netw. (ICTON)*, Jul. 2019, pp. 1–4. DOI: 10.1109/ICTON.2019.8840355.

[26] F. Da Ros, S. M. Ranzini, R. Dischler, *et al.*, "Machine-learning-based equalization for short-reach transmission: Neural networks and reservoir computing," in *Metro and Data Center Optical Networks and Short-Reach Links IV*, Proc. SPIE Int. Soc. Opt. Photonics, vol. 11712, Mar. 2021, p. 1 171 205. DOI: 10.1117/12.2583011.

[27] T. A. Eriksson, H. Bülow, and A. Leven, "Applying neural networks in optical communication systems: Possible pitfalls," *IEEE Photonics Technol. Lett.*, vol. 29, no. 23, pp. 2091–2094, Sep. 2017. DOI: 10.1109/LPT.2017.2755663.

[28] B. Karanov, M. Chagnon, F. Thouin, *et al.*, "End-to-end deep learning of optical fiber communications," *IEEE J. Lightw. Technol.*, vol. 36, no. 20, pp. 4843–4855, Aug. 2018. DOI: 10.1109/JLT.2018.2865109.

[29] T. Koike-Akino, Y. Wang, D. S. Millar, K. Kojima, and K. Parsons, "Neural turbo equalization: Deep learning for fiber-optic nonlinearity compensation," *IEEE J. Lightw. Technol.*, vol. 38, no. 11, pp. 3059–3066, Feb. 2020. DOI: 10.1109/JLT.2020.2976479.

[30] R. Hecht-Nielsen, "Theory of the backpropagation neural network," in *Neural networks Percept.* Elsevier, Jan. 1992, pp. 65–93. DOI: 10.1016/B978-0-12-741252-8.50010-8.

[31] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

[32] S. Deligiannidis, C. Mesaritakis, and A. Bogris, "Performance and complexity analysis of bi-directional recurrent neural network models versus Volterra nonlinear equalizers in digital coherent systems," *IEEE J. Lightw. Technol.*, vol. 39, no. 18, pp. 5791–5798, Jun. 2021. DOI: 10.1109/JLT.2021.3092415.

[33] X. Liu, Y. Wang, X. Wang, H. Xu, C. Li, and X. Xin, "Bi-directional gated recurrent unit neural network based nonlinear equalizer for coherent optical communication system," *Opt. Express*, vol. 29, no. 4, pp. 5923–5933, Feb. 2021. DOI: 10.1364/OE.416672.

[34] P. J. Freire, Y. Osadchuk, B. Spinnler, *et al.*, "Performance versus complexity study of neural network equalizers in coherent optical systems," *IEEE J. Lightw. Technol.*, vol. 39, no. 19, pp. 6085–6096, 2021. DOI: 10.1109/JLT.2021.3096286.

[35] S. Deligiannidis, A. Bogris, C. Mesaritakis, and Y. Kopsinis, "Compensation of fiber nonlinearities in digital coherent systems leveraging long short-term memory neural networks," *IEEE J. Lightw. Technol.*, vol. 38, no. 21, pp. 5991–5999, Jul. 2020. DOI: 10.1109/JLT.2020.3007919.

[36] Priyamstudycentre, *Optical fiber*, https://www.priyamstudycentre.com/2021/07/optical-fiber.html/, (accessed 10 4 2022), 2021.

[37] R. G. H. van Uden, "MIMO digital signal processing for optical spatial division multiplexed transmission systems," Ph.D. dissertation, Deptartment Electrical Engineering, Eindhoven University of Technology, 2014.

[38] G. Agrawal and a. O. M. C. Safari, *Nonlinear Fiber Optics, 6th Edition*. Academic Press, 2019.

[39] J. Toulouse, "Optical nonlinearities in fibers: Review, recent examples, and systems applications," *J. Light. Technol.*, vol. 23, no. 11, pp. 3625–3641, Dec. 2005. DOI: 10.1109/JLT.2005.855877.

