Neuromorphic Computing and Machine Learning for Nonlinearity Mitigation in Optical Communication Systems

Diego Argüello Ron

University student number: [redacted]

Doctor of Philosophy

Aston University

April 2025

Aston University

Title: Neuromorphic Computing and Machine Learning for Nonlinearity Mitigation in Optical Communication Systems
Name: Diego Argüello Ron
Degree: Doctor of Philosophy
Year of submission for examination: 2025

Nonlinear fibre propagation and transceiver impairments limit the performance of coherent optical communication systems. Established digital signal processing techniques such as digital back-propagation and Volterra equalisation can mitigate Kerr-induced nonlinear distortions, but their computational complexity and energy consumption hinder real-time deployment. This thesis develops machine-learning and neuromorphic approaches to optical channel equalisation, focusing on performance–complexity trade-offs and practical hardware feasibility.

A unified training and benchmarking framework is introduced for coherent transmission scenarios, employing Q-factor and bit-error-rate alongside implementation measures including operation counts, memory footprint, and inference latency. Within this framework, a broad class of equalisers is analysed, including multi-layer perceptrons, convolutional and recurrent neural networks, and complex-valued neural networks that process in-phase and quadrature components. Model compression is studied: pruning, quantisation, and weight clustering are jointly optimised using Bayesian optimisation to identify Pareto-efficient configurations that preserve equalisation performance while reducing computational load, memory usage, and latency. Experimental evaluations on edge-device platforms demonstrate feasibility under realistic receiver constraints.

The thesis also explores hardware–software co-design. Optical phase conjugation is integrated with neural equalisers to offload part of the nonlinearity compensation to the optical domain, enabling smaller models and lower digital complexity. To address analogue noise in photonic neuromorphic computing, robustness is assessed under additive and signal-dependent noise; noise-aware training, stochastic-resonance neurons, and ensemble-based “crowd equalisation” are proposed to enhance resilience. Finally, a neuromorphic equaliser combining spiking neural networks with a streaming RWKV time-mixing module is introduced. By exploiting event-driven sparsity and constant-memory sequential processing, this architecture achieves competitive equalisation performance with reduced computational and energy requirements compared to conventional deep learning models.

Keywords: coherent optical communications; nonlinearity compensation; neural networks; pruning; quantisation; neuromorphic computing

Dedication

To my parents, for raising me by example with hard work and integrity,
and to my sister, Pati, my best friend.
In memory of my grandfathers and loved ones who could not be here to see this completed.

Personal Acknowledgements

This thesis would not have been possible without the support, guidance, and patience of many people.

I am deeply grateful to Professor Sergei K. Turitsyn for his steady mentorship and for believing in me when I struggled to believe in myself. I also thank Dr Jaroslaw E. Prilepsky for teaching me how to write and refine research papers, and for his consistent support and generosity. I am grateful to Dr Morteza Kamalian-Kopae for his mentorship and for always taking the time to listen, and to Dr Egor Manuylovich for his help, encouragement, and openness.

At Aston, I am especially grateful to Karina Nurlybayeva and Dr Pedro J. Freire for their friendship, for helping me feel part of the community, and for many conversations that kept me grounded.

Finally, I thank my parents for their unconditional support throughout the PhD, and for raising me by example. I thank my sister, Pati, for being my closest friend. I carry this achievement with the memory of my grandfathers, whom I wish could have shared this moment.

Collaborator Acknowledgements

This thesis is based in part on research undertaken collaboratively with colleagues at Aston University and partner institutions. I am grateful to all co-authors and collaborators for their technical input, constructive discussions, and support throughout the development, validation, and dissemination of the work presented here.

For the research on reducing the computational cost and improving the energy efficiency of neural-network-based optical channel equalisation, I thank Dr Pedro J. Freire for implementing multiple neural network models, producing key figures, and providing both simulated and experimental datasets. I also acknowledge Dr Antonio Napoli, Dr Bernhard Spinnler, Dr Michael Anderson, Dr Wolfgang Schairer, Dr Thomas Bex, and Dr Nelson Costa for valuable feedback and technical guidance that shaped the scope and positioning of this work. In particular, I wish to recognise Dr Pedro J. Freire for originating and championing the core framing that combines Bayesian Optimisation with model-compression techniques (notably pruning, quantisation, and weight clustering) as a principled route to identifying performance–complexity Pareto-efficient equalisers; this perspective and the supporting data generation provided a strong foundation for the optimisation and deployment analyses developed in this thesis.

For the work on noise-resistant crowd equalisation and noise-aware approaches, I thank Karina Nurlybayeva for leading the study, performing thorough optimisation, and preparing key results and plots. I further acknowledge the contributions of Dr Morteza Kamalian-Kopae, Dr Elena Turitsyna, and Professor Sergei K. Turitsyn, and I thank Dr Egor Manuylovich for contributions related to stochastic resonance neuron modelling that enriched the noise-utilisation direction of this research.

For the study on the impact of optical phase conjugation on the computational complexity of neural-network-based equalisers, I thank Abdallah Alsayed Ismail Ali for generating the experimental datasets underpinning validation, and I acknowledge Karina Nurlybayeva, Dr Morteza Kamalian-Kopae, Dr Elena Turitsyna, and Professor Sergei K. Turitsyn for their contributions to analysis and interpretation. For the work on hardware-restricted implementation using pruning and quantisation, I again thank Dr Pedro J. Freire for datasets and figures, and I acknowledge guidance and input from Dr Jaroslaw E. Prilepsky, Dr Morteza Kamalian-Kopae, Dr Antonio Napoli, and Professor Sergei K. Turitsyn.

Across these collaborative works, I led the overall method development, experiments and/or simulations, analysis, and manuscript preparation, with my collaborators contributing complementary

expertise and resources as noted above.

Finally, I offer special recognition to Dr Jaroslaw E. Prilepsky, Dr Morteza Kamalian-Kopae, and Professor Sergei K. Turitsyn for their sustained mentorship on research strategy and academic writing; their guidance materially strengthened both the scientific narrative and the clarity of presentation throughout the thesis. I also acknowledge the broader research community whose discussions and feedback helped refine the work, and the institutional and funding support that enabled the research to be carried out.

Contents

List of Abbreviations

1D One-dimensional.

AC Accumulate Operation.

ADC Analog-to-Digital Converter.

AI Artificial Intelligence.

ANN Artificial Neural Network.

ASE Amplified Spontaneous Emission.

ASIC Application-Specific Integrated Circuit.

AWGN Additive White Gaussian Noise.

BE Binary Embedding.

BER Bit Error Rate.

BO Bayesian Optimisation.

BPTT Backpropagation Through Time.

CC Computational Complexity.

CD Chromatic Dispersion.

CDC Chromatic Dispersion Compensation.

CDR Clock and Data Recovery.

CMOS Complementary Metal-Oxide-Semiconductor.

CNN Convolutional Neural Network.

CO-OFDM Coherent Orthogonal Frequency Division Multiplexing.

CPR Carrier Phase Recovery.

CPU Central Processing Unit.

CVNN Complex-Valued Neural Network.

DARTS Differentiable Architecture Search.

DBP Digital Back-Propagation.

DFB Distributed Feedback.

DFE Decision Feedback Equalisation.

DP Dual-Polarisation.

DSP Digital Signal Processing.

EAM Electro-Absorption Modulator.

ECOC European Conference on Optical Communication.

EDFA Erbium-Doped Fibre Amplifier.

EGN Enhanced Gaussian Noise.

EI Expected Improvement.

EOTM End-of-Time Membrane.

FCNN Fully Connected Neural Network.

FEC Forward Error Correction.

FFE Feed-Forward Equalisation.

FFT Fast Fourier Transform.

FIR Finite Impulse Response.

FWM Four-Wave Mixing.

GP Gaussian Process.

GPU Graphics Processing Unit.

GRU Gated Recurrent Unit.

GVD Group Velocity Dispersion.

HD-FEC Hard-Decision Forward Error Correction Codes.

HPC High-Performance Computing.

IFFT Inverse Fast Fourier Transform.

IM/DD Intensity Modulation / Direct Detection.

ISI Inter-Symbol Interference.

LASSO Least Absolute Shrinkage and Selection Operator.

LEAF Large Effective Area Fibre.

LIF Leaky Integrate-and-Fire.

LLM Large Language Model.

LMMSE Linear Minimum Mean-Square Estimator.

LMS Least Mean Square.

LSTM Long Short-Term Memory.

MAC Multiply-Accumulate Operation.

ML Machine Learning.

MLP Multi-Layer Perceptron.

MOTM Max-over-Time Membrane.

MSE Mean Squared Error.

MZI Mach-Zehnder Interferometer.

MZM Mach-Zehnder Modulator.

NAS Neural Architecture Search.

NF Noise Figure.

NLIN Nonlinear Interference Noise.

NLSE Nonlinear Schrödinger Equation.

NN Neural Network.

OPC Optical Phase Conjugation.

OSNR Optical Signal-to-Noise Ratio.

PAM Pulse Amplitude Modulation.

PCTW Phase Conjugated Twin Waves.

PDL Polarisation-Dependent Loss.

PDM Polarisation Division Multiplexing.

PMD Polarisation Mode Dispersion.

PPLN Periodically Poled Lithium Niobate.

PRBS Pseudo-Random Binary Sequence.

QAM Quadrature Amplitude Modulation.

QPSK Quadrature Phase Shift Keying.

RAM Random Access Memory.

ReLU Rectified Linear Unit.

RLS Recursive Least Squares.

RMpS Real Multiplications per Symbol.

RNN Recurrent Neural Network.

RRC Root-Raised Cosine.

RVNN Real-Valued Neural Network.

RWKV Receptance Weighted Key Value.

SC Single-Carrier.

SE Spectral Efficiency.

SGD Stochastic Gradient Descent.

SMF Single-Mode Fibre.

SNN Spiking Neural Network.

SNR Signal-to-Noise Ratio.

SOA Semiconductor Optical Amplifier.

SOI Symbol of Interest.

SPM Self-Phase Modulation.

SRFFN Spiking Receptance Feed-Forward Network.

SRWKV Spiking Receptance Weighted Key Value.

SSFM Split-Step Fourier Method.

SSMF Standard Single-Mode Fibre.

STDP Spike-Timing Dependent Plasticity.

STpS 1 Span per Symbol.

Tanh Hyperbolic Tangent.

TIA Transimpedance Amplifier.

TTFS Time-To-First-Spike.

TWC True Wave Classic Fibre.

UCB Upper Confidence Bound.

VCSEL Vertical-Cavity Surface-Emitting Laser.

VSTF Volterra Series Transfer Function.

WDM Wavelength-Division Multiplexed.

XPM Cross-Phase Modulation.

ZFA Zero Forcing Algorithm.

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Increasing Demand for High-Speed, High-Capacity Optical Communication Systems

The rapid proliferation of digital technologies, coupled with the rise of computationally intensive applications such as Large Language Models (LLMs), advanced deep learning architectures, and cloud-based services, has ushered in an era of unprecedented data generation, processing, and transmission. These advancements have fundamentally reshaped the global technological landscape, with major corporations such as Google, OpenAI, and Meta relying on geographically distributed data centre clusters to meet the ever-growing computational demands [8–10].

The exponential increase in the size and complexity of Artificial Intelligence (AI) models serves as a primary driver of this transformation. State-of-the-art LLMs now consist of hundreds of billions of parameters and require training datasets spanning terabytes of data, with the energy and computational resources required for training being staggering [11, 12]. Beyond training, inference tasks demand immense computational power and low-latency data processing across consumer applications such as voice assistants, recommendation engines, and real-time translation [13]. Recent projections indicate that AI will drive a 165% increase in data centre power demand by 2030, with global data centre energy consumption expected to double due to AI applications [12, 14].

These computational demands are matched by equally rigorous requirements for data storage and transmission infrastructure. Modern AI workflows involve extensive data movement—from preprocessing datasets and transferring intermediate results during training, to synchronising distributed systems in real time. Such operations place extraordinary pressure on data centres and the networks interconnecting them, with AI-enabled data centres requiring over 10 times more optical fibre connections than traditional data centres [15, 16]. Data centres require significant energy not only to power processors—with next-generation AI chips consuming up to 1.2 kW compared to 150 W to 200 W for traditional Central Processing Units (CPUs)—but also to cool the vast arrays of servers that generate substantial heat during operation [17]. Moreover, inefficiencies in traditional electrical interconnects compound the overall energy footprint, highlighting the urgent need for more energy-efficient optical alternatives to sustain the scaling of AI technologies [18, 19].

As AI models grow in complexity and size, optical interconnects have become indispensable for supporting the massive data transfer requirements between remote data centres as well as within the data centre environment itself for rack-to-rack or even board-to-board interconnection. Emerging technologies targeting 800 Gb/s and beyond, with demonstrations of 1.6 Tb/s and 3.2 Tb/s coherent transmission systems, are under continuous development to satisfy these requirements [20,21]. The convergence of High-Performance Computing (HPC) and optical communications has escalated in importance, motivating research into novel methods for pushing the boundaries of achievable data rates and reliability while meeting the stringent energy efficiency requirements of modern AI infrastructure.

This AI-driven transformation coincides with the exponential growth of global internet traffic, which has surged from approximately 100 GB/d in 1992 to over 150 TB/s by 2022, with projections suggesting that this exponential growth will continue [22]. The dramatic increase in data traffic encompasses not only AI-driven applications but also bandwidth-intensive applications and emerging technologies like virtual and augmented reality, where ultra-reliable, high-speed networks are essential for maintaining performance and user satisfaction [20].

To address this rising demand, optical communication systems have seen substantial advancements over the years. In the 1990s, the introduction of Wavelength-Division Multiplexed (WDM) technologies, combined with EDFAs, significantly expanded system capacity, allowing multiple wavelength channels to be transmitted over a single fibre [23]. This breakthrough resulted in a 1000-fold increase in capacity within a decade, achieving transmission rates of 1 Tb/s by the year 2000, laying the foundation for addressing today's demands for high-speed, AI-driven data traffic [24].

More recently, research has concentrated on improving SE, which is defined as the number of bits transmitted per second per unit of bandwidth bit/s/Hz. This has been realised through the adoption of advanced modulation formats that exploit both polarisation states and multiple amplitude and phase levels of the optical carrier [3]. For example, the implementation of PDM with Quadrature Phase Shift Keying (QPSK) enabled the transmission of 100 Gb/s signals with a spectral efficiency of $2\,\mathrm{bit/s/Hz}$ [6]. Further developments in modulation techniques have led to even higher spectral efficiencies, with laboratory demonstrations reaching $11\,\mathrm{bit/s/Hz}$ using 128-ary Quadrature Amplitude Modulation (128-QAM) with PDM [4], and more recently, ultra-high-order QAM formats such as 4096-QAM achieving spectral efficiencies of up to $17.3\,\mathrm{bit/s/Hz}$ [5]. These high spectral efficiencies are critical for modern applications such as streaming, real-time AI inference, and large-scale data synchronisation, where maximizing data throughput within limited bandwidth is essential to meet escalating user demands and maintain network performance.

However, the pursuit of higher SE introduces several significant challenges. As the density of the signal constellation increases, a higher SNR becomes necessary to maintain acceptable

BER [24]. This requirement makes it necessary increases in signal launch power, which in turn exacerbates nonlinear effects in the optical fibre, particularly those related to Kerr nonlinearity [25]. The interaction between elevated launch powers and nonlinear distortions imposes a fundamental limit on the achievable information capacity of optical fibre channels, often referred to as the "nonlinear Shannon limit" [24, 26].

To address these capacity limitations, researchers have investigated several approaches:

1. **Expanding the usable optical bandwidth:** Efforts have been directed towards utilising multiple transmission bands (O, E, S, C, L, and U) in optical fibres. For instance, recent demonstrations have achieved the transmission of 54 Tb/s over 9,150 km using the C+L bands with hybrid Raman-EDFA amplification [27]. However, the inclusion of additional bands introduces new challenges, such as the need for more sophisticated component designs and the potential for increased nonlinear interactions [28–30].

2. **Improving amplification schemes:** Advanced amplification techniques, such as distributed Raman amplification, have been employed to reduce amplifier Noise Figures (NF) and enhance the SNR without significantly increasing launch powers [31].

3. **Nonlinearity compensation techniques:** Various methods have been proposed to mitigate or compensate for nonlinear effects, including DBP [32, 33], OPC [34], and machine learning-based approaches [35].

4. **Space-Division Multiplexing (SDM):** Research into multi-core and multi-mode fibres has shown promise in overcoming the capacity limitations of single-mode fibres by leverageing the spatial dimension for parallel transmission [36].

While all four strategies are critical to addressing capacity limitations, they each present distinct challenges and opportunities. Expanding bandwidth (solution 1) and increasing spatial dimensions (solution 4) can be used simultaneously to scale capacity but come with higher infrastructure complexity. Amplification improvements (solution 2) are essential but do not address the computational and energy costs associated with manageing the nonlinearities. Among these solutions, nonlinearity compensation (solution 3) offers a distinct advantage as it complements all the others. However, traditional techniques like DBP are computationally intensive and power-hungry, making them challenging to scale in energy-efficient systems [37].