[40] A. D. Boardman, M. Bertolotti, and T. Twardowski, *Nonlinear waves in solid state physics*. Springer Science & Business Media, 2012, vol. 247.

[41] R. G. Gallager, *Principles of Digital Communication*. Cambridge University Press, 2008.

[42] R. Paschotta, A. Schlatter, S. Zeller, H. Telle, and U. Keller, "Optical phase noise and carrier-envelope offset noise of mode-locked lasers," *Appl. Phys. B*, vol. 82, no. 2, pp. 265–273, Feb. 2006. DOI: 10.1007/s00340-005-2041-9.

[43] C. Menyuk, "Nonlinear pulse propagation in birefringent optical fibers," *IEEE J. Quantum Electron.*, vol. 23, no. 2, pp. 174–176, Feb. 1987. DOI: 10.1109/JQE.1987.1073308.

[44] D. Marcuse, C. Manyuk, and P. Wai, "Application of the Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence," *IEEE J. Lightw. Technol.*, vol. 15, no. 9, pp. 1735–1746, Sep. 1997. DOI: 10.1109/50.622902.

[45] I. FiberLabs Inc, *Spontaneous and stimulated emission*, https://www.fiberlabs.com/glossary/spontaneous-and-stimulated-emission/, (accessed 1 5 2022), 2021.

[46] E. Ip, A. P. T. Lau, D. J. Barros, and J. M. Kahn, "Coherent detection in optical fiber systems," *Opt. Express*, vol. 16, no. 2, pp. 753–791, Jan. 2008. DOI: 10.1364/OE.16.000753.

[47] S. J. Savory, "Digital coherent optical receivers: Algorithms and subsystems," *IEEE J. Sel. Top. Quantum Electron.*, vol. 16, no. 5, pp. 1164–1179, May 2010. DOI: 10.1109/JSTQE.2010.2044751.

[48] M. Ready and R. Gooch, "Blind equalization based on radius directed adaptation," in *Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Apr. 1990, 1699–1702 vol.3. DOI: 10.1109/ICASSP.1990.115806.

[49] T. Xu, P. Bayvel, T. Liu, *et al.*, "Carrier phase estimation in dispersion-unmanaged optical transmission systems," in *IEEE Adv. Inf. Technol. Electron. Autom. Control Conf. (IAEAC)*, Mar. 2017, pp. 1860–1864. DOI: 10.1109/IAEAC.2017.8054335.

[50] E. Serpedin, P. Ciblat, G. Giannakis, and P. Loubaton, "Performance analysis of blind carrier phase estimators for general qam constellations," *IEEE Trans. Signal Process.*, vol. 49, no. 8, pp. 1816–1823, Aug. 2001. DOI: 10.1109/78.934152.

[51] P.-K. A. Wai, C. R. Menyuk, and H. H. Chen, "Stability of solitons in randomly varying birefringent fibers," *Opt. Lett.*, vol. 16, no. 16, pp. 1231–1233, Aug. 1991. DOI: 10.1364/OL.16.001231.

[52] D. Michie, ""memo" functions and machine learning," *Nature*, vol. 218, no. 5136, pp. 19–22, Apr. 1968. DOI: 10.1038/218019a0.

[53] T. Hastie, R. Tibshirani, J. H. Friedman, and J. H. Friedman, *The elements of statistical learning: data mining, inference, and prediction*. Springer, 2009, vol. 2.

[54] T. O'Shea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Trans. Cognit. Commun. Networking*, vol. 3, no. 4, pp. 563–575, Oct. 2017. DOI: 10.1109/TCCN.2017.2758370.

[55] H. W. Lin, M. Tegmark, and D. Rolnick, "Why does deep and cheap learning work so well?" *J. Stat. Phys.*, vol. 168, no. 6, pp. 1223–1247, Sep. 2017. DOI: 10.1007/s10955-017-1836-5.

[56] N. Qian, "On the momentum term in gradient descent learning algorithms," *Neural networks*, vol. 12, no. 1, pp. 145–151, Jan. 1999. DOI: 10.1016/S0893-6080(98)00116-6.

[57] W. Rudin, *Real and Complex Analysis* (Mathematics series). McGraw-Hill, 2006, ISBN: 9780071002769.

[58] P. C. Petersen, "Neural network theory," *University of Vienna*, Apr. 2022.

[59] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural networks*, vol. 2, no. 5, pp. 359–366, Mar. 1989. DOI: 10.1016/0893-6080(89)90020-8.

[60] C. Kong and S. Lucey, "Take it in your stride: Do we need striding in cnns?" *arXiv preprint arXiv:1712.02502*, 2017.

[61] Y. Yu, X. Si, C. Hu, and J. Zhang, "A review of recurrent neural networks: LSTM cells and network architectures," *Neural Comput.*, vol. 31, no. 7, pp. 1235–1270, Jul. 2019. DOI: 10.1162/neco_a_01199.

[62] M. Ravanelli, P. Brakel, M. Omologo, and Y. Bengio, "Light gated recurrent units for speech recognition," *IEEE Trans. Emerg. Top. Comput. Intell.*, vol. 2, no. 2, pp. 92–102, Mar. 2018. DOI: 10.1109/TETCI.2017.2762739.

[63] Z. Luo, J. Xiong, F. Luo, *et al.*, "UAV spectrum sensing of normalized spectrum based on gru network," in *IEEE Int. Conf. Unmanned Syst. (ICUS)*, Oct. 2021, pp. 328–333. DOI: 10.1109/ICUS52573.2021.9641069.

[64] D. Nadig, B. Ramamurthy, B. Bockelman, and D. Swanson, "APRIL: An application-aware, predictive and intelligent load balancing solution for data-intensive science," in *IEEE Conf. Comput. Commun. Workshops (INFOCOM)*, May 2019, pp. 1909–1917. DOI: 10.1109/INFOCOM.2019.8737537.

[65] M. Popel, M. Tomkova, J. Tomek, *et al.*, "Transforming machine translation: A deep learning system reaches news translation quality comparable to human professionals," *Nat. Commun.*, vol. 11, no. 1, pp. 1–15, Sep. 2020. DOI: 10.1038/s41467-020-18073-9.

[66] A. Graves, "Generating sequences with recurrent neural networks," *arXiv preprint*, Jun. 2013. DOI: 10.48550/arXiv.1308.0850.

[67] A. C.-C. Yao, "On constructing minimum spanning trees in k-dimensional spaces and related problems," *SIAM J. Comput.*, vol. 11, no. 4, pp. 721–736, 1982.

[68] K. G. Larsen and J. Nelson, "Optimality of the johnson-lindenstrauss lemma," in *IEEE Annu. Symp. Found. Comput. Sci. (FOCS)*, IEEE, 2017, pp. 633–638.

[69] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley Interdiscip. Rev.: Comput. Stat.*, vol. 2, no. 4, pp. 433–459, Jul. 2010. DOI: 10.1002/wics.101.

[70] J. C. Cartledge, F. P. Guiomar, F. R. Kschischang, G. Liga, and M. P. Yankov, "Digital signal processing for fiber nonlinearities," *Opt. Express*, vol. 25, no. 3, pp. 1916–1936, Feb. 2017. DOI: 10.1364/OE.25.001916.

[71] P. J. Winzer, D. T. Neilson, and A. R. Chraplyvy, "Fiber-optic transmission and networking: The previous 20 and the next 20 years," *Opt. Express*, vol. 26, no. 18, pp. 24 190–24 239, Sep. 2018. DOI: 10.1364/OE.26.024190.

[72] M. I. Yousefi and F. R. Kschischang, "Information transmission using the nonlinear Fourier transform, part i, ii, iii," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4312–4369, Apr. 2014. DOI: 10.1109/TIT.2014.2321143.