Despite these technological advances, understanding the fundamental physical limits of optical fibre transmission is crucial for contextualizing the challenges that modern optical networks face. These limitations will be examined in detail in the context of nonlinear impairments and their compensation techniques in Section 1.2.3.

The convergence of exponential AI growth and massive data transmission requirements necessitates sophisticated signal processing techniques that can mitigate nonlinear impairments in optical communication systems. Modern optical networks must simultaneously address the capacity limitations imposed by fibre nonlinearities while supporting the ultra-high bandwidth and low-latency requirements of distributed AI systems.

Emerging photonic neuromorphic computing paradigms offer promising solutions where optical hardware mimics neural architectures to achieve high-speed, energy-efficient computations. These approaches, including reservoir computing and optical neural networks, can process data in the analogue domain, bypassing traditional digital-electronic bottlenecks for tasks such as channel equalisation and real-time fault detection [38–41].

Addressing these challenges requires a multidisciplinary approach that bridges computational intelligence and photonic engineering to extend the performance boundaries of optical communication systems beyond current limitations [37, 38, 41, 42].

To understand how these advanced signal processing techniques can be effectively integrated into optical networks, it is essential to first establish a comprehensive understanding of the fundamental principles governing optical communication systems and the specific impairments they introduce.

1.2 Fundamentals of Optical Communication Systems

1.2.1 Fibre Principles and Key Parameters

Optical fibre communication systems have transformed the telecommunications industry, enabling the transmission of vast amounts of data at high speeds over long distances. These systems exploit the unique properties of light propagation in optical fibres to achieve unprecedented data capacity with minimal errors [7]. Understanding the fundamental principles of optical fibre transmission requires examination of both the physical structure of fibres and the key parameters that govern their performance.

The basic component of every optical communication system is the optical fibre itself. A typical optical fibre consists of three main structural elements [7]: a core region made from silica (SiO_2) that is doped with materials such as germanium to increase its refractive index, a cladding layer surrounding the core composed of pure silica with a lower refractive index, and an outer protective coating typically made of polymer materials. The difference in refractive indices between the core and cladding facilitates total internal reflection, a phenomenon that confines light within the core and allows it to propagate along the length of the fibre. This behaviour is governed by Snell's law [7]:

$$n_1 \sin\theta_1 = n_2 \sin\theta_2 \tag{1.1}$$

where n_1 and n_2 represent the refractive indices of the core and cladding, respectively, and θ_1 and θ_2 denote the angles of incidence and refraction.

The propagation of light inside an optical fibre can be described by solving Maxwell's equations, which completely define electromagnetic wave behaviour [43, 44]. In the case of weakly guiding fibres, where the refractive-index difference between core and cladding is small, this can be approximated by the wave equation:

$$\nabla^2 E + n^2(r) k_0^2 E = 0 \tag{1.2}$$

where E is the electric field, $n(r)$ is the refractive index profile, $k_0 = 2\pi/\lambda$ is the wave number in vacuum, and λ is the wavelength of light. The solutions to this equation yield the modes of propagation in the fibre, with the fundamental LP_{01} mode being the primary mode of interest in single-mode fibres [43].

Optical fibres are broadly classified into two categories based on their light propagation characteristics: Single-Mode Fibre (SMF) with small core diameters (typically 8-10 µm) that support only one mode of light propagation, and multi-mode fibres (MMF) with larger core diameters (typically 50-62.5 µm) that support multiple modes. SMFs are widely used in long-distance, high-speed communications due to their low dispersion and attenuation characteristics, while MMFs are often used in short-distance applications due to their ease of alignment with light sources and lower cost [7].

The performance of optical fibres in communication systems is governed by several critical parameters that determine their suitability for different applications. These parameters can be broadly categorised into linear and nonlinear effects, each playing a crucial role in determining the overall system performance.

Among the linear effects, attenuation represents the optical power loss during propagation and is typically expressed in dB/km as a wavelength-dependent quantity. The attenuation coefficient α is derived from the relationship between input and output powers: $\alpha = -10\log_{10}(P_{out}/P_{in})/L$, where L is the fibre length and P_{in} and P_{out} are the input and output powers, respectively.

Chromatic dispersion represents another fundamental linear effect that significantly impacts system performance. This phenomenon arises from the wavelength-dependent variation in the refractive index of the fibre material, combined with waveguide dispersion effects. The group velocity, defined as the speed at which the envelope of an optical pulse travels along the fibre, varies with wavelength, causing different spectral components of a pulse to travel at different speeds and leading to pulse broadening. This effect is quantified by the Group Velocity Dispersion (GVD) parameter $\beta_2 = -\lambda^2(d^2n/d\lambda^2)/(2\pi c)$, where c is the speed of light in vacuum, and n is the refractive index [45]. When β_2 is non-zero, different spectral components of a pulse travel at different speeds, leading to pulse broadening. Positive β_2 (normal dispersion) means longer wavelengths travel faster than shorter ones, while negative β_2 (anomalous dispersion) means shorter wavelengths travel faster. At 1550 nm, standard single-mode fibres typically exhibit negative β_2, corresponding to anomalous dispersion. This property is commonly quantified using the dispersion parameter $D = -2\pi c\beta_2/\lambda^2$ [45].

Typical values for SSMF are around 17 ps/(nm km) at 1550 nm wavelength.

Polarisation-related effects constitute another important class of linear impairments. Polarisation Mode Dispersion (PMD) arises from the difference in propagation speeds between different polarisation states of light, leading to pulse broadening. These impairments are a consequence of imperfections in the fibre's structure, such as asymmetry in the core or external stresses, which create different refractive indices for different polarisation states (birefringence). PMD is quantified by the PMD coefficient D_p, measured in $\mathrm{ps}/\sqrt{\mathrm{km}}$ [45, 46]. The first-order PMD, which is the primary component of this effect, is characterised by the Differential Group Delay (DGD) $\Delta\tau = D_p\sqrt{L}$, where $\Delta\tau$ represents the DGD, D_p is the PMD coefficient, and L is the fibre length. Lower values of D_p indicate better performance in terms of PMD.

Closely related to PMD is Polarisation-Dependent Loss (PDL), which represents the variation in attenuation that different polarisation modes experience as they propagate through optical components. One polarisation of light might be absorbed or scattered more than another, leading to an imbalance in the power of the light waves. This phenomenon is particularly detrimental in systems that use polarisation to carry information, such as in high-speed polarisation division multiplexed communications, as optical components do not affect all polarisation states equally [47].

The nonlinear properties of optical fibres become increasingly significant at high optical powers and fundamentally limit the capacity of communication systems. The nonlinear refractive index, denoted as n_2, is an intensity-dependent component of the refractive index arising from the Kerr effect, where the refractive index increases with the intensity of the light passing through it [45]. This phenomenon is responsible for various nonlinear effects in optical fibres, including SPM, XPM, and FWM, which can cause significant signal distortion particularly in high optical power and long-distance communication systems. SPM results in a phase shift proportional to the optical intensity, leading to spectral broadening of the pulse. XPM occurs when the phase of one optical signal is modulated by the intensity of another co-propagating signal, leading to crosstalk between channels in wavelength-division multiplexed systems. FWM is a nonlinear process where two or more optical signals interact to generate new frequency components, causing interference and signal quality degradation. A detailed analysis of these effects is provided in Section 1.2.3.

The magnitude of these nonlinear effects is inversely proportional to the effective area A_{eff} of the fibre, which represents the area over which the optical power is distributed in the fibre cross-section. Fibres with larger A_{eff} generally exhibit reduced nonlinear distortions, allowing for higher power transmission or longer distances before nonlinear effects become significant. However, this comes with trade-offs, as larger A_{eff} fibres are more susceptible to bending losses and, in few-mode or multimode fibres, can exacerbate mode coupling, leading to modal dispersion and signal degradation. Therefore, large-A_{eff} fibres are often preferred for long-haul, high-power transmission systems to mitigate nonlinear distortions, while smaller A_{eff} may be used in applications where nonlinear effects are less critical [7].

Optical fibre transmission typically occurs in specific wavelength ranges known as transmission windows, each characterised by distinct attenuation and dispersion properties [7, 48]. These windows include the O-band (Original, 1260-1360 nm), E-band (Extended, 1360-1460 nm), S-band

(Short, 1460-1530 nm), C-band (Conventional, 1530-1565 nm), L-band (Long, 1565-1625 nm), and U-band (Ultra-long, 1625-1675 nm). The C-band and L-band are most commonly used in long-haul transmission systems due to their low attenuation characteristics and the availability of efficient optical amplifiers, particularly EDFAs, in these ranges [48].As illustrated in Figure 1.1, the attenuation varies significantly across different bands, with the C-band exhibiting minimum attenuation at approximately 1550 nm (0.15 dB/km). This wavelength region coincides with the gain spectrum of EDFAs, highlighted in the figure, which explains its dominance in commercial optical networks. The attenuation increases in both the O-band (shorter wavelengths) and U-band (longer wavelengths), limiting their use in ultra-long-haul applications without additional amplification technologies.

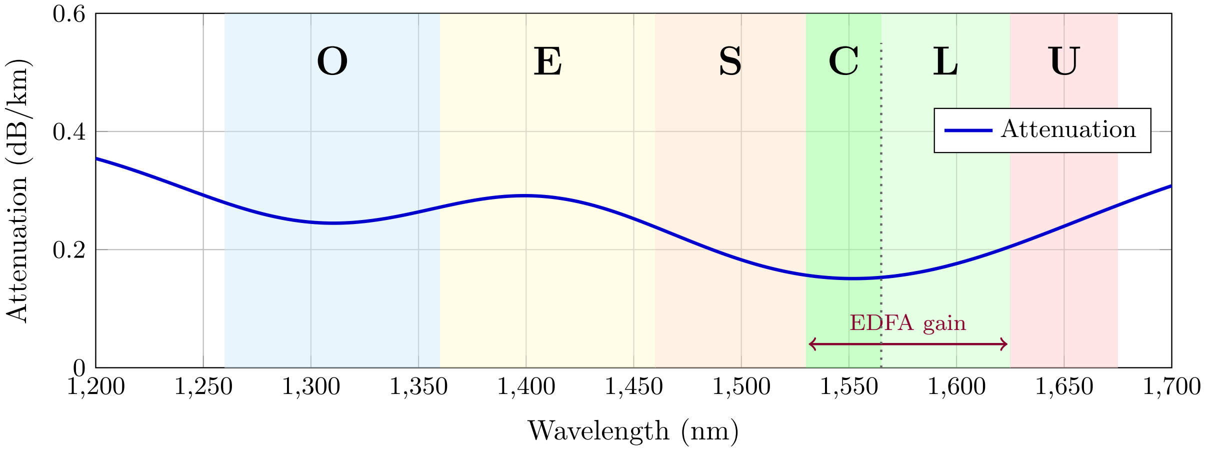


Figure 1.1: Optical communication bands (O, E, S, C, L, U) showing wavelength ranges and attenuation characteristics. The C-band region exhibits the lowest attenuation (approximately 0.15 dB/km) and overlaps with the EDFA gain window, making it ideal for long-haul transmission.

1.2.2 System Components and Modulation Formats

Optical communication systems comprise several essential components that work together to enable high-speed, long-distance data transmission through sophisticated modulation techniques. The fundamental architecture consists of transmitters, optical fibres, amplifiers, and receivers, each playing a crucial role in the overall system performance [49].

The transmitter serves as the interface between electrical and optical domains, converting electrical signals into optical signals for fibre transmission. The primary component is the light source, typically a semiconductor laser such as a Distributed Feedback (DFB) laser or Vertical-Cavity Surface-Emitting Laser (VCSEL). The optical carrier generated by this laser is modulated with data through various techniques depending on the required data rate and modulation format. For lower data rates, direct modulation of the laser current is employed, while higher speeds and complex modulation formats require external modulators such as Mach-Zehnder Modulator (MZM) or Electro-Absorption Modulator (EAM) [50]. Driver circuits provide the necessary electrical signals to control both the laser and modulator, while wavelength control mechanisms, including temperature control and wavelength-locking techniques, maintain the laser's output wavelength within the desired range [7, 29, 49].

The choice of modulation format fundamentally affects the system's SE, reach, and overall performance. Modern optical communication systems employ various modulation techniques that encode information onto different properties of the optical field, including amplitude, phase, frequency, and polarisation [3]. The spectral efficiency, defined as $SE = R_b/B = \log_2(M)/(T_s B)$ [bits/s/Hz], where R_b is the bit rate, B is the signal bandwidth, M is the number of constellation points, and T_s is the symbol period, serves as a critical performance metric [7].

QAM represents one of the most widely used modulation formats, combining both amplitude and phase modulation to encode information efficiently. In M-QAM, the complex-valued signal can be represented as $s_k = A_k e^{i\phi_k}$, where A_k and ϕ_k are the amplitude and phase of the k-th symbol, respectively. The number of bits per symbol is given by $\log_2(M)$, making QAM highly efficient for encoding large amounts of data. High-order QAM formats (64-QAM, 256-QAM, and beyond) require sophisticated hardware including dual-polarisation IQ modulators with precise bias control, high-linearity driver amplifiers, and advanced digital signal processing for pre-distortion and linearisation [3].

Phase-Shift Keying (PSK) provides another important modulation approach, where information is encoded solely in the phase of the carrier. Binary PSK (BPSK) and QPSK can be implemented using simpler hardware configurations, typically requiring single or dual MZM with appropriate phase control. Higher-order PSK formats benefit from integrated modulators such as nested MZMs or ring modulators for enhanced phase accuracy and reduced hardware complexity [49].

To further enhance data capacity, these modulation formats are often combined with PDM, which exploits the polarisation of light to transmit multiple signals simultaneously over the same optical fibre, effectively doubling the data capacity. PDM implementations require polarisation beam splitters, dual-polarisation modulators, and coherent receivers with polarisation diversity to manage the orthogonal polarisation states [51].

The hardware complexity for generating different modulation formats varies significantly with the required spectral efficiency and transmission distance. Simple On-Off Keying (OOK) can be implemented using direct laser modulation or a DFB laser with an EAM, requiring only basic driver electronics and offering a straightforward, cost-effective solution for short-reach applications [7]. Phase-Shift Keying formats (BPSK, QPSK) demand lasers with narrow linewidth ($<$100 kHz) to minimise phase noise, MZM—one for BPSK or a dual-drive configuration for QPSK—along with RF amplifiers and precise bias control circuits to maintain optimal operating points [3]. High-order QAM formats (64-QAM, 256-QAM, and beyond) impose stringent requirements: ultra-narrow linewidth lasers ($<$10 kHz) to ensure phase coherence over extended transmission distances, integrated dual-polarisation IQ modulators containing four nested MZMs for independent control of in-phase and quadrature components on both polarisations, wideband high-linearity RF drivers ($>$25 GHz bandwidth) to preserve signal fidelity, sophisticated bias stabilisation with feedback mechanisms to compensate for temperature and ageing effects, and pre-compensation DSP to linearise the transmitter response and reduce constellation distortions [3, 4]. Polarisation Division Multiplexing further requires polarisation beam splitters to separate orthogonal polarisation states, combiners to multiplex them onto the fibre, and polarisation-maintaining fibre components throughout the signal path to independently modulate and preserve the integrity of each polarisation state [51].

Table 1.1 provides a comprehensive comparison of commonly used modulation formats, showing their spectral efficiency characteristics, required OSNR, primary nonlinear impairments, and typical application scenarios. This comparison illustrates the fundamental trade-off between spectral efficiency and robustness against nonlinear effects, where higher-order modulation formats achieve greater data rates but require higher OSNR and exhibit increased sensitivity to Kerr-effect-induced distortions.

Table 1.1: Comparison of modulation formats: spectral efficiency, hardware requirements, nonlinear impairments, and typical applications in optical communication systems. SE: Spectral Efficiency (single polarisation, doubled for PDM); OSNR: Required optical signal-to-noise ratio at BER $= 10^{-3}$ (pre-FEC); SPM: Self-Phase Modulation; XPM: Cross-Phase Modulation; FWM: Four-Wave Mixing [3–7].