[73] M. Secondini, D. Marsella, and E. Forestieri, "Enhanced split-step Fourier method for digital backpropagation," in *Eur. Conf. Opt. Commun. (ECOC)*, 2014, pp. 1–3.

[74] A. Napoli, Z. Maalej, V. A. Sleiffer, *et al.*, "Reduced complexity digital back-propagation methods for optical communication systems," *IEEE J. Lightw. Technol.*, vol. 32, no. 7, pp. 1351–1362, 2014.

[75] A. Amari, X. Lin, O. A. Dobre, R. Venkatesan, and A. Alvarado, "A machine learning-based detection technique for optical fiber nonlinearity mitigation," *IEEE Photonics Technol. Lett.*, vol. 31, no. 8, pp. 627–630, Mar. 2019. DOI: 10.1109/LPT.2019.2902973.

[76] A. Balatsoukas-Stimming and C. Studer, "Deep unfolding for communications systems: A survey and some new directions," in *IEEE Int. Workshop Signal Process. Syst. (SiPS)*, Oct. 2019, pp. 266–271. DOI: 10.1109/SiPS47522.2019.9020494.

[77] C. Häger and H. D. Pfister, "Nonlinear interference mitigation via deep neural networks," in *Opt. Fiber Commun. Conf. Expo. (OFC)*, Mar. 2018, pp. 1–3. DOI: 10.1364/OFC.2018.W3A.4.

[78] C. Häger and H. D. Pfister, "Deep learning of the nonlinear Schrödinger equation in fiber-optic communications," in *IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2018, pp. 1590–1594. DOI: 10.1109/ISIT.2018.8437734.

[79] R. M. Butler, C. Hager, H. D. Pfister, G. Liga, and A. Alvarado, "Model-based machine learning for joint digital backpropagation and PMD compensation," *IEEE J. Lightw. Technol.*, pp. 1–1, 2020. DOI: 10.1109/JLT.2020.3034047.

[80] O. Sidelnikov, A. Redyuk, S. Sygletos, M. Fedoruk, and S. Turitsyn, "Advanced convolutional neural networks for nonlinearity mitigation in long-haul WDM transmission systems," *IEEE J. Lightw. Technol.*, vol. 39, no. 8, pp. 2397–2406, Jan. 2021. DOI: 10.1109/JLT.2021.3051609.

[81] X. Jiang, D. Wang, Q. Fan, M. Zhang, C. Lu, and A. P. T. Lau, "Solving the nonlinear schrödinger equation in optical fibers using physics-informed neural network," in *Opt. Fiber Commun. Conf. Expo. (OFC)*, Jun. 2021, pp. 1–3. DOI: 10.1364/OFC.2021.M3H.8.

[82] C. Catanese, R. Ayassi, E. Pincemin, and Y. Jaouën, "A fully connected neural network approach to mitigate fiber nonlinear effects in 200G DP-16-QAM transmission system," in *Int. Conf. Transparent Opt. Netw. (ICTON)*, 2020, pp. 1–4. DOI: 10.1109/ICTON51198.2020.9203197.

[83] C. Chuang, L. Liu, C. Wei, *et al.*, "Convolutional neural network based nonlinear classifier for 112-Gbps high speed optical link," in *Opt. Fiber Commun. Conf. Expo. (OFC)*, Mar. 2018, pp. 1–3. DOI: 10.1364/OFC.2018.W2A.43.

[84] H. Zhang, Z. Yu, L. Shu, Z. Wan, Y. Zhao, and K. Xu, "Fiber nonlinearity equalizer using MLP-ANN for coherent optical OFDM," in *Int. Conf. Opt. Commun. Netw. (ICOCN)*, 2019, pp. 1–3. DOI: 10.1109/ICOCN.2019.8934433.