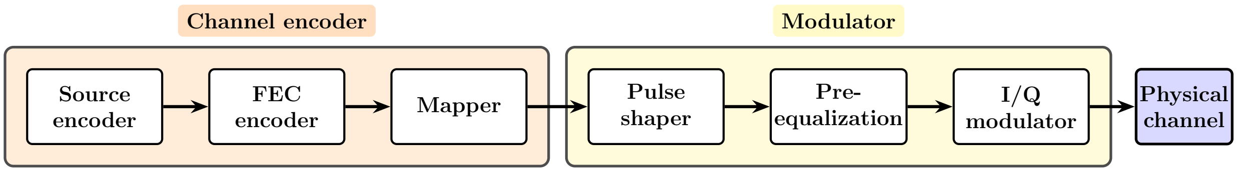
Format	**Bits/ Symb**	**SE (b/s/Hz)**	**OSNR (dB)**	**Primary Nonlinear Impairments**	**Kerr Sens./ Applications**
OOK	1	1	12–15	SPM, XPM in WDM	Low / Short-reach
BPSK	1	1	10–13	SPM (phase dist.)	Low–Med. / Long-haul
QPSK	2	2	13–16	SPM, XPM, phase noise	Med. / Long-haul, 100G
PDM-QPSK	4	4	13–16	SPM, XPM, nonlin. phase	Med. / Submarine
16-QAM	4	4	18–21	SPM, XPM, const. warp.	Med.–High / Metro
PDM-16-QAM	8	8	18–21	SPM, XPM, FWM, amp./phase	High / 400G, metro
64-QAM	6	6	23–26	Severe SPM/XPM, collapse	High / Datacenter
PDM-64-QAM	12	12	23–26	Severe crosstalk, tight spac.	Very High / 800G/1.6T
256-QAM	8	8	28–32	Extreme Kerr sensitivity	Very High / Lab demos
PDM-256-QAM	16	16	28–32	Catastrophic at high power	Extreme / Research

Optical amplifiers play a crucial role in compensating for fibre losses and enabling the implementation of multi-span systems. The two main types of optical amplifiers used in fibre-optic communications are EDFAs and Raman Amplifiers [49]. EDFAs utilise erbium-doped fibre as the gain medium and are typically operated in the C-band (1530-1565 nm) and L-band (1565-1625 nm), with gain described by $G = \exp(g_0 L)$, where g_0 represents the small-signal gain coefficient and L is the length of the doped fibre [48]. Raman Amplifiers exploit stimulated Raman scattering within the transmission fibre itself to provide distributed amplification, which enhances the signal-to-noise ratio of the system. Both amplifier types introduce Amplified Spontaneous Emission (ASE) noise, which accumulates along the transmission link and can ultimately limit overall system performance [49].

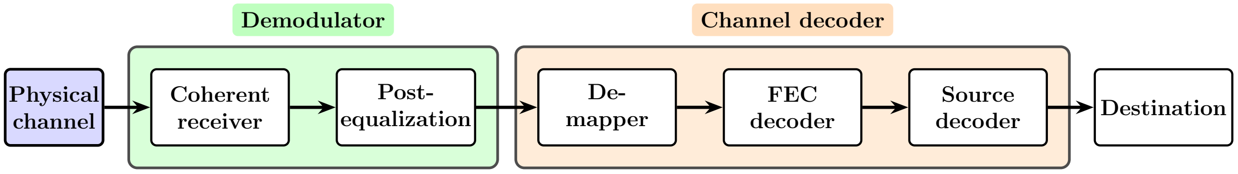
At the receiver end, the optical signal is converted back into an electrical signal for further processing. Modern coherent receivers comprise several key components including the photodetector, Transimpedance Amplifier (TIA), Clock and Data Recovery (CDR) circuits, Analog-to-Digital Converters (ADCs), and the DSP unit [29, 37, 49]. The photodetector converts optical power into electrical current with photocurrent $I_p = RP_{opt}$, where R is the responsivity and P_{opt} is the incident optical power. The TIA converts the photocurrent into a voltage signal and provides initial amplification, while the CDR circuit extracts timing information to ensure optimal data sampling. ADCs digitize the received signal for advanced DSP functions such as chromatic dispersion compensation, polarisation demultiplexing, carrier recovery, and nonlinearity compensation.

System performance is characterised by key metrics including the BER and Q-factor. The BER, defined as the probability of an error occurring in the reception of a digital bit, can be computed as BER $= \frac{1}{2}\text{erfc}(Q/\sqrt{2})$, where Q is the Q-factor representing the signal-to-noise ratio at the decision circuit input, defined as $Q = |\mu_1 - \mu_0|/(\sigma_1 + \sigma_0)$, with μ_1 and μ_0 being the mean values of the received signal for bits '1' and '0', and σ_1 and σ_0 their respective standard deviations [29, 52].

While the system components and modulation formats discussed above enable high-capacity transmission, their performance is ultimately limited by various impairments that occur during signal propagation through optical fibres. The most significant of these limitations arise from nonlinear effects that become increasingly pronounced as systems push toward higher data rates and spectral efficiencies.



(a) Transmitter (channel encoder)



(b) Receiver (channel decoder)

Figure 1.2: Coherent optical communication system architecture showing the complete signal processing chain from source to destination.

Figure 1.2 illustrates the complete architecture of a coherent optical communication system, highlighting the critical signal processing stages that enable high-capacity transmission. As shown in Figure 1.2(a), the transmitter chain begins with source encoding for data compression, followed by FEC encoding that adds redundancy to combat channel impairments. The mapper converts binary data into complex symbols according to the chosen modulation format (e.g., QAM or PSK), while the pulse shaper controls the signal's spectral characteristics to optimise bandwidth utilisation. Pre-equalisation compensates for known channel distortions, and the in-phase/quadrature (I/Q) modulator generates the optical signal by combining in-phase and quadrature components, encoding both amplitude and phase information onto the optical carrier. After transmission through the physical channel, Figure 1.2(b) shows how the coherent receiver uses a local oscillator to detect both amplitude and phase of the incoming signal, enabling recovery of the transmitted data. The post-equalisation stage, typically implemented through advanced DSP algorithms, compensates for accumulated linear and nonlinear impairments including chromatic dispersion, polarisation mode dispersion, and fibre nonlinearities. Finally, the de-mapper, FEC decoder, and source decoder complete the recovery of the transmitted information, delivering it to the destination.

1.2.3 Nonlinear Impairments and Compensation Techniques

The pursuit of higher data rates and spectral efficiency in optical communication systems has led to nonlinear impairments in optical fibres becoming a fundamental limitation. These effects become increasingly significant as signal power levels are increased to achieve the higher signal-to-noise ratio required for advanced modulation formats [25].

The primary source of nonlinear impairments is the Kerr effect, which causes the refractive index of the fibre to vary with the intensity of the propagating optical field. This intensity-dependent refractive index, expressed as $n(I) = n_0 + n_2 I$ where n_0 is the linear refractive index, n_2 is the nonlinear-index coefficient (typically around 2.6×10^{-20} m^2/W for silica fibres), and I is the optical intensity, becomes significant due to the high optical intensities achieved in the small core of optical fibres and the long interaction lengths enabled by low fibre losses [25, 37]. The accumulated nonlinear phase shift over a fibre length L is given by $\phi_{NL} = \gamma P L_{eff}$, where $\gamma = 2\pi n_2/(\lambda A_{eff})$ is the nonlinear coefficient, P is the optical power, and $L_{eff} = [1 - \exp(-\alpha L)]/\alpha$ is the effective length [49].

The Kerr effect manifests through three primary nonlinear phenomena that significantly impact signal transmission. SPM occurs when the phase of an optical pulse is modulated by its own intensity, experiencing a time-dependent phase shift $\phi_{SPM}(t) = -\gamma P(t) L_{eff}$ that leads to spectral broadening and potential intersymbol interference when combined with dispersion. XPM becomes significant in wavelength division multiplexed systems, where the intensity of one channel modulates the phase of co-propagating channels according to $\phi_{XPM,j} = 2\gamma L_{eff} \sum_{k \neq j} P_k$, with the factor of 2 indicating that XPM is twice as effective as SPM for the same power level, resulting in timing jitter and amplitude distortions particularly in systems with tight channel spacing. FWM represents a nonlinear process where three optical waves at frequencies f_1, f_2, and f_3 interact to generate a fourth wave at frequency $f_4 = f_1 \pm f_2 \pm f_3$, with power given by $P_{FWM} = (\gamma L_{eff})^2 \eta P_1 P_2 P_3 \exp(-\alpha L)$, where η is the FWM efficiency factor dependent on phase mismatch and channel spacing [25, 44, 53, 54].

The complex interaction between dispersion and nonlinearity in optical fibre transmission can be described by the Nonlinear Schrödinger Equation (NLSE):

$$\frac{\partial A}{\partial z} + \frac{\alpha}{2} A + i \frac{\beta_2}{2} \frac{\partial^2 A}{\partial T^2} - i\gamma |A|^2 A = 0 \tag{1.3}$$

where A is the slowly varying envelope of the optical field, z is the propagation distance, T is the time in a moving frame, α is the attenuation coefficient, β_2 is the GVD parameter, and $\gamma = 2\pi n_2/(\lambda A_{eff})$ is the nonlinear coefficient. This equation captures the interplay between various linear and nonlinear effects, with the interaction governed by two characteristic length scales: the dispersion length $L_D = T_0^2/|\beta_2|$ and the nonlinear length $L_{NL} = 1/(\gamma P_0)$, where their relative magnitudes determine whether the system operates in a dispersion-dominated ($L_D \ll L_{NL}$), nonlinearity-dominated ($L_D \gg L_{NL}$), or strongly interacting regime ($L_D \approx L_{NL}$) [25, 45, 52, 55].

These nonlinear impairments impose several detrimental effects on optical communication systems, including reduced transmission reach where nonlinear effects limit the maximum distance for signal transmission without regeneration, constrained spectral efficiency through the nonlinear Shannon limit that imposes an upper bound on achievable spectral efficiency particularly for long-haul systems, increased system complexity and power consumption due to sophisticated signal processing requirements, limited network flexibility in dynamically allocating resources, signal distortion causing amplitude and phase distortions that increase bit error rates, spectral broadening from SPM and XPM leading to inter-channel interference, and crosstalk from FWM-generated frequency components interfering with existing channels [20, 24, 32, 37, 52, 56, 57].

The relationship between the system capacity and these nonlinear effects can be quantified using the nonlinear Shannon limit, which provides an upper bound on the SE as a function of SNR: $C = 2\log_2(1 + \text{SNR}/(1 + \eta P^2))$, where η is a parameter that depends on system and fibre properties, and P is the signal power [26, 58]. From this equation it is easy to understand that increasing the signal power improves the SNR, but also enhances nonlinear effects, ultimately leading to a capacity limit.

This fundamental limitation manifests as what is commonly referred to as the "capacity crunch" [59, 60], representing the convergence of exponentially growing data demands with the fundamental capacity limits imposed by fibre nonlinearities. The nonlinear Shannon limit arises from the interaction between the Kerr nonlinearity in optical fibres and the quantum noise from optical amplifiers, creating a fundamental barrier where increasing signal power to improve signal-to-noise ratio simultaneously increases nonlinear distortion [24].

The theoretical foundation of this limit can be understood through the nonlinear Schrödinger equation presented in Equation 1.3. As signal power increases, the nonlinear phase shift becomes significant, leading to spectral broadening and inter-channel crosstalk in wavelength-division multiplexed systems. Recent research has shown that even with ideal distributed Raman amplification and optimised fibre parameters, the nonlinear Shannon limit remains relatively insensitive to changes in fibre loss or nonlinearity coefficient [61]. This fundamental limitation has profound implications for future network capacity, as Internet traffic continues to grow at approximately

40% per year, with projections indicating that current bandwidth scaling strategies may become insufficient within the next decade [60].

To address these limitations, several advanced signal processing techniques have been developed to mitigate nonlinear impairments and push system performance closer to fundamental limits. DBP attempts to reverse fibre nonlinearities by numerically solving the inverse nonlinear Schrödinger equation, providing significant performance improvements but with computational complexity that scales with transmission distance and signal bandwidth, making real-time implementation challenging. Volterra Series Transfer Function (VSTF) approaches model the fibre channel as a nonlinear system with memory, using mathematical expansions to compensate for both linear and nonlinear impairments while offering a good trade-off between performance and complexity. Perturbation-based nonlinear compensation uses first-order perturbation approaches to estimate and compensate for nonlinear distortions, offering lower complexity than DBP but with reduced effectiveness in strong nonlinear regimes. OPC provides an all-optical signal processing technique that compensates for both chromatic dispersion and nonlinear effects by creating a phase-conjugated copy of the signal at the transmission link midpoint [62–65].

Emerging approaches to overcome capacity limitations include nonlinear frequency division multiplexing (NFDM), which exploits the integrability of the nonlinear Schrödinger equation to transmit data in the nonlinear domain without suffering degradation from increased power or transmission distance. Machine learning-based approaches have shown particular promise in addressing the stochastic nature of nonlinear distortions, with deep neural networks demonstrating the ability to compensate for nonlinear inter-carrier crosstalk effects and adapt to changing channel conditions. Neuromorphic computing architectures, inspired by biological neural systems, offer new possibilities for real-time, energy-efficient nonlinearity compensation by processing high-dimensional optical signals with low latency and power consumption, making them attractive for next-generation optical networks [35, 38, 66–69].

Table 1.2 provides a comprehensive comparison of state-of-the-art digital equalisation techniques for nonlinearity compensation in optical communication systems. This comparison highlights the fundamental trade-offs between computational complexity, performance gains measured in Q-factor improvement, and implementation feasibility. Linear equalisation techniques serve as the baseline for chromatic dispersion and PMD compensation but offer no nonlinearity mitigation. Advanced techniques such as DBP and machine learning approaches provide significant performance improvements but at the cost of increased computational requirements, while emerging optical methods like OPC offer the potential for all-optical processing with reduced latency.

Table 1.2: State-of-the-art digital equalisation techniques for nonlinearity compensation in optical communication systems. **Complexity notation:** N = samples; M = DBP steps/span; K = CD taps; H = neurons; L = layers; F = CNN filters; K_{CNN} = kernel size; S = spike events. **Q-gain:** Performance improvement vs. linear baseline (dB Q-factor, ~1000 km, PDM-16QAM). **Status:** Commercial = deployed; Field trials = operational tests; Research = lab demos; Early research = proof-of-concept [1, 32–34, 37, 38, 42, 58, 66, 67].

Technique	Operating Principle	Complexity	Q-gain	Status
Linear Equalisation				
FFE/DFE	Time-domain filtering (CD/PMD)	$O(N)$	0 (baseline)	Commercial
Freq-domain EQ	FFE frequency-domain CD compensation	$O(N \log N)$	0 (baseline)	Commercial
Digital Nonlinearity Compensation				
DBP (full)	Inverse NLSE via SSFM	$O(NMK)$	1.5–3.5 dB	Research
DBP (single-step)	Simplified back-propagation	$O(NK)$	0.8–1.5 dB	Field trials
Perturbation-based	First-order approximation	$O(N)$	0.3–1.2 dB	Research
Volterra (VSTF)	2nd/3rd-order expansion	$O(N^2)$–$O(N^3)$	0.5–2.0 dB	Research
Machine Learning Approaches				
NN-Equaliser (MLP)	Feedforward neural network	$O(NHL)$	1.0–2.5 dB	Research
NN-Equaliser (CNN)	Convolutional filters	$O(NFK_{CNN})$	1.2–3.0 dB	Research
NN-Equaliser (RNN)	Recurrent memory	$O(NHL)$	1.5–3.2 dB	Research
Hybrid (OPC+NN)	Optical + neural refinement	$O(NH)$	2.5–4.5 dB	Lab demo
Optical Processing				
OPC (mid-span)	Spectral inversion	Optical	2.0–4.0 dB	Field trials
Phase-sens. amp.	Optical phase squeezing	Optical	1.0–2.5 dB	Research
Emerging Techniques				
NFDM	Nonlinear Fourier transform	$O(N \log N)$	Variable	Research
Neuromorphic (SNN)	Event-driven spiking	$O(NS)$	1.0–2.5 dB	Early research

1.3 Overview of Machine Learning in Optical Communication Systems

Machine Learning and particularly ANNs have emerged as a powerful tool for addressing the complex challenges in modern optical communication systems, particularly in the realm of nonlinear impairments compensation. As larger data rates are needed and transmission distances extend, the optical fibre channel introduces increasingly significant nonlinear distortions that conventional DSP techniques struggle to mitigate effectively. This has led to a growing interest in leverageing the capabilities of ANNs to enhance the performance and efficiency of optical communication links [68].

Traditional DSP in optical systems follows a modular approach, with distinct algorithms addressing specific types of distortions at the transmitter and receiver [33, 37, 70]. However, as transmission rates increase significantly, the limitations of this conventional approach become apparent, particularly in handling the complex interplay of linear and nonlinear impairments [52]. The nonlinear effects in optical fibres (SPM, XPM, and FWM) interact with linear impairments like chromatic dispersion and polarisation mode dispersion, creating a highly complex transmission environment [29]. ANNs, with their ability to model and adapt to intricate, nonlinear relationships, offer a promising solution to this multifaceted problem [68].

The application of ANNs in optical communications, however, is not without its challenges. Unlike many common machine learning tasks, optical systems require extremely low bit error rates, typically around 10^{-2} before forward error correction. This demand for high precision, coupled with the need for real-time operation in ultrafast photonic systems, imposes strict requirements on ANN implementations in terms of throughput, energy efficiency, and latency [37, 38, 68]. To address these unique demands, researchers have developed various strategies to adapt ANNs for optical communication tasks. These strategies can be broadly categorised into two main groups: those aimed at optimising traditional ANN architectures through compression techniques such as pruning, weight clustering, knowledge distillation, and quantisation, and those exploring emerging computational paradigms such as neuromorphic computing. The first group focuses on reducing computational demands without significantly compromising performance, while the second group pushes the boundaries of traditional models by mimicking biological neural systems [37, 38, 52].

One of the main challenges of taking these approaches into real-world implementations is their computational complexity. Compression techniques have become crucial for making ANNs more suitable for real-time implementation in optical systems, while neuromorphic computing approaches, particularly SNNs, offer potential advantages in terms of energy efficiency and processing speed through event-driven processing [38,42,52]. Additionally, hybrid approaches that combine traditional ANNs with other solutions like optical phase conjugation further enhance capabilities [1]. The pursuit of more efficient implementations has led to the emerging field of neuromorphic photonics, which aims to overcome electronic hardware limitations by implementing ANNs using photonic components. Photonic Neural Networks (PNNs) offer compelling advantages including enhanced processing speed, reduced latency, and lower power consumption, with the ability to process optical signals directly in the optical domain [38, 71]. However, these systems also face unique challenges, particularly dealing with analogue noise in optical components, which requires developing robust architectures and algorithms to maintain high performance [71–73].

The integration of ANNs, particularly neuromorphic implementations, into optical communication systems represents a significant shift in approach to signal processing and nonlinearity compensation. These technologies hold the potential to dramatically improve the performance and efficiency of next-generation optical networks [38]. The detailed technical implementations, architectures, and performance evaluations of these approaches are explored extensively in Chapters 2-5 of this thesis.

1.4 Contributions of the Thesis

The field of optical communication has seen tremendous growth in recent years, driven by the ever-increasing demand for high-speed, high-capacity data transmission. As traditional signal processing techniques approach their limits in compensating for channel impairments and nonlinearities, researchers have turned to advanced machine learning methods, particularly neural networks, as a promising solution. While neural network-based equalisers have shown potential in numerical simulations and experimental demonstrations, their full capabilities and limitations in practical, industrial-scale applications remain largely unexplored.

This thesis aims to address this critical gap in knowledge by conducting a comprehensive, large-scale investigation into neural network-based equalisers for coherent fibre-optic communication systems. Our work spans from the development of novel architectures inspired by the physics of optical systems to the evaluation of black-box solutions, and from enhancing operational flexibility to optimising performance-complexity trade-offs. Furthermore, we delve into the challenges of implementing these advanced equalisers on resource-constrained hardware, a crucial step towards their adoption in next-generation digital signal processing systems. The contributions of this thesis represent a significant step forward in the field, offering insights and solutions that could pave the way for the widespread industrial application of neural network-based equalisers in optical communications. Our findings not only advance the theoretical understanding of these systems but also provide practical approaches to overcome the hurdles in their real-world implementation.

The key contributions of this thesis can be summarised as follows:

- This thesis presents a comprehensive large-scale investigation of neural network-based equalisers in coherent fibre-optic communication systems, addressing the gap between promising experimental results and industrial application.
- We propose novel physical-inspired neural network architectures specifically designed for optical channel equalisation, including complex-valued neural networks and adaptations of

advanced architectures such as spiking neural networks and transformers.

- A thorough investigation of various “black box” neural network solutions is conducted, evaluating their potential for optical signal equalisation and comparing their performance with traditional methods.
- This work provides an extensive performance versus complexity trade-off analysis for a wide range of neural network architectures and configurations, including feed-forward, recurrent, and advanced models.
- We introduce and evaluate novel techniques for reducing the computational complexity of neural network equalisers, including pruning strategies, quantisation techniques, and weight clustering approaches, making them more attractive for industrial implementation.
- The thesis presents a pioneering study on the synergy between optical phase conjugation and neural network-based equalisation, demonstrating how these techniques can be combined to achieve superior performance with reduced complexity.
- We conduct a detailed analysis of the resilience to noise in analogue neural network implementations, proposing novel techniques such as crowd equalisation to enhance performance in noisy environments.
- This work introduces the application of neuromorphic computing concepts, specifically spiking neural networks, to optical signal equalisation, demonstrating their potential for energy-efficient implementations.
- Finally, we provide a comprehensive framework for evaluating the performance, computational complexity, and energy efficiency of neural network-based equalisers on resource-constrained hardware, paving the way for their practical deployment in next-generation digital signal processing systems for optical communications.

These contributions collectively represent a significant advancement in the field of neural network-based equalisation for optical communication systems, bridging the gap between theoretical potential and practical implementation.

To provide a clear roadmap for presenting these contributions and their supporting research, the following section outlines the structure and organisation of this thesis.

1.5 Organisation of the Thesis

This thesis is organised as follows:

Chapter 1 introduces the increasing demand for high-capacity optical communication systems, reviews optical communication fundamentals, and presents an overview of machine learning applications in optical communications.

Chapter 2 presents the fundamentals of neural network-based equalisation in optical communication systems, covering traditional algorithms and ANN-based equaliser principles.

Chapter 3 focuses on performance optimisation and complexity reduction of neural network-based equalisers through advanced architectures (MLPs, CNNs, RNNs) and compression techniques including pruning, quantisation, and weight clustering for resource-constrained hardware deployment.

Chapter 4 addresses noise resilience in photonic neuromorphic computing, exploring analogue ANN implementations, noise models, and novel techniques such as stochastic resonance neurons for noise-resistant optical channel equalisation.

Chapter 5 adapts spiking neural networks for enhanced optical signal equalisation, presenting the SNN-RWKV equaliser design and comparing computational complexity and energy consumption with traditional methods.

Chapter 6 provides conclusions and directions for further work, including a summary of findings, limitations, future research directions, and broader impact.

Chapter 2

Neural Network-Based Equalisation in Optical Systems

2.1 Introduction

As it has already been explained in the previous Chapter, in optical communication systems the transmitted signal undergoes various forms of distortion as it propagates through the fibre. These linear and nonlinear distortions can significantly degrade the quality of the received signal, limiting the achievable data rates and transmission distances [37, 68]. Some examples of sources of impairment include Chromatic Dispersion (CD), PMD, SPM, or XPM, among others [52].

Equalisation is the process of compensating for these channel-induced distortions to recover the original transmitted signal. It plays a crucial role in modern optical communication systems, enabling higher data rates, longer transmission distances, and improved SE. Traditional equalisation techniques, such as Feed-Forward Equalisations (FFEs) and Decision Feedback Equalisations (DFEs), have been widely used. However, these linear methods cannot address the problems associated with the complex nonlinear distortions present in high-speed, long-haul optical systems [37].

The advent of coherent detection and DSP has opened new avenues for more sophisticated equalisation techniques [33, 38, 68]. Among these, ANNs-based equalisation has emerged as a promising approach due to its ability to model and compensate for both linear and nonlinear distortions without explicit knowledge of the channel characteristics [68, 74] .

2.2 The Equalisation Problem in Digital Communications

Equalisers are integral components in advanced digital communication systems, as they are employed to mitigate channel interference and distortion. This process facilitates the accurate retrieval of the transmitted data originally input into the channel. The evolution of communication technology during the last decades has accelerated the development of numerous equalisation techniques. Among these, it is worth mentioning three main linear equaliser algorithms: Zero Forcing Algorithm (ZFA), Least Mean Square (LMS), and Recursive Least Squares (RLS). Each of these algorithms employ different methodologies for addressing the challenges posed by signal interference and distortion in data transmission channels [37, 75].

In the context of this Thesis, where the focus is on coherent transmission systems, the equaliser at the receiver's end operates on a sequence $\{t_k\}$, which represents T-spaced complex symbols from an M-high-order QAM constellation. These symbols have in-phase $\Re\{t_k\}$ and quadrature $\Im\{t_k\}$ components, each taking values from the set $\{\pm 1, \pm 2, ..., \pm(\sqrt{M}-1)\}$. Here, $1/T$ is the symbol rate and k is the discrete time index. The "channel" encompasses the effects of both the transmitter's filter and the transmission medium, potentially introducing linear and nonlinear distortions to the signal t. A common model for a linear dispersive channel is the Finite Impulse Response (FIR) filter. This model is particularly useful for representing channels with memory, where the output at any given time depends not only on the current input but also on past inputs. Therefore, an output at the k-th instant is given by [75]:

$$a_k = \sum_{i=0}^{N_h-1} h_i t_{k-i}, \tag{2.1}$$

where h_i are the FIR filter coefficients, also known as tap weights or channel impulse response, and N_h is the filter length. This model can capture linear effects such as Inter-Symbol Interference (ISI) caused by multipath propagation or bandwidth limitations. The FIR filter model is widely used due to its simplicity and effectiveness in representing many real-world channels. It forms the basis for many equalisation techniques, such as linear equalisers and decision feedback equalisers. In order to model channels that contain nonlinear effects, it is necessary to modify the filter's output as follows [75, 76]:

$$b_k = \psi\left(a_k, a_{k-1}, a_{k-2}, ..., a_{k-N_h+1}; h_0, h_1, ..., h_{N_h-1}\right), \tag{2.2}$$

with $\psi\{\cdot\}$ being the nonlinear function applied by the channel. The output is also affected by Additive White Gaussian Noise (AWGN) q_k with variance σ^2. The signal t_k, after traversing the nonlinear channel and noise addition, is received as r_k, which at time k is $r_k = r_{k,l} + jr_{k,Q}$, where $r_{k,l}$ and $r_{k,Q}$ denote the in-phase and quadrature components, respectively [75].

The equaliser's task is to reconstruct the original transmitted sequence t_k or its time-delayed version $t_{k-\tau}$, with τ being the channel's propagation delay. It takes the noisy sequence r_k and its time-shifted versions as inputs, and outputs the compensated sequence y_k. The performance of the equaliser is measured by a loss function, typically the MSE [Loss = MSE(d, y)], calculated using the desired output ($d_k = t_{k-\tau}$) and the equalised output y_k. During the training phase, this loss function guides the adjustment of the equaliser's coefficients to minimise errors using an adaptive algorithm, such as the LMS algorithm. After training, the optimised coefficients are fixed and deployed to estimate the transmitted sequence in subsequent operations. The final criterion for evaluating the equaliser's performance is the BER. It quantifies the fraction of bits that are incorrectly received after equalisation and is a direct measure of the system's reliability in recovering the transmitted data. During deployment, BER serves as a key metric for assessing the effectiveness of the equaliser in mitigating noise and distortions introduced by the channel. A lower BER indicates higher performance and robustness of the equalisation process [37, 75].

2.3 Traditional Algorithms for Equalisation in Optical Communications

The field of nonlinear equalisation has gained significant attention as a means to enhance fibre capacity. This surge in interest is mainly a consequence of the fact that Kerr nonlinearities are a predominant factor in signal degradation within fibre optic channels, particularly when higher-order modulation formats are employed. To mitigate the signal distortions induced by fibre nonlinearities, a variety of advanced optical signal equalisers have been developed. This section reviews two key traditional nonlinear compensation strategies employed in fibre optic communication systems: DBP and Volterra Equaliser [37, 77].

2.3.1 Digital Back Propagation

Understanding DBP requires first examining why traditional linear compensation methods fail for optical fibre channels. When we consider the fundamental physics of light propagation in optical fibres, we encounter the NLSE, which governs the evolution of optical pulses in dispersive and nonlinear media. The challenge lies in the fact that this equation cannot be solved analytically for general cases, particularly when both dispersion and nonlinearity are significant [78, 79].

The conceptual foundation of DBP emerges from a profound insight: if we can mathematically model the forward propagation through the fibre using the NLSE, then we can theoretically reverse this process by solving the same equation with inverted parameters. This digital emulation of a hypothetical "inverse fibre" represents one of the most elegant approaches to nonlinear compensation in optical communications [70, 78].

To understand why DBP works, we must first appreciate the mathematical structure of the NLSE. For a dual-polarisation system, the coupled NLSE can be written as:

$$\frac{\partial \mathbf{A}}{\partial z} = -\frac{\alpha}{2}\mathbf{A} - i\frac{\beta_2}{2}\frac{\partial^2 \mathbf{A}}{\partial T^2} + i\gamma\left(\frac{8}{9}|\mathbf{A}|^2\mathbf{A}\right) \tag{2.3}$$

where $\mathbf{A} = [A_x, A_y]^T$ represents the complex field envelopes in both polarisations, z is the propagation distance, T is the retarded time, α is the attenuation coefficient, β_2 is the group velocity dispersion parameter, and γ is the nonlinear coefficient.

The fundamental challenge in solving this equation lies in the coupling between the linear dispersive term ($i\frac{\beta_2}{2}\frac{\partial^2 \mathbf{A}}{\partial T^2}$) and the nonlinear term ($i\gamma|\mathbf{A}|^2\mathbf{A}$). The Split-Step Fourier Method (SSFM) addresses this challenge through a mathematically elegant approach: it recognises that while the combined linear-nonlinear equation is intractable, each component can be solved independently if we assume they act over sufficiently small propagation steps [78].

The SSFM algorithm operates on the principle of operator splitting, where the propagation over a small step h is decomposed into sequential linear and nonlinear operations:

$$\mathbf{A}(z+h,T) = \exp(h\hat{N})\exp(h\hat{L})\mathbf{A}(z,T) + O(h^3) \tag{2.4}$$

where $\hat{L}$ represents the linear operator (dispersion and attenuation) and $\hat{N}$ represents the nonlinear operator (Kerr effect). The key insight is that this decomposition introduces only third-order errors in the step size, making it highly accurate for practical implementations.

The linear step addresses dispersion and attenuation effects in the frequency domain, where these operations become multiplicative rather than differential. This transformation is the mathematical foundation for why the Fast Fourier Transform (FFT) becomes central to the algorithm:

$$\mathbf{A}_{linear}(z+h,T) = \mathcal{F}^{-1}\left\{\mathcal{F}[\mathbf{A}(z,T)]\exp\left(-\frac{\alpha h}{2} + i\frac{\beta_2\omega^2 h}{2}\right)\right\} \tag{2.5}$$

where $\mathcal{F}[\cdot]$ and $\mathcal{F}^{-1}[\cdot]$ denote the FFT and Inverse Fast Fourier Transform (IFFT) operations, respectively, ω is the angular frequency variable, and h is the step size. The exponential term captures both the attenuation (real part) and dispersion (imaginary part) effects that accumulate over the step.

Following the linear step, the nonlinear phase shift due to the Kerr effect is applied in the time domain, where the intensity-dependent refractive index change manifests as a phase rotation:

$$\mathbf{A}(z+h,T) = \mathbf{A}_{linear}(z+h,T)\exp\left(i\gamma h\frac{8}{9}|\mathbf{A}_{linear}(z+h,T)|^2\right) \tag{2.6}$$

The factor $\frac{8}{9}$ arises from the statistical averageing over random polarisation rotations in the fibre, representing the effective nonlinear interaction between polarisation modes.

Understanding the step size selection is crucial for balancing computational efficiency with numerical accuracy. The error in the SSFM algorithm scales as $O(h^3)$, meaning that halving the step size reduces the error by a factor of eight. However, this comes at the cost of doubling the computational burden. The optimal step size must satisfy two critical conditions [78, 79]:

1. Dispersion Accuracy Condition: The step size must be small enough that dispersion-induced pulse broadening over one step remains small compared to the pulse width:

$$h < \frac{T_0^2}{|\beta_2|} \tag{2.7}$$

where T_0 is the characteristic pulse width.

2. Nonlinear Accuracy Condition: The nonlinear phase shift accumulated over one step should be much less than π:

$$\gamma P_0 h \ll \pi \tag{2.8}$$

where P_0 is the peak power.

The computational complexity of DBP scales linearly with the number of steps, which is inversely proportional to the step size. For a fibre link of length L, the total number of steps is $N_{steps} = L/h$, and each step requires two FFT operations, leading to a total computational complexity of:

$$\mathcal{O}_{DBP} = \frac{2L}{h} \cdot N_{samples} \log_2(N_{samples}) \tag{2.9}$$

where $N_{samples}$ is the number of time samples. This scaling reveals why DBP becomes computationally prohibitive for very accurate step sizes or very long transmission distances.

Despite its theoretical elegance, DBP faces several fundamental limitations that reveal the inherent trade-offs in nonlinear optical communication systems. The most significant challenge arises from the stochastic nature of ASE noise, which introduces signal-noise interactions that cannot be reversed through deterministic algorithms. When ASE noise mixes with the signal during nonlinear propagation, it generates additional noise components that effectively reduce the signal-to-noise ratio. This noise-nonlinearity interaction represents a fundamental physical limitation that no digital algorithm can completely overcome [79, 80].