[85] O. Sidelnikov, A. Redyuk, and S. Sygletos, "Equalization performance and complexity analysis of dynamic deep neural networks in long haul transmission systems," *Opt. Express*, vol. 26, no. 25, pp. 32 765–32 776, Dec. 2018. DOI: 10.1364/OE.26.032765.

[86] P. Li, L. Yi, L. Xue, and W. Hu, "56 Gbps IM/DD PON based on 10G-class optical devices with 29 dB loss budget enabled by machine learning," in *Opt. Fiber Commun. Conf. Expo. (OFC)*, Mar. 2018, pp. 1–3. DOI: 10.1364/OFC.2018.M2B.2.

[87] P. Li, L. Yi, L. Xue, and W. Hu, "100Gbps IM/DD transmission over 25km SSMF using 20G-class DML and PIN enabled by machine learning," in *Opt. Fiber Commun. Conf. Expo. (OFC)*, Mar. 2018, pp. 1–3. DOI: 10.1364/OFC.2018.W2A.46.

[88] C. Gagné, M. Zeng, L. A. Rusch, *et al.*, "Recurrent neural networks achieving mlse performance for optical channel equalization," *Opt. Express*, vol. 29, no. 9, pp. 13 033–13 047, Apr. 2021. DOI: 10.1364/OE.423103.

[89] B. Karanov, D. Lavery, P. Bayvel, and L. Schmalen, "End-to-end optimized transmission over dispersive intensity-modulated channels using bidirectional recurrent neural networks," *Opt. Express*, vol. 27, no. 14, pp. 19 650–19 663, Jul. 2019. DOI: 10.1364/OE.27.019650.

[90] B. Karanov, M. Chagnon, V. Aref, *et al.*, "Experimental investigation of deep learning for digital signal processing in short reach optical fiber communications," in *IEEE Int. Workshop Signal Process. Syst. (SiPS)*, Sep. 2020, pp. 1–6. DOI: 10.1109/SiPS50750.2020.9195215.

[91] I. Cisco Systems, *Introduction to optical fibers, db, attenuation and measurements*,
https://www.cisco.com/c/en/us/support/docs/optical/synchronous-digital-hierarchy-sdh/29000-db-29000.html/, (accessed 20 5 2022), 2022.

[92] TeleGeography, *C-lion1 submarine cable*, https://www.submarinecablemap.com/submarine-cable/c-lion1/, (accessed 20 5 2022), 2022.

[93] I. Fatadin, D. Ives, and S. J. Savory, "Blind equalization and carrier phase recovery in a 16-QAM optical coherent system," *IEEE J. Lightw. Technol.*, vol. 27, no. 15, pp. 3042–3049, May 2009. DOI: 10.1109/JLT.2009.2021961.

[94] A. Shahkarami, M. I. Yousefi, and Y. Jaouen, *Fiber-optic transmission system modeling, accompanied with the implementation of TX and RX DSP chain*, https://github.com/FONTE-EID/fiber-optic-transmission-system-modeling/, (accessed 14 April 2022), 2022.

[95] A. Shahkarami, M. I. Yousefi, and Y. Jaouen, *Monitoring report of CPU cycles and performance counters in 1-epoch training of recurrent-based neural network equalizers*, Available: https://www.kaggle.com/datasets/performancedata/nn-nonlinearity-mitigation/, [Online; accessed 10-July-2022], 2022.

[96] C. Bluemm, M. Schaedler, S. Calabrò, *et al.*, "Equalizing nonlinearities with memory effects: Volterra series vs. deep neural networks," in *Eur. Conf. Opt. Commun. (ECOC)*, Sep. 2019, pp. 1–4. DOI: 10.1049/cp.2019.0945.

[97] M. A. Jarajreh, E. Giacoumidis, I. Aldaya, *et al.*, "Artificial neural network nonlinear equalizer for coherent optical ofdm," *IEEE Photonics Technol. Lett.*, vol. 27, no. 4, pp. 387–390, Dec. 2015. DOI: 10.1109/LPT.2014.2375960.

[98] L. Liu, M. Bi, S. Xiao, J. Fang, T. Huang, and W. Hu, "Ols-based rbf neural network for nonlinear and linear impairments compensation in the co-ofdm system," *IEEE Photonics J.*, vol. 10, no. 2, pp. 1–8, Feb. 2018. DOI: 10.1109/JPHOT.2018.2808919.

[99] J. Zhang, P. Lei, S. Hu, *et al.*, "Functional-link neural network for nonlinear equalizer in coherent optical fiber communications," *IEEE Access*, vol. 7, pp. 149 900–149 907, Oct. 2019. DOI: 10.1109/ACCESS.2019.2947278.

[100] O. Kotlyar, M. Kamalian-Kopae, M. Pankratova, A. Vasylchenkova, J. E. Prilepsky, and S. K. Turitsyn, "Convolutional long short-term memory neural network equalizer for nonlinear fourier transform-based optical transmission systems," *Opt. Express*, vol. 29, no. 7, pp. 11 254–11 267, Mar. 2021. DOI: 10.1364/OE.419314.

[101] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

[102] R. Bellman, "On the theory of dynamic programming," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 38, no. 8, p. 716, 1952.

[103] C. J. Watkins and P. Dayan, "Q-learning," *Machine learning*, vol. 8, no. 3-4, pp. 279–292, 1992.

[104] R. S. Sutton, D. A. McAllester, S. P. Singh, and Y. Mansour, "Policy gradient methods for reinforcement learning with function approximation," in *Conf. Neural Inf. Process. Syst. (NeurIPS)*, 2000, pp. 1057–1063.

[105] V. Mnih, K. Kavukcuoglu, D. Silver, *et al.*, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.

[106] T. P. Lillicrap, J. J. Hunt, A. Pritzel, *et al.*, "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, 2015.

[107] M. Goutay, F. A. Aoudia, and J. Hoydis, "Deep reinforcement learning autoencoder with noisy feedback," in *Int. Symp. Model. Optim. Mob. Ad Hoc Wireless Networks (WiOPT)*, Jun. 2019, pp. 1–6. DOI: 10.23919/WiOPT47501.2019.9144089.

[108] F. A. Aoudia and J. Hoydis, "End-to-end learning of communications systems without a channel model," in *Asilomar Conf. Signals Syst. Comput. (ACSSC)*, Oct. 2018, pp. 298–303. DOI: 10.1109/ACSSC.2018.8645416.

[109] M. Li, D. Wang, Q. Cui, Z. Zhang, L. Deng, and M. Zhang, "End-to-end learning for optical fiber communication with data-driven channel model," in *Opto-Electron. Commun. Conf. (OECC)*, Oct. 2020, pp. 1–3. DOI: 10.1109/OECC48412.2020.9273665.

[110] V. Neskorniuk, A. Carnio, V. Bajaj, *et al.*, "End-to-end deep learning of long-haul coherent optical fiber communications via regular perturbation model," in *Eur. Conf. Opt. Commun. (ECOC)*, Sep. 2021, pp. 1–4. DOI: 10.1109/ECOC52684.2021.9605928.

[111] O. Kotlyar, M. Pankratova, M. Kamalian-Kopae, A. Vasylchenkova, J. E. Prilepsky, and S. K. Turitsyn, "Combining nonlinear Fourier transform and neural network-based processing in optical communications," *Opt. Lett.*, vol. 45, no. 13, pp. 3462–3465, Jul. 2020. DOI: 10.1364/OL.394115.

ECOLE
DOCTORALE

Titre : Réduction de la complexité de l'égalisation de la non-linéarité Kerr dans les communications sur fibre optique à double polarisation par une approche de réseaux de neurones récurrents convolutifs

Mots clés : Communications sur fibre optique, égalisation de non-linéarité, réseaux de neurones, réduction de la complexité de calcul, réseaux de neurones récurrents convolutifs.