The computational complexity scaling presents another critical limitation. Real-time implementation requires processing rates that match or exceed the symbol rate, typically in the range of tens of GBaud for modern systems. The computational burden becomes particularly acute when we consider that practical systems require step sizes on the order of 100 meters or less for accurate compensation, leading to thousands of steps for typical transmission distances. This computational requirement has driven the development of specialised hardware implementations and algorithmic approximations [70].

For WDM systems, the challenge becomes even more complex. Inter-channel nonlinear effects such as XPM and FWM require joint processing of multiple channels, expanding the computational requirements by the number of channels. Multichannel DBP theoretically provides optimal performance but demands computational resources that scale superlinearly with the number of channels, making it impractical for dense WDM systems [33, 80].

These practical constraints have led to the widespread adoption of Single-Channel DBP (SC-DBP), which focuses exclusively on compensating intra-channel nonlinearities such as SPM. While SC-DBP sacrifices some performance compared to full multichannel DBP, it provides a reasonable balance between compensation effectiveness and computational feasibility. The performance gap between SC-DBP and full DBP depends on the system parameters, with the difference becoming more pronounced in dense WDM systems with high launch powers [80].

Understanding these limitations motivates the exploration of alternative approaches to nonlinear compensation, including polynomial-based methods such as Volterra equalisers and, more recently, machine learning-based approaches that can potentially achieve similar performance with reduced computational complexity.

2.3.2 Volterra Equaliser

The choice of polynomial expansions in Volterra equalisers reflects a deeper principle about approximating nonlinear systems. When we consider how nonlinear distortions manifest in optical systems, we recognise that the Kerr effect creates intensity-dependent phase shifts that can be mathematically represented through polynomial expansions of the signal power. This fundamental insight leads us to the VSTF approach, which provides a systematic framework for characterising nonlinear systems with memory effects [81].

Understanding why Volterra series work for optical channels requires examining the Taylor series expansion of the nonlinear refractive index. The intensity-dependent refractive index can be written as $n(I) = n_0 + n_2 I + n_4 I^2 + \ldots$, where higher-order terms typically become negligible for moderate power levels. This polynomial relationship directly translates to the Volterra series structure, making it a natural mathematical tool for modelling optical nonlinearities [81, 82].

For a discrete-time system with input $y(n)$ and output $\tilde{y}(n)$, the complete Volterra series expansion can be written as:

$$\tilde{y}(n) = \sum_{p=1}^{P} \sum_{m_1=0}^{M_1} \cdots \sum_{m_p=0}^{M_p} h_p(m_1, \ldots, m_p) \prod_{k=1}^{p} y(n - m_k) \tag{2.10}$$

where $h_p(m_1, \ldots, m_p)$ represents the Volterra kernel of order p, and M_p is the memory depth for the p-th order term. To understand which terms contribute significantly, we must analyse each order's physical meaning and typical magnitude. The First-Order Terms ($p = 1$) represent the linear channel response, capturing dispersion, attenuation, and linear filtering effects:

$$y_1(n) = \sum_{m_1=0}^{M_1} h_1(m_1) y(n - m_1) \tag{2.11}$$

This term is always significant and typically dominates the system response. In optical systems, $h_1(m_1)$ captures the impulse response of the linear channel, including chromatic dispersion effects.

The Second-Order Terms ($p = 2$) capture signal-signal interaction effects, primarily SPM in single-channel systems:

$$y_2(n) = \sum_{m_1=0}^{M_2} \sum_{m_2=0}^{M_2} h_2(m_1, m_2) y(n - m_1) y(n - m_2) \tag{2.12}$$

For optical channels, many second-order terms are zero due to physical constraints. Specifically, terms where $m_1 \neq m_2$ often contribute minimally because SPM primarily depends on the instantaneous power, not cross-correlations between different time instances.

The Third-Order Terms ($p = 3$) represent more complex nonlinear interactions:

$$y_3(n) = \sum_{m_1=0}^{M_3} \sum_{m_2=0}^{M_3} \sum_{m_3=0}^{M_3} h_3(m_1, m_2, m_3) y(n - m_1) y(n - m_2) y(n - m_3) \tag{2.13}$$

Many third-order terms are also negligible in typical optical systems. The dominant contributions come from terms where multiple indices are equal, representing higher-order intensity-dependent effects.

The computational complexity analysis reveals why practical Volterra equalisers are typically limited to low orders and short memory lengths. For a system with maximum polynomial order P and memory depth M, the total number of coefficients scales as:

$$N_{coefficients} = \sum_{p=1}^{P} \binom{M + p - 1}{p} \approx \frac{M^P}{P!} \tag{2.14}$$

This exponential growth in complexity explains why most practical implementations are limited to second or third-order terms with memory depths of 10-20 samples. For example, a third-order Volterra equaliser with $M = 15$ memory taps requires approximately 1,000 coefficients, while extending to fifth order would require over 100,000 coefficients.

The coefficient calculation presents another significant challenge. Unlike linear equalisers where coefficients can be determined through matrix inversion, Volterra coefficients require more sophisticated optimisation techniques. The most common approaches include:

LMS Adaptation: The coefficients are updated iteratively using:

$$h_p(m_1, \ldots, m_p)(n + 1) = h_p(m_1, \ldots, m_p)(n) + \mu \cdot e(n) \cdot \prod_{k=1}^{p} y(n - m_k) \tag{2.15}$$

where μ is the adaptation step size and $e(n)$ is the error signal.

RLS: This provides faster convergence but requires matrix operations that scale quadratically with the number of coefficients.

The memory depth selection involves balancing compensation performance against computational complexity. In optical systems, the effective memory is determined by the chromatic dispersion-induced pulse broadening and the nonlinear length scale. For typical fibre parameters, memory depths of 5-15 samples often capture the dominant nonlinear interactions while maintaining computational tractability [81, 83]. To manage complexity, several truncation approaches are employed [84–86]:

1. Order Truncation: Limiting to second or third-order terms only

2. **Diagonal Truncation**: Keeping only terms where all memory indices are equal (capturing mainly SPM effects)
3. **Sparse Truncation**: Selecting only statistically significant coefficients through techniques like Least Absolute Shrinkage and Selection Operator (LASSO) regression

The frequency domain implementation offers computational advantages, particularly for WDM systems. By transforming to the frequency domain, each WDM channel can be processed independently, and the convolution operations become simple multiplications. However, this approach requires careful windowing to avoid circular convolution artefacts and adds FFT/IFFT computational overhead [87, 88].

In terms of the performance and complexity trade-offs, it is demonstrated that that Volterra equalisers typically achieve 70-90% of the performance of full DBP with 10-100 times lower computational complexity, making them attractive for real-time implementations [82].

These performance and complexity characteristics have motivated exploration of alternative approaches to nonlinear equalisation. Reservoir computing represents an alternative paradigm for processing nonlinear systems with memory, where computation occurs through the dynamics of a fixed, randomly connected recurrent neural network (the "reservoir"), with only the output layer trained through linear regression [89, 90]. This approach has gained significant attention in photonic implementations due to its potential for energy-efficient, high-speed signal processing, as detailed extensively in Chapter 4.

While both Volterra equalisers and reservoir computers are capable of handling nonlinear systems with memory, they employ fundamentally different approaches to process temporal information and compensate for channel distortions [89, 90]. Understanding their similarities and differences is crucial for selecting the most appropriate technique for optical channel equalisation applications.

Both methods excel at processing nonlinear systems with memory effects, making them suitable for optical channel equalisation where temporal correlations and nonlinear distortions are prevalent. Both approaches can handle temporal sequences effectively, learning from historical data to predict current outputs. Additionally, both techniques can model complex input-output relationships without requiring explicit mathematical models of the underlying physical processes, making them attractive for applications where the channel characteristics are difficult to model analytically [83, 89].

The fundamental distinction lies in their mathematical frameworks. Volterra equalisers use polynomial expansions to represent nonlinear systems, expressing the output as a weighted sum of products of delayed input signals up to a specified order. In contrast, reservoir computers employ recurrent neural network dynamics with a fixed, randomly connected reservoir that transforms the input into a high-dimensional representation, from which the desired output is extracted through linear regression [89, 90].

Regarding memory processing, Volterra equalisers handle memory through explicit memory taps, where each coefficient corresponds to a specific combination of delayed inputs. The memory depth is directly controlled by the number of taps and the polynomial order. Reservoir computers, however, implement implicit temporal dynamics through their recurrent connections, where the reservoir's internal state naturally incorporates temporal information through its dynamic evolution [90].

In terms of computational complexity, Volterra equalisers face exponential growth in computational complexity as the memory length and polynomial order increase, with the number of coefficients scaling as $O(M^P)$ where M is the memory depth and P is the polynomial order. Reservoir computers maintain relatively constant computational complexity during operation, as only the output weights require adaptation while the reservoir connections remain fixed. However, the reservoir size must be sufficiently large to provide adequate computational capacity [83, 89].

Furthermore, when evaluating the implementation for optical channel equalisation, Volterra equalisers offer the advantage of explicit parameter control and interpretability, allowing engineers to directly relate coefficients to specific nonlinear effects. Reservoir computers provide greater flexibility in handling complex temporal dynamics and can potentially achieve better performance for channels with long memory, but may require more sophisticated training procedures and larger computational resources [37, 90].

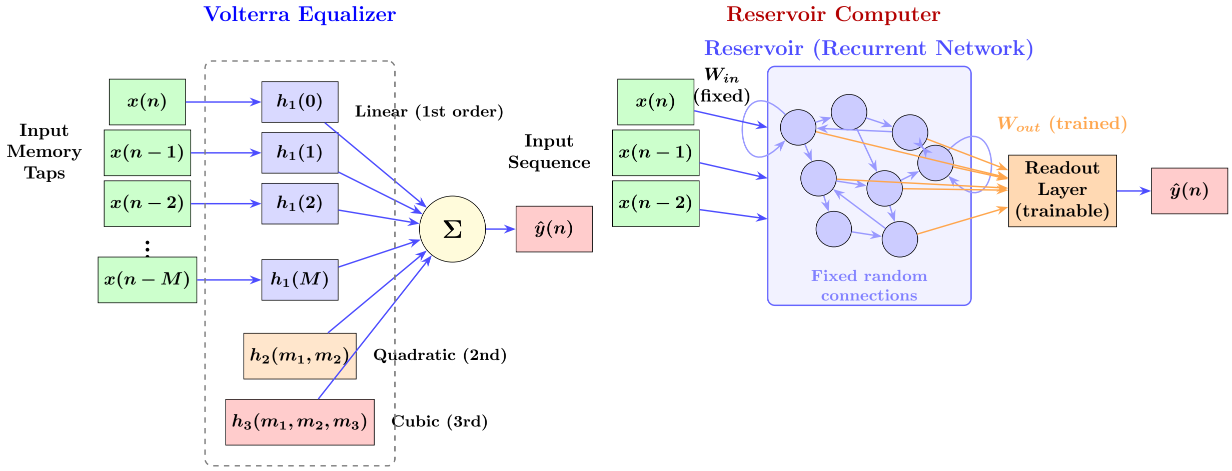


Figure 2.1: Architectural comparison between Volterra Equaliser and Reservoir Computer approaches for nonlinear equalisation. Left: Volterra equaliser employs explicit memory taps $\{x(n), x(n-1), \ldots, x(n-M)\}$ with polynomial expansion through linear (h_1), quadratic (h_2), and cubic (h_3) kernels, where all coefficients are trainable parameters optimised simultaneously. Right: Reservoir computer processes input sequences through a fixed, randomly connected recurrent neural network (reservoir, W_{in}) that implicitly encodes temporal memory through recurrent dynamics, with only the linear readout layer (W_{out}) trained via regression. The Volterra approach offers explicit control over nonlinearity order and memory depth at the cost of exponentially growing parameter complexity, while the reservoir computing paradigm achieves efficient training through fixed dynamics and linear output optimisation, though with less interpretable internal representations.

2.4 Optical Solutions for Equalisation in Optical Communication Systems

Understanding why optical phase conjugation emerges as a compelling alternative to digital compensation requires recognising the fundamental trade-offs between computational complexity and physical implementation. While digital approaches like DBP and Volterra equalisation provide flexibility and precision, they demand substantial computational resources that scale poorly with system complexity. This computational burden becomes particularly acute in high-speed optical systems where real-time processing requirements can exceed the capabilities of practical digital signal processors. In this context, optical nonlinear compensation represents a paradigm shift: instead of digitally reversing channel effects after detection, we can exploit the same physical phenomena that cause distortions to actively compensate for them within the optical domain itself [91, 92].

The conceptual foundation of OPC emerges from the time-reversal symmetry inherent in the Nonlinear Schrödinger Equation that governs optical propagation. When we examine the NLSE more carefully, we discover that if we can create the complex conjugate of an optical field and propagate it through a medium with inverted parameters, the accumulated nonlinear phase can theoretically be undone. This insight leads us to understand that OPC is not merely a signal processing technique, but rather a fundamental exploitation of the reciprocity principles underlying optical wave propagation [92, 93].

To understand the mathematical foundation of OPC, we must first examine how nonlinear phase accumulation occurs in optical fibres. As an optical signal propagates through a fibre span of length L, it accumulates a nonlinear phase shift given by:

$$\phi_{NL}(T) = \gamma \int_0^L |A(z,T)|^2 dz \tag{2.16}$$

where γ is the nonlinear coefficient, $A(z,T)$ is the complex field envelope, and the integration accounts for power variations along the fibre due to attenuation. The key insight is that this phase accumulation is deterministic and depends on the signal's intensity profile.

The phase conjugation operation mathematically transforms the complex field $A(T)$ into its conjugate $A^*(T)$, effectively reversing the sign of the accumulated nonlinear phase. When this conjugated signal propagates through a second fibre span with identical characteristics, it experiences a nonlinear phase shift of opposite sign:

$$\phi_{NL,conjugated}(T) = -\gamma \int_0^L |A^*(z,T)|^2 dz = -\phi_{NL}(T) \tag{2.17}$$

This mathematical reciprocity forms the theoretical basis for perfect nonlinear compensation, provided that certain symmetry conditions are met.

However, the practical implementation of OPC reveals the complexity underlying this seemingly elegant solution. The symmetry requirements are quite stringent and encompass several critical aspects. Power symmetry demands that the signal power profile in the second fibre span must mirror that of the first span. Since fibre attenuation causes exponential power decay, achieving power symmetry requires either distributed Raman amplification or careful placement of discrete amplifiers to recreate the power evolution profile. Dispersion symmetry requires that the accumulated chromatic dispersion in both spans be identical, which typically necessitates dispersion-managed fibre layouts or dispersion compensation modules. Length symmetry mandates that both fibre spans have identical effective lengths, accounting for variations in fibre parameters.

The mathematical formulation of these symmetry conditions can be expressed through the condition:

$$\int_0^{L_1} \gamma_1(z)P_1(z)dz = \int_0^{L_2} \gamma_2(z)P_2(z)dz \tag{2.18}$$

where the subscripts 1 and 2 refer to the fibre spans before and after the OPC device, and $P(z)$ represents the power evolution along each span.

The performance analysis of OPC systems reveals both the potential benefits and fundamental limitations. Reach Extension is one of the most significant advantages, with demonstrations showing transmission distance improvements of 50-100% compared to linear compensation alone. This improvement arises because OPC can handle higher launch powers without nonlinearity penalties, effectively pushing the system operating point to higher signal-to-noise ratios. Power Scaling Benefits allow for increased launch powers that would otherwise be limited by nonlinear distortions, leading to improved system performance in noise-limited scenarios.

Q-factor Improvement in OPC systems can be quantified through the relationship:

$$Q_{OPC} = Q_{linear}\sqrt{1 + \frac{P_{launch,OPC}}{P_{launch,linear}}} \tag{2.19}$$

where $P_{launch,OPC}$ and $P_{launch,linear}$ represent the optimal launch powers with and without OPC, respectively.

Despite these advantages, practical implementation challenges reveal the complexity of translating theoretical benefits into real-world systems. Broadband Operation requires that the phase conjugation process work effectively across the entire optical bandwidth, which can span several tens of nanometers in dense WDM systems. This broadband requirement places stringent demands on the wavelength conversion efficiency and bandwidth of the OPC device. Noise Figure Penalty arises because most practical OPC implementations introduce additional noise through the wavelength conversion process, partially offsetting the SNR improvements from higher launch powers. Conversion Efficiency limitations mean that not all of the input signal power is successfully converted to the conjugated wavelength, introducing insertion losses that must be compensated by additional amplification.

The implementation of OPC typically relies on wavelength conversion techniques, with FWM in highly nonlinear fibres or Periodically Poled Lithium Niobate (PPLN) waveguides being the most common approaches. FWM-based OPC leverages the third-order nonlinearity to generate a conjugated idler wave through the interaction:

$$\omega_{idler} = 2\omega_{pump} - \omega_{signal} \tag{2.20}$$

This approach offers broad bandwidth operation but requires careful management of pump power and fibre dispersion. PPLN-based OPC utilises quasi-phase-matched second-order nonlinearity, providing higher conversion efficiency but with more limited bandwidth.