Résumé : Les dégradations dues à la non-linéarité de Kerr dans les fibres optiques limitent les débits d'information des systèmes de communications. Les effets linéaires, tels que la dispersion chromatique et la dispersion modale de polarisation, peuvent être compensés par égalisation linéaire, de mise oeuvre relativement simple, au niveau du récepteur. A l'inverse, la complexité de calcul des techniques classiques de réduction de la non-linéarité, telles que la rétro-propagation numérique, peut être considérable.
Les réseaux neuronaux ont récemment attiré l'attention, dans ce contexte, pour la mise en oeuvre d'égaliseurs non-linéaires à faible complexité. Cette thèse porte sur l'étude des réseaux neuronaux récurrents pour compenser efficacement les dégradations des canaux dans les transmissions à longue distance multiplexés en polarisation. Nous présentons une architecture hybride de réseaux neuronaux récurrents convolutifs (CRNN), comprenant un encodeur basé sur un réseau neuronal convolutif (CNN) suivie d'une couche récurrente travaillant en tandem. L'encodeur basé sur CNN représente efficacement la mémoire de canal à court terme résultant de la dispersion chromatique, tout en faisant passer le signal vers un espace latent avec moins de caractéristiques pertinentes. La couche récurrente suivante est implémentée sous la forme d'un RNN unidirectionnel de type vanille, chargé de capturer les interactions à longue portée négligées par l'encodeur CNN. Nous démontrons que le CRNN proposé atteint la performance des égaliseurs actuels dans la communication par fibre optique, avec une complexité de calcul significativement plus faible selon le modèle du système. Enfin, le compromis performance-complexité est établi pour un certain nombre de modèles, y compris les réseaux neuronaux multicouches entièrement connectés, les CNN, les réseaux neuronaux récurrents bidirectionnels, les réseaux long short-term memory bidirectionnels (bi-LSTM), les réseaux gated recurrent units bidirectionnels, les modèles bi-LSTM convolutifs et le modèle hybride proposé.

Title : Complexity reduction over bi-RNN-based Kerr nonlinearity equalization in dual-polarization fiber-optic communications via a CRNN-based approach

Keywords : Optical fiber communications, nonlinearity equalization, neural networks, complexity reduction, convolutional recurrent neural networks.

Abstract :
The impairments arising from the Kerr nonlinearity in optical fibers limit the achievable information rates in fiber-optic communication. Unlike linear effects, such as chromatic dispersion and polarization-mode dispersion, which can be compensated via relatively simple linear equalization at the receiver, the computational complexity of the conventional nonlinearity mitigation techniques, such as the digital back-propagation, can be substantial.
Neural networks have recently attracted attention, in this context, for low-complexity nonlinearity mitigation in fiber-optic communications. This Ph.D. dissertation deals with investigating the recurrent neural networks to efficiently compensate for the nonlinear channel impairments in dual-polarization long-haul fiber-optic transmission. We present a hybrid convolutional recurrent neural network (CRNN) architecture, comprising a convolutional neural network (CNN) -based encoder followed by a recurrent layer working in tandem. The CNN-based encoder represents the short-term channel memory arising from the chromatic dispersion efficiently, while transitioning the signal to a latent space with fewer relevant features. The subsequent recurrent layer is implemented in the form of a unidirectional vanilla RNN, responsible for capturing the long-range interactions neglected by the CNN encoder. We demonstrate that the proposed CRNN achieves the performance of the state-of-the-art equalizers in optical fiber communication, with significantly lower computational complexity depending on the system model. Finally, the performance-complexity trade-off is established for a number of models, including multi-layer fully-connected neural networks, CNNs, bidirectional recurrent neural networks, bidirectional long short-term memory (bi-LSTM), bidirectional gated recurrent units, convolutional bi-LSTM models, and the suggested hybrid model.

Institut Polytechnique de Paris
91120 Palaiseau, France