Recent advances in silicon photonics have opened new possibilities for integrated OPC implementations. Symbol of Interest (SOI) platforms can achieve high nonlinear coefficients in compact form factors, enabling chip-scale OPC devices with reduced power consumption and improved scalability. However, silicon's two-photon absorption at telecommunications wavelengths introduces additional loss mechanisms that must be carefully managed.

The comparison between lumped and distributed phase conjugation approaches reveals different trade-offs in system performance. Lumped OPC places the phase conjugation device at the midpoint of the transmission link, requiring precise symmetry in the fibre spans before and after conjugation. Distributed OPC embeds the conjugation functionality throughout the transmission medium, potentially offering better compensation but with increased system complexity.

A related technique that exploits similar principles is Phase Conjugated Twin Waves (PCTW), which provides an alternative implementation approach for nonlinear compensation. PCTW involves transmitting a signal and its conjugate on separate polarisations within dual-polarisation systems, implemented at the receiver end through digital signal processing. This technique assumes that nonlinear distortions in twin waves are inversely correlated, allowing cancellation of the first-order nonlinear phase shift through signal superposition at the receiver. While PCTW efficiently counters deterministic nonlinear phase shifts and phase noise caused by signal-noise interactions, its primary limitation lies in the reduction in spectral efficiency due to the transmission of the signal's conjugate. Recent developments in PCTW have addressed this spectral efficiency concern through advanced encoding techniques, though they may exhibit slightly reduced compensation performance [94, 95].

Integration with digital signal processing techniques has emerged as a promising hybrid approach that combines the benefits of optical and digital compensation. By using OPC to handle the majority of nonlinear distortions and digital techniques to address residual impairments and linear effects, these hybrid systems can achieve superior performance while maintaining computational efficiency [96–98]. This integration strategy is particularly attractive for next-generation systems where both reach and data rate requirements exceed the capabilities of purely digital or purely optical solutions.

The cost-benefit analysis of OPC implementation reveals complex trade-offs between performance improvements and system complexity. While OPC can significantly reduce the computational requirements for nonlinear compensation, it introduces additional hardware components, wavelength management complexity, and potential reliability concerns [96, 99]. The economic viability of OPC depends heavily on the specific application requirements and the relative costs of optical components versus digital processing power.

Practical deployment considerations encompass network architecture implications, including the need for dispersion-managed fibre layouts, additional wavelength allocation for pump and idler waves, and network management systems capable of optimising OPC performance across varying traffic conditions. These considerations highlight that while OPC offers compelling technical benefits, its practical deployment requires careful system-level integration and optimisation [99].

In the context of modern high-capacity optical networks, OPC represents a complementary technology that can work synergistically with advanced digital signal processing techniques. Rather than replacing digital compensation entirely, OPC can reduce the computational burden on digital equalisers, enabling more sophisticated algorithms to be deployed in real-time systems. This complementary approach aligns with the broader trend toward hybrid optical–digital signal processing architectures that leverage the unique advantages of each domain to achieve optimal system performance [98].

2.5 Fundamentals of ANN-based Equalisers

ANN-based equalisers have emerged as a promising alternative to traditional equalisation techniques in optical communication systems. While traditional approaches like DBP and Volterra equalisers provide effective solutions for specific channel impairments, they face fundamental limitations in handling the complex, multidimensional nature of optical channel distortions. DBP, although theoretically optimal, suffers from computational complexity that scales with the number of channels and step size requirements, making it impractical for real-time, high-speed applications. Volterra equalisers, despite their ability to model nonlinear systems with memory, experience exponential growth in computational complexity as channel memory increases, limiting their applicability for long-haul transmission systems.

In contrast, ANN-based equalisers leverage the universal approximation theorem, which states that neural networks with sufficient width and depth can approximate any continuous function to arbitrary precision [100]. This theoretical foundation provides a powerful framework for learning the inverse channel function without requiring explicit mathematical models of the underlying physical processes. The key advantage of ANN-based equalisers lies in their ability to learn complex channel characteristics directly from data, automatically discovering optimal compensation strategies through supervised learning rather than relying on predefined mathematical formulations [37].

ANNs-based equalisers operate on the principle of supervised learning. The basic process involves:

1. Data collection: Gathering pairs of transmitted and received symbols from the optical channel.

2. Network training: Adjusting the ANN parameters to minimise the difference between the network's output and the desired (transmitted) signal.

3. Deployment: Using the trained ANN to equalise new, unseen received signals.

Supervised learning is particularly advantageous in optical communication systems because we have a robust mathematical model for the fibre channel, such as the NLSE. Therefore, by simulating the NLSE, we can generate highly accurate training datasets under various conditions, including different power levels, modulation formats, and noise levels. This allows the ANN to learn the intrinsic behaviours of the channel and effectively compensate for both linear and nonlinear distortions. Furthermore, supervised learning leverages this simulated data to train the ANN without requiring extensive real-world measurements, making it a cost-effective and scalable approach to nonlinearity mitigation in high-speed optical networks [37, 101].

The mathematical foundation of this process can be described as follows:

Let $x(n) \in \mathbb{C}$ be the transmitted signal and $y(n) \in \mathbb{C}$ be the received signal after passing through the optical channel. The goal of the ANN-based equaliser is to find a function $f : \mathbb{C}^{M+1} \to \mathbb{C}$ such that:

$$\hat{x}(n) = f(y(n), y(n-1), ..., y(n-M)) \tag{2.21}$$

where $\hat{x}(n)$ is the estimated transmitted signal and M is the memory of the equaliser. The function $f(\cdot)$ is parameterised by the neural network weights $\boldsymbol{\theta} = \{\mathbf{W}_1, \mathbf{b}_1, ..., \mathbf{W}_D, \mathbf{b}_D\}$, where D is the network depth [68].

The optimisation objective is to minimise the expected risk:

$$\mathcal{R}(\boldsymbol{\theta}) = \mathbb{E}_{(x,y)\sim\mathcal{D}}\left[\mathcal{L}(x(n), f(y(n), ..., y(n-M); \boldsymbol{\theta}))\right] \tag{2.22}$$

where $\mathcal{D}$ represents the underlying data distribution, $\mathcal{L}(\cdot, \cdot)$ is the loss function, and $\mathbb{E}[\cdot]$ denotes the expectation operator. In practice, this is approximated using empirical risk minimisation over the training dataset.

The function $f(\cdot)$ that the ANN aims to approximate is, in fact, the inverse of the channel transfer function. This inverse function can be highly complex, especially in the presence of nonlinear impairments. The channel transfer function can be generally expressed as:

$$y(n) = h(x(n), x(n-1), ..., x(n-M)) + w(n) \tag{2.23}$$

where $h(\cdot)$ represents the combined effects of linear and nonlinear distortions in the channel, M is the channel memory, and $w(n)$ is additive noise [37].

The function $h(\cdot)$ encompasses a comprehensive range of physical phenomena that affect signal propagation in optical fibres. Linear effects, primarily CD, cause frequency-dependent phase

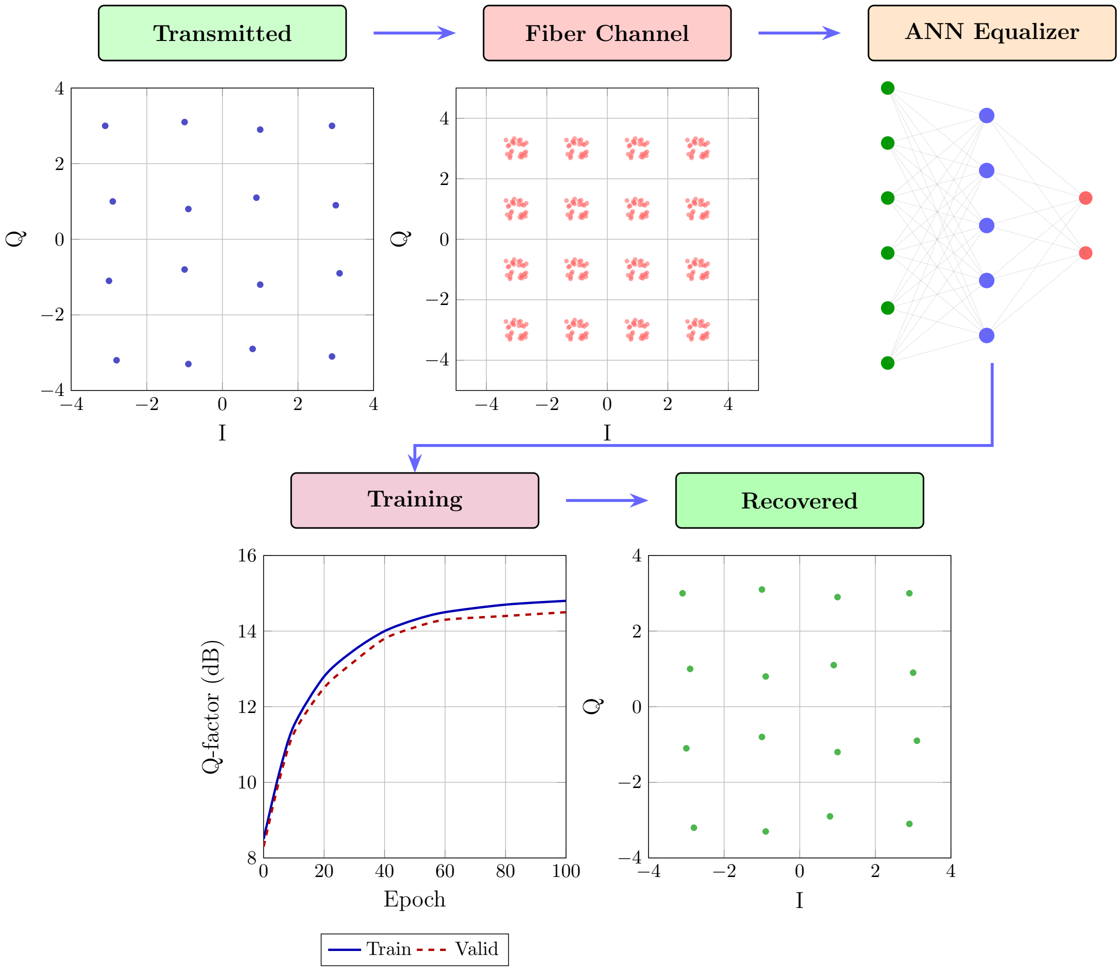


Figure 2.2: Complete equalisation process using ANN-based equalisers: The first stage shows the transmitted signal constellation showing the original 16-QAM data, then it comes the fibre channel output after nonlinear distortion, followed by the ANN equaliser architecture, as well as the training evolution, showing Q-factor improvement with epochs for both training and validation sets. Finally, the recovered signal constellation after equalisation.

shifts that lead to pulse broadening and intersymbol interference. Nonlinear effects, including SPM, XPM, and FWM, arise from the intensity-dependent refractive index changes in the fibre, creating complex distortions that depend on signal power and spectral content. Additionally, memory effects manifest as ISI, where the distortion of each symbol depends not only on its own characteristics but also on the temporal sequence of neighbouring symbols, creating dependencies that extend across multiple symbol periods.

As already mentioned, in the case of a simplified model considering only CD and Kerr nonlinearity, $h(\cdot)$ can be approximated by the solution to the NLSE.

While discussing channel impairments, it is important to note that PMD, a dynamic phenomenon that changes on a microsecond scale, poses significant challenges for machine learning-based equalisers, as adaptive filtering is needed to handle it in practice. As Machine Learning (ML)-based equalisers are typically trained on large, static datasets, their fixed nature makes them less effective in adapting to rapidly changing PMD [102].

The ANN-based equaliser must learn to invert this complex function, which is typically not analytically tractable. The fundamental challenge in nonlinear compensation lies in the fact that the channel function $h(\cdot)$ is a nonlinear operator that maps the input signal space to the output signal space in a complex, history-dependent manner. This nonlinearity arises from the intensity-dependent refractive index changes in the fibre (Kerr effect), which cause the propagation characteristics to vary with the signal power and temporal distribution.

The nonlinear compensation process can be understood through the lens of functional analysis. The channel can be viewed as a nonlinear operator $\mathcal{H} : \mathcal{X} \to \mathcal{Y}$, where $\mathcal{X}$ and $\mathcal{Y}$ are appropriate function spaces representing the transmitted and received signal spaces, respectively. The equaliser seeks to approximate the inverse operator $\mathcal{H}^{-1} : \mathcal{Y} \to \mathcal{X}$ such that $\mathcal{H}^{-1}(\mathcal{H}(x)) \approx x$ for signals $x \in \mathcal{X}$.

The power of ANNs lies in their ability to approximate such complex, nonlinear functions through the composition of simpler functions, leverageing the universal approximation theorem:

$$f(y) = f_D(f_{D-1}(...f_1(y))) \tag{2.24}$$

where D is the number of layers in the ANN, and each f_i represents the operation of a single layer, typically involving a linear transformation followed by a nonlinear activation function:

$$f_i(z) = \sigma(W_i z + b_i) \tag{2.25}$$

Here, W_i and b_i are the weight matrix and bias vector for layer i, respectively, and $\sigma(\cdot)$ is a nonlinear activation function such as the ReLU or hyperbolic tangent (tanh) [100].

The challenge in designing effective ANN-based equalisers lies in selecting appropriate model architectures, activation functions, and training algorithms that can accurately approximate the inverse channel function while maintaining an appropriate Computational Complexity (CC) for real-time operation in high-speed optical communication systems.

Based on the theoretical foundations described above, ANN-based equalisers offer several key advantages over traditional equalisation techniques [100, 103]. Their adaptability represents a fundamental departure from conventional approaches: unlike DBP, which requires explicit knowledge of fibre parameters, or Volterra equalisers, which require predetermined polynomial orders, ANNs can adapt to changing channel conditions through data-driven learning without requiring explicit channel models or manual parameter tuning. This adaptive capability extends to their superior nonlinearity compensation, where ANNs can effectively model and compensate for the complex nonlinear distortions arising from the Kerr effect in optical fibres. Unlike traditional approaches that rely on specific mathematical formulations (e.g., NLSE solutions for DBP or polynomial expansions for Volterra equalisers), ANNs learn the inverse channel function directly from training data, potentially capturing higher-order nonlinear interactions that may be missed by traditional methods.

Furthermore, ANNs enable joint optimisation for multiple impairments simultaneously, leverageing end-to-end training to balance trade-offs across different channel effects. This contrasts with traditional approaches that typically address impairments sequentially (e.g., separate CD and nonlinearity compensation stages), often leading to suboptimal overall performance. The versatility of neural network architectures allows the same fundamental design to be applied to different modulation formats and transmission scenarios with minimal modifications. This flexibility represents a significant advantage over specialised approaches like DBP, which requires different implementations for different fibre types and transmission scenarios, making ANNs particularly attractive for dynamic and heterogeneous optical networks [37].

Nevertheless, despite these advantages, ANN-based equalisers face several significant challenges [37, 68, 104]. Computational complexity represents a primary concern, as both training and inference can be computationally intensive, especially for high-speed optical systems where real-time implementation becomes challenging due to stringent hardware requirements. While DBP has deterministic computational complexity that scales predictably with system parameters, ANNs require careful architecture design to balance performance and computational efficiency, often necessitating specialised hardware or algorithmic approximations for practical deployment.

2.5.1 Neural Network Architecture Considerations

The effectiveness of ANN-based equalisers depends critically on the choice of neural network architecture, which must be matched to the specific characteristics of the optical channel impairments. Different architectural paradigms—including feed-forward networks (Multi-Layer Perceptrons), convolutional networks, and recurrent networks (vanilla RNNs, Long Short-Term Memory networks, and Gated Recurrent Units)—offer distinct advantages for processing the temporal and nonlinear structure of optical signals. The selection of appropriate architectures, their optimisation, and complexity reduction strategies are extensively discussed in Chapter 3, where we explore how architectural choices directly impact both equalisation performance and computational efficiency.

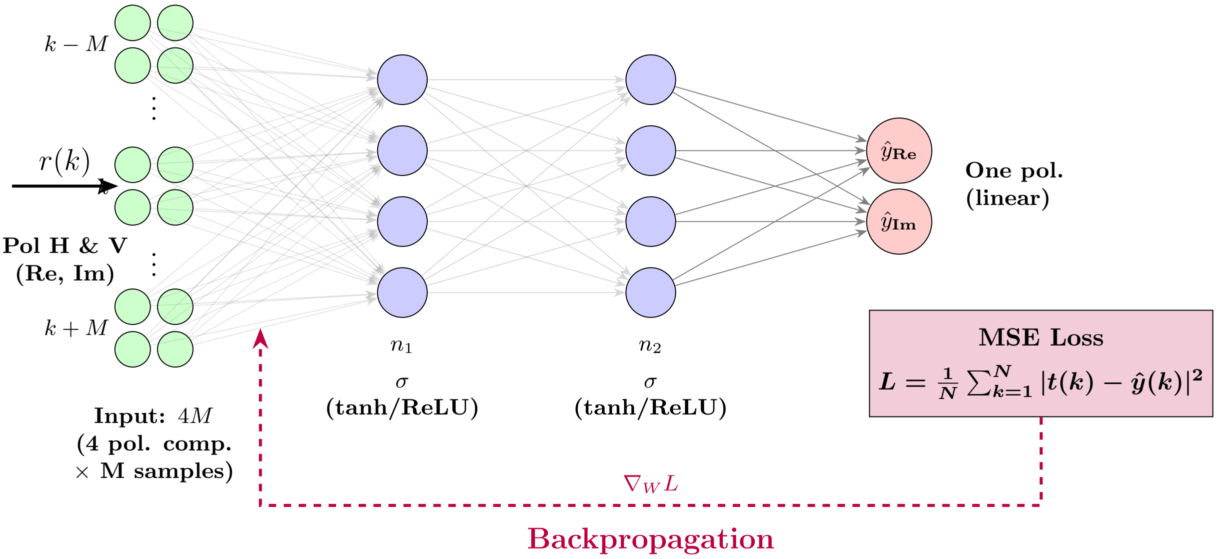


Figure 2.3: Fundamental architecture of a feedforward ANN for optical signal equalisation. The received signal $r(k)$ from both polarisations (H and V) is processed over a temporal window of $2M+1$ samples, providing the real and imaginary components as inputs. Two hidden layers with n_1 and n_2 neurons apply nonlinear transformations (tanh or ReLU activation functions) to extract relevant features. The output layer estimates the transmitted symbol components $\hat{y}_{\mathrm{Re}}$ and $\hat{y}_{\mathrm{Im}}$ for linear polarisation. Training employs backpropagation with MSE loss to minimise the error between estimated $\hat{y}(k)$ and transmitted $t(k)$ symbols.

In coherent dual-polarisation systems, the ANN processes the received signal $r(k)$ over a temporal window centred at the symbol of interest. This window spans $2M+1$ consecutive symbol periods, where M represents the number of past and future symbols considered on each side of the current symbol. Each symbol contributes four real-valued features to the input: the real and imaginary components from both the horizontal (H) and vertical (V) polarisations. Therefore, the complete input to the neural network has a total dimension of $(2M+1)\times 4$ features, representing the in-phase and quadrature components (I_H, Q_H, I_V, Q_V) for each of the $2M+1$ temporal samples. This input structure captures both the temporal dependencies introduced by chromatic dispersion and inter-symbol interference, as well as the polarisation coupling effects that arise from fibre birefringence and nonlinear interactions [103].

For single-polarisation processing, the input dimension reduces to $(2M+1)\times 2$ features, considering only the real and imaginary components from one polarisation. The parameter M, often referred to as the memory depth or half-window size, is typically optimised during the hyperparameter tuning process and commonly ranges from 5 to 20 symbols depending on the channel characteristics and transmission distance [103]. The choice of M directly impacts both the equalisation performance and computational complexity, as larger values capture longer-range temporal dependencies at the cost of increased parameter count and processing latency. This trade-off between memory depth and computational burden represents a fundamental design consideration in ANN-based equalisers, where the number of input features grows linearly with M, while the total number of network parameters typically scales quadratically with the input dimension for fully-connected architectures.

For the purposes of establishing the fundamental equalisation framework, it suffices to recognise that the neural network function $f(\cdot)$ in the preceding equations can be implemented through various architectural choices, each with specific computational characteristics and modelling capabilities. The subsequent sections in this chapter focus on the loss functions, optimisation algorithms, hyperparameter optimisation, and data requirements that are common to all neural network-based equalisers regardless of the specific architecture employed.

2.5.2 Loss Functions and Optimisation

The choice of loss function is critical in training ANN-based equalisers. For regression tasks in equalisation, the MSE is commonly used. The MSE loss function measures the average of the squares of the errors—that is, the average squared difference between the estimated values ($\hat{x}(n)$) and the actual value $x(n)$. Mathematically, it is defined as [37]:

$$L_{\mathsf{MSE}}(x, \hat{x}) = \frac{1}{N} \sum_{n=1}^{N} (x(n) - \hat{x}(n))^2 \tag{2.26}$$

where N is the number of samples. This loss function is particularly suitable for regression problems where the goal is to minimise the error between predicted and true values.

In the context of optical communication systems, the terms of Equation 2.26 have specific physical interpretations [105]. The term $x(n)$ represents the transmitted signal consisting of digital symbols that were originally sent through the optical channel, while $\hat{x}(n)$ denotes the estimated transmitted signal produced by the neural network equaliser after processing the received distorted signal. The parameter N indicates the total number of samples in the dataset used for training or evaluation, determining the statistical significance of the error measurement.

Selecting an appropriate loss function is vital for training ANN–based equalisers, as it directly influences the learning process and ultimately the system's performance. In this work, the MSE was chosen as the principal loss function, aligning well with the regression-style nature of the equalisation task, where the goal is to minimise the discrepancy between predicted and actual transmitted symbols. Moreover. it has been thoroughly tested in other works and demonstrated consistently led to rapid convergence and reliable symbol-level reconstruction [37]. Moreover, we evaluated the effectiveness of this choice by monitoring the BER throughout training and inference. Although BER is often viewed as the ultimate performance metric in digital communication systems, it is difficult to integrate directly into the gradient-based optimisation process. This arises because BER is a discrete measure that does not lend itself easily to calculus-based approaches. Consequently, an approach that combines MSE for backpropagation and BER as an ongoing performance indicator has been followed [37,100].

To further validate this dual-metric strategy, we explored the feasibility of incorporating a BER-based objective at select stages of training for parameter fine-tuning. BER-based loss functions can offer a more direct optimisation toward the communication-centric performance goal, their implementation demands a careful trade-off between computational complexity and gradient smoothness. In practice, in this project it was confirmed that MSE provides stable and efficient learning, yet incorporating periodic checks of BER enables an effective gauge of real-world performance. Through this procedure, it was possible to ensure that the system does not merely excel in reducing MSE, but also achieves minimal BER under realistic operating conditions [37].

2.5.3 Backpropagation in ANN-based Equalisers

Understanding backpropagation in the context of optical channel equalisation requires recognising that we are essentially solving an inverse problem: given the distorted received signal, we must learn to recover the original transmitted signal. This inverse mapping is inherently complex because optical channels exhibit both memory effects and nonlinear distortions that create intricate relationships between transmitted and received symbols. The backpropagation algorithm provides the mathematical framework for automatically discovering these relationships through data-driven learning [37,100].

The fundamental challenge lies in computing how each of the potentially millions of parameters in a deep neural network contributes to the final equalisation error. Without an efficient method to compute these sensitivities, training would require evaluating the network's performance with perturbed parameters for each weight individually—a computationally prohibitive approach. Backpropagation solves this challenge through the elegant application of the chain rule of calculus, enabling simultaneous computation of all parameter gradients in a single forward-backward pass [106].

To understand the mathematical foundation, consider a multi-layer neural network equaliser with D layers. The forward propagation through the network can be expressed as:

$$\mathbf{z}^{(1)} = \mathbf{W}^{(1)}\mathbf{y} + \mathbf{b}^{(1)} \tag{2.27}$$

$$\mathbf{a}^{(1)} = \sigma(\mathbf{z}^{(1)}) \tag{2.28}$$

$$\vdots \tag{2.29}$$

$$\mathbf{z}^{(D)} = \mathbf{W}^{(D)}\mathbf{a}^{(D-1)} + \mathbf{b}^{(D)} \tag{2.30}$$

$$\hat{\mathbf{x}} = \mathbf{z}^{(D)} \tag{2.31}$$

where $\mathbf{y}$ is the input received signal, $\mathbf{W}^{(i)}$ and $\mathbf{b}^{(i)}$ are the weight matrix and bias vector for layer i, $\mathbf{z}^{(i)}$ is the pre-activation output, $\mathbf{a}^{(i)}$ is the post-activation output, $\sigma(\cdot)$ is the activation function, and $\hat{\mathbf{x}}$ is the estimated transmitted signal.

The backpropagation algorithm computes gradients by working backwards through the network, applying the chain rule to decompose the complex gradient computation into simpler local derivatives. For the loss function $L(\mathbf{x}, \hat{\mathbf{x}})$, the gradient with respect to the weights in layer ℓ is:

$$\frac{\partial L}{\partial \mathbf{W}^{(\ell)}} = \frac{\partial L}{\partial \mathbf{z}^{(\ell)}} \frac{\partial \mathbf{z}^{(\ell)}}{\partial \mathbf{W}^{(\ell)}} = \boldsymbol{\delta}^{(\ell)} (\mathbf{a}^{(\ell-1)})^T \tag{2.32}$$

where $\boldsymbol{\delta}^{(\ell)} = \frac{\partial L}{\partial \mathbf{z}^{(\ell)}}$ represents the error signal at layer ℓ. The key insight is that these error signals can be computed recursively:

$$\boldsymbol{\delta}^{(D)} = \frac{\partial L}{\partial \hat{\mathbf{x}}} \tag{2.33}$$

$$\boldsymbol{\delta}^{(\ell)} = (\mathbf{W}^{(\ell+1)})^T \boldsymbol{\delta}^{(\ell+1)} \odot \sigma'(\mathbf{z}^{(\ell)}) \tag{2.34}$$

where $\odot$ denotes element-wise multiplication and $\sigma'(\cdot)$ is the derivative of the activation function.

Weight Sensitivity Analysis in Optical Equalisation: Understanding how different layers contribute to the equalisation performance reveals important insights about the network's learning process. In optical channel equalisation, early layers typically learn to capture linear dispersion effects, while deeper layers focus on nonlinear distortion compensation. This hierarchical learning can be analysed through the magnitude of gradients at different layers.

The sensitivity of the loss function to weights in layer ℓ can be quantified through the gradient norm:

$$S^{(\ell)} = \| \frac{\partial L}{\partial \mathbf{W}^{(\ell)}} \|_F = \| \boldsymbol{\delta}^{(\ell)} (\mathbf{a}^{(\ell-1)})^T \|_F \tag{2.35}$$

where $\| \cdot \|_F$ denotes the Frobenius norm. This sensitivity metric reveals which layers are most critical for reducing equalisation error and guides architectural design decisions.

Gradient Flow Analysis: The vanishing and exploding gradient problems become particularly challenging in optical equalisation networks due to the need for deep architectures to capture long memory effects. The gradient flow through the network can be analysed by examining the recursive relationship:

$$\| \boldsymbol{\delta}^{(\ell)} \| = \| (\mathbf{W}^{(\ell+1)})^T \boldsymbol{\delta}^{(\ell+1)} \odot \sigma'(\mathbf{z}^{(\ell)}) \| \tag{2.36}$$

When $\| \mathbf{W}^{(\ell+1)} \| \cdot \| \sigma'(\mathbf{z}^{(\ell)}) \| < 1$ for multiple consecutive layers, gradients vanish exponentially. Conversely, when this product exceeds 1, gradients explode. In optical systems, this problem is exacerbated by the need to process sequences with long temporal dependencies.

Mitigation Strategies for Optical Equalisers: 1. **Residual Connections**: Adding skip connections that allow gradients to flow directly to earlier layers:

$$\mathbf{a}^{(\ell)} = \sigma(\mathbf{W}^{(\ell)} \mathbf{a}^{(\ell-1)} + \mathbf{b}^{(\ell)}) + \mathbf{a}^{(\ell-1)} \tag{2.37}$$

2. **Gradient Clipping**: Limiting gradient magnitude to prevent explosion:

$$\nabla_\theta L \leftarrow \begin{cases} \nabla_\theta L & \text{if } \|\nabla_\theta L\| \leq \tau \\ \frac{\tau}{\|\nabla_\theta L\|} \nabla_\theta L & \text{otherwise} \end{cases} \tag{2.38}$$

3. **Normalised Initialisation**: Setting initial weights to maintain gradient variance across layers, such as Xavier initialisation:

$$W_{ij} \sim \mathcal{N}(0, \frac{2}{n_{in} + n_{out}}) \tag{2.39}$$

Computational Complexity in Optical Systems: The computational demands of backpropagation scale quadratically with network width and linearly with depth. For a network with D layers, each containing n neurons, the computational complexity is approximately:

$$\mathcal{O}_{backprop} = \mathcal{O}(D \cdot n^2 \cdot B) \tag{2.40}$$

where B is the batch size. In optical systems operating at symbol rates of 32-64 GBaud, this computational burden becomes significant, often requiring specialised hardware implementations or algorithmic approximations [103, 107].

Once computed, gradients are used to update the parameters. The general update rule for the parameters θ using gradient descent is given by [100]:

$$\theta^{(t+1)} = \theta^{(t)} - \eta \frac{\partial L}{\partial \theta} \tag{2.41}$$

In this fundamental update equation, $\theta^{(t)}$ represents the parameter value at iteration t, while $\theta^{(t+1)}$ denotes the updated parameter value after applying the gradient-based correction. The learning rate η serves as a hyperparameter that controls the step size of the update, determining how aggressively the parameters are modified in response to the computed gradients. The term $\frac{\partial L}{\partial \theta}$ represents the gradient of the loss function with respect to the parameter θ, providing the direction and magnitude of the required parameter adjustment to minimise the equalisation error. The update rule is a crucial component of the backpropagation algorithm, dictating how the weights and biases are adjusted during the training process to minimise the loss function. Therefore, it determines how ANNs learn from data. On the other hand, the learning rate η plays a critical role in the update rule, as it greatly affects the training stability. It determines how large a step is taken in the direction of the gradient. A learning rate that is too high can cause the training process to become unstable and diverge, while a learning rate that is too low can result in slow convergence, making the training process inefficient [100].

While the basic update rule is derived from simple gradient descent, various optimisation algorithms build upon this foundation to improve convergence and performance. Often, techniques like learning rate schedules or adaptive learning rates (e.g., Adam optimiser) are used to dynamically adjust the learning rate during training [100, 108]. Stochastic Gradient Descent (SGD) represents the foundational approach, updating parameters using a single or a small batch of training examples, which leads to faster but noisier updates that can help escape local minima. Momentum methods accelerate convergence by adding a fraction of the previous update to the current update, effectively helping the optimisation process navigate ravines in the loss landscape and maintain consistent progress toward the minimum. RMSprop adapts the learning rate for each parameter based on a moving average of the squared gradients, reducing the learning rate for parameters with large gradients and increasing it for those with small gradients, thereby achieving more balanced parameter updates. Adam (Adaptive Moment Estimation) combines the benefits of both momentum and RMSprop by maintaining moving averages of both the gradients and their squared values, resulting in an optimisation algorithm that is both robust and efficient for a wide range of optical equalisation problems [100].

Among the algorithms listed above, the Adaptive Moment Estimation (Adam) optimiser, which combines the advantages of both AdaGrad and RMSProp algorithms to handle sparse gradients on noisy problems, is the most popular one [100, 109].

The Adam optimiser updates parameters based on adaptive estimates of lower-order moments, using the following equations [100, 110]:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1)\nabla_\theta L(\theta_t) \tag{2.42}$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2)(\nabla_\theta L(\theta_t))^2 \tag{2.43}$$

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t} \tag{2.44}$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t} \tag{2.45}$$

$$\theta_{t+1} = \theta_t - \alpha \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon} \tag{2.46}$$

In these equations, m_t represents the first moment estimate that tracks the mean of the gradients, while v_t denotes the second moment estimate that monitors the uncentred variance of the gradients. The exponential decay rates β_1 and β_2 control the exponential moving averages for the moment estimates, with typical values of $\beta_1 = 0.9$ and $\beta_2 = 0.999$ providing effective balance between responsiveness and stability. The bias-corrected estimates $\hat{m}_t$ and $\hat{v}_t$ compensate for the initialisation bias that occurs during the early stages of training when the moment estimates are close to zero. Finally, ϵ serves as a small constant (typically 10^{-8}) to prevent division by zero in the parameter update step, ensuring numerical stability throughout the optimisation process.

2.6 Hyperparameter Optimisation for ANN-based Equalisers

The performance of ANNs-based equalisers in optical communication systems is highly dependent on the selection of appropriate hyperparameters. These hyperparameters include architectural choices such as the number of layers and neurons, as well as training parameters like learning rate and batch size. Optimising these hyperparameters is crucial for achieving the best possible performance. Moreover, it has a direct impact on the CC of the final solution. There are several approaches to hyperparameter optimisation, differentiated by how the different combination of hyperparameters are searched in order to find the appropriate one. The simplest one is grid search, which involves exhaustively searching through a manually specified subset of the hyperparameter space. Thus, a set of possible values for each hyperparameter is specified and the model's performance for every combination of these hyperparameter values is tested. The primary advantage of grid search is its simplicity and guaranteed coverage of the specified search space. However, it suffers from the curse of dimensionality as it becomes computationally intractable when the number of hyperparameters increases. In order to try to avoid this disadvantage, random search was proposed. It involves randomly sampling hyperparameters from a defined distribution, instead of exhaustively evaluating all combinations. The number of evaluations is set beforehand, and each evaluation uses a random combination of hyperparameter values. While this method has been shown to be more efficient than grid search, especially when only a few hyperparameters significantly impact model performance, it may miss optimal configurations in dense regions of the search space [103, 111]. BO is a more sophisticated approach that uses probabilistic models to guide the search for optimal hyperparameters. The optimisation process can be described as follows [112]:

1. Define a prior distribution over the function mapping hyperparameters to model performance.
2. Iteratively:
   (a) Use an acquisition function to determine the next set of hyperparameters to evaluate.
   (b) Evaluate the model performance with these hyperparameters.
   (c) Update the posterior distribution based on the observed performance.
3. Select the hyperparameters that yielded the best performance.

A common choice for the probabilistic model in BO is the Gaussian Process (GP). The acquisition function balances exploration of unknown regions of the hyperparameter space with exploitation of regions known to perform well. Popular acquisition functions include Expected Improvement (EI) and Upper Confidence Bound (UCB). BO can find good hyperparameters in fewer evaluations than grid or random search. However, it can be computationally expensive for high-dimensional hyperparameter spaces. For ANN-based equalisers in optical systems, BO has shown promising results, achieving superior performance compared to manual tuning [103, 112, 113].

Recent advancements have enabled gradient-based optimisation of certain hyperparameters, which attempts to compute the gradient of the model's performance with respect to its hyperparameters. This allows for the use of efficient gradient descent algorithms to optimise hyperparameters. For example, Differentiable Architecture Search (DARTS) is one of such methods and relaxes the discrete architecture search space to a continuous one, allowing for efficient gradient-based optimisation. These methods can be very efficient, especially for continuous hyperparameters. They can potentially find optimal hyperparameters much faster than other methods. However, they are primarily suitable for architectural hyperparameters and less so for training hyperparameters like learning rate and batch size, which do not have a clear gradient with respect to model performance [111, 114].

The selection of hyperparameter optimisation techniques must account for the specific constraints and requirements of the optical system. While Neural Architecture Search (NAS) is an active research area aiming to discover efficient neural network architectures, the high CC associated with evaluating numerous candidate models can be prohibitive, especially when considering the vast hyperparameter space and the expense of performance evaluations [115].

To illustrate the trade-offs involved, consider the decision regarding the number of layers in a neural network. Increasing the number of layers can enhance the model's capacity to capture complex patterns, potentially improving performance. However, this also leads to deeper architectures that are more computationally intensive to train and may require larger datasets to prevent overfitting. Conversely, opting for a shallower architecture might reduce computational demands and training time but could compromise the model's ability to learn intricate data representations [116].

Recent advancements in NAS have sought to address these challenges by developing more efficient search strategies. For instance, differentiable architecture search methods, such as DARTS, transform the discrete search space into a continuous one, enabling the use of gradient-based optimisation techniques to expedite the search process [115]. Additionally, multi-objective optimisation approaches have been introduced to balance trade-offs between model accuracy and resource consumption, facilitating the deployment of machine learning models in resource-constrained environments [117].

Despite these innovations, the computational expense of NAS remains a significant consideration. Techniques like weight-sharing among candidate architectures have been proposed to mitigate this issue, allowing multiple architectures to be evaluated without training each one from scratch. Nonetheless, the complexity of the search space and the necessity for extensive computational resources continue to pose challenges in the practical application of NAS [115].

Therefore, while NAS offers promising avenues for discovering efficient neural network architectures, it is crucial to carefully consider the trade-offs between model complexity, computational cost, and performance. Tailoring the hyperparameter optimisation approach to the specific constraints of the optical system is essential to achieve an optimal balance between these factors [112, 118].

2.7 Data Requirements for Training ANN-based Equalisers

The effectiveness of ANNs in mitigating nonlinear impairments in optical communication systems is heavily dependent on the quality and quantity of training data [101]. Therefore, a desirable dataset must possess several key characteristics that ensure robust and generalisable equaliser performance [103].

Representativeness constitutes the most fundamental requirement, demanding that the training data adequately represent the full range of signal distortions encountered in the optical channel. This encompasses not only the primary Kerr effect but also its complex interplay with CD and PMD, while accounting for various fibre types, transmission distances, and power levels that may be encountered in practical deployments. Volume represents another critical consideration, as large datasets are typically required to capture the complexity of nonlinear interactions, with dataset sizes often ranging from 10^{18} to 10^{20} symbols depending on the channel complexity and desired performance targets [103].

Diversity ensures that the dataset includes a comprehensive set of modulation formats, symbol rates, and channel conditions to guarantee the generalizability of the trained equaliser across different operational scenarios. Finally, balanced representation requires that the dataset maintains equilibrium across different impairment scenarios, preventing bias toward specific channel conditions and ensuring that the equaliser can properly handle the full spectrum of distortions it may encounter in real-world deployments [37].

Several methods can be employed to acquire the necessary training data [101, 103]:

1. Experimental measurements: Direct measurement of transmitted and received signals in a real optical system provides the most accurate representation of channel impairments. However, this method can be time-consuming and expensive.

2. Numerical simulations: SSFM simulations of the NLSE can generate large volumes of data efficiently. While not perfectly accurate, this approach offers flexibility in exploring various system parameters.

3. Hybrid approaches: Combining experimental measurements with data augmentation techniques or model-based simulations can provide a good balance between accuracy and dataset size.

On the other hand, there are several relevant points that must be considered when manageing data to train ANN-based equalisers in optical communication systems [7, 29, 37, 103]:

1. High Dimensionality: Optical signals in modern communication systems are characterised by high symbol rates, multiple polarisations, and WDM. This results in high-dimensional data spaces that require careful consideration in data generation and preprocessing.

2. Nonlinear Complexity: The NLSE and its variants (e.g., the Manakov equation for dual-polarisation systems) that govern signal propagation in optical fibres are complex and computationally intensive to solve. Accurately capturing these nonlinear dynamics in the training data is crucial but challenging.

3. Stochastic Noise Processes: Optical systems are subject to various noise sources, including ASE noise, shot noise, and thermal noise. Properly incorporating these stochastic processes into the training data is essential for developing robust ANNs.

4. Memory Effects: Fibre nonlinearities introduce memory effects, where the distortion of a given symbol depends on neighbouring symbols. Capturing these inter-symbol interactions in the dataset requires careful consideration of sequence lengths and data structure.

5. Varying Transmission Conditions: Optical networks operate under diverse conditions, including different fibre types, span configurations, and power levels. The training data must encompass this variability to ensure the ANN's adaptability.

6. Experimental Limitations: While numerical simulations can generate vast amounts of data, experimental data collection is often limited by equipment availability and measurement time. Balancing between simulated and experimental data is a key challenge.

7. Data Augmentation Constraints: Unlike in some other fields (e.g., image processing), data augmentation in optical communications must adhere to the physical constraints of the optical channel, limiting the applicable augmentation techniques.

Numerical simulations are a powerful tool when training ANN-based equalisers in optical communication systems, as they make it possible to create extensive and diverse datasets. These simulations are based on the Manakov equation [29], a vectorial extension of the NLSE that accurately describes the propagation of light in optical fibres. It is worth mentioning its variant known as the Manakov equation, which focuses on dual-polarisation systems. The SSFM is the numerical technique employed to efficiently solve them.

The Manakov equation for a dual-polarisation signal can be expressed as:

$$\frac{\partial \mathbf{A}}{\partial z} = -\frac{\alpha}{2}\mathbf{A} - i\frac{\beta_2}{2}\frac{\partial^2 \mathbf{A}}{\partial T^2} + i\gamma\left(\frac{8}{9}|\mathbf{A}|^2\mathbf{A}\right) \tag{2.47}$$

where $\mathbf{A} = \begin{bmatrix} A_x \\ A_y \end{bmatrix}$ is the complex envelope of the optical field in the two orthogonal polarisation states (horizontal and vertical), z is the propagation distance, T is the retarded time, α is the fibre attenuation coefficient, β_2 is the GVD parameter, and γ is the nonlinear coefficient. The factor $\frac{8}{9}$ accounts for the averageing over the fast polarisation rotations.

This equation models how the complex field envelope $\mathbf{A}$ evolves as it propagates through the fibre, considering the effects of loss (α), dispersion (β_2), and nonlinearity (γ). The term $\frac{\beta_2}{2}\frac{\partial^2 \mathbf{A}}{\partial T^2}$

corresponds to dispersion, while the term $\gamma\left(\frac{8}{9}|\mathbf{A}|^2\mathbf{A}\right)$ represents the nonlinear interaction between the two polarisation modes.

To solve this equation numerically, the SSFM is employed, which alternates between solving the linear and nonlinear parts of the equation in small steps along the propagation distance z [55].

The linear effects, including dispersion and attenuation, are handled in the frequency domain:

$$\tilde{\mathbf{A}}(z+h,\omega) = \exp\left(-\frac{\alpha h}{2} + i\frac{\beta_2\omega^2}{2}h\right)\tilde{\mathbf{A}}(z,\omega) \tag{2.48}$$

where $\tilde{\mathbf{A}}$ denotes the FFT of $\mathbf{A}$, h is the step size, and ω is the angular frequency.

The nonlinear effects are addressed in the time domain:

$$\mathbf{A}(z+h,T) = \exp\left(i\gamma\frac{8}{9}|\mathbf{A}(z,T)|^2 h\right)\mathbf{A}(z,T) \tag{2.49}$$

This step models the nonlinear phase shift induced by the intensity of the light.

The parameters α, β_2, and γ are crucial for determining the behaviour of the signal as it propagates through the fibre, each governing different aspects of the physical propagation process. The attenuation coefficient α governs the exponential decay of power along the fibre length, determining how quickly the signal power diminishes due to material absorption and scattering losses. The dispersion parameter β_2 describes how different frequency components of the pulse spread out in time, leading to pulse broadening that creates intersymbol interference and limits the achievable data rates over long distances. The nonlinear coefficient γ controls the strength of nonlinear effects, such as SPM and XPM, which become increasingly significant at higher signal powers and can either enhance or degrade system performance depending on the specific operating conditions and compensation strategies employed.

In numerical simulations, these parameters are varied to model different transmission scenarios, such as using different types of optical fibres (e.g., SSMF or Large Effective Area Fibre (LEAF)) or varying the launch power to explore linear and nonlinear regimes.

The fibre hyperparameters α, β_2, and γ play crucial roles in determining the nature and strength of the signal impairments. The attenuation coefficient α, typically around 0.2 dB/km for SSMF at 1550 nm, determines the rate of power decay along the fibre. The dispersion coefficient β_2, often expressed in terms of the dispersion parameter $D \approx -\frac{2\pi c}{\lambda^2}\beta_2$, characterises the temporal broadening of pulses. For SSMF, D is approximately 17 ps/(nm·km) at 1550 nm. The nonlinear coefficient γ, typically around 1.3 $W^{-1}km^{-1}$ for SSMF, governs the strength of nonlinear effects [74, 119].

In the simulations employed in this Thesis, we systematically vary these parameters to capture a wide range of transmission scenarios. For instance, we might consider dispersion-shifted fibres with D close to zero, or LEAFs with reduced γ values. The interplay between dispersion and nonlinearity is particularly important, as it determines the dominant impairment regime and the effectiveness of various equalisation strategies.

The transmission distance is another critical parameter in our simulations. We typically consider multi-span systems, where each span consists of a length of fibre followed by an optical amplifier. The span length is usually set between 50 km and 100 km. At each amplifier, we model the addition of ASE noise as a complex AWGN process with power spectral density:

$$S_{ASE} = n_{sp} h\nu (G - 1) \tag{2.50}$$

where n_{sp} is the spontaneous emission factor, $h\nu$ is the photon energy, and G is the amplifier gain.

The modulation format and symbol rate are other key transmitter-side parameters that significantly influence the signal's susceptibility. In this Thesis, QAM constellations (16-QAM, 32-QAM and higher), and symbol rates (from 32 to 64 GBaud) are employed. For each format, we generate Pseudo-Random Binary Sequences (PRBSs) of sufficient length to avoid periodicity issues, typically using a PRBS order of 15 or higher.

Pulse shaping is applied to the symbol sequences using a Root-Raised Cosine (RRC) filter. The roll-off factor of this filter, typically is set to 0.2, affects the signal's SE and its robustness against inter-symbol interference. The shaped pulses are then oversampled, usually by a factor of 4 or 8, to accurately capture high-frequency components during propagation.

The launch power per channel is a key parameter that determines the balance between nonlinear effects and noise. We typically sweep the launch power from -5 dBm to +5 dBm per channel, encompassing both the linear and highly nonlinear regimes. This power sweep is crucial for capturing the nonlinear threshold and the optimal operating point of the system.

At the receiver end, our simulations include matched filtering, sampling, and optional linear equalisation to compensate for CD.

In conclusion, the previously described numerical simulation framework provides a flexible and comprehensive approach to generating the diverse and extensive datasets required for training effective ANN-based equalisers. By carefully controlling a wide range of parameters and incorporating detailed models of various impairments, we create datasets that closely mimic the complexities of real-world optical communication systems, thereby enabling the development of robust and effective equalisers capable of addressing the myriad challenges in modern high-capacity optical networks.

Chapter 3

Performance Optimisation and Complexity Reduction of Neural Network-based Equalisers for Optical Communication Systems

Having established the fundamental principles of ANN-based equalisation in Chapter 2, we now turn to the critical question of architecture selection and optimisation. The choice of neural network architecture directly determines both the equalisation performance and computational complexity, making it a central concern for practical deployment. This chapter provides comprehensive analysis of the major architectural families employed for optical channel equalisation, examining their mathematical foundations, implementation considerations, and performance trade-offs.

The architectural landscape for optical equalisation encompasses three major paradigms: feed-forward architectures (MLPs and CNNs) that process temporal windows in parallel, recurrent architectures (RNNs, LSTMs, GRUs) that maintain internal state across time, and specialised designs (complex-valued networks, bidirectional processing) that exploit specific properties of optical signals. Understanding these architectures in depth is essential for making informed optimisation decisions throughout the remainder of this chapter.

This chapter focuses on optimising the performance of ANN-based equalisers for optical channel equalisation, specifically targeting the balance between accuracy, computational complexity, and resource constraints. We present a comprehensive analysis of the trade-offs between performance and complexity in ANNs tailored for optical communication channels, alongside detailed architectural considerations for various neural network types including feed-forward, convolutional, and recurrent designs.

We introduce several model compression techniques—pruning, quantisation, and weight clustering—and demonstrate their application to reduce computational complexity (CC) and memory demands. Each technique is first presented theoretically, then evaluated empirically across multiple transmission scenarios. These compression methods are further enhanced through BO, enabling fine-tuned trade-offs between performance and efficiency.

The chapter also covers practical deployment considerations, providing empirical measurements of inference time, energy consumption, and resource utilisation on various hardware platforms including Raspberry Pi 4 and Nvidia Jetson Nano. We demonstrate how compressed neural networks can be effectively deployed in resource-constrained environments while maintaining acceptable equalisation performance.

3.1 Introduction

In previous sections of this Thesis, we have established the necessity for a new generation of optical communication networks driven by the soaring demands of data-intensive applications and the exponential rise in global internet traffic. To meet these challenges, advanced DSP techniques are being developed to effectively mitigate impairments in optical transmission systems, with a special focus on nonlinear distortions. ANNs have emerged as a powerful tool within this new wave of DSP solutions for nonlinearity compensation, owing to their exceptional ability to model complex nonlinear relationships.

This section revisits the chapter's goals and defines the evaluation scope.

1. Examine various neural network architectures—feed-forward, convolutional, and recurrent—and their suitability for optical channel equalisation
2. Introduce metrics for evaluating the computational complexity of neural network equalisers, establishing a standardised framework for comparison
3. Evaluate the performance-complexity trade-offs of various ANN architectures for optical channel equalisation across multiple transmission scenarios
4. Investigate the effectiveness of compression techniques (pruning, quantisation, and weight clustering) in reducing the computational complexity of ANN-based equalisers
5. Apply Bayesian Optimisation to enhance the efficiency of these compression techniques, enabling optimal trade-offs between performance and complexity
6. Assess the deployment of optimised ANN-based equalisers on edge computing devices, with detailed measurements of inference time, energy consumption, and resource utilisation
7. Explore specialised applications including low-complexity RNN-based equalisers for systems with SOAs

8. Investigate a synergistic approach combining OPC with ANN-based equalisers, demonstrating the benefits of hardware-software co-design
9. Discuss the broader implications of these findings for the future of optical communication systems

The chapter begins with a detailed examination of neural network architectures (Sections 3.2 and 3.3), establishing the foundation for our complexity reduction strategies. We then introduce computational complexity metrics (Section 3.4) that will be used throughout the chapter to evaluate the efficiency of various approaches. Moreover, we discuss different compression techniques—pruning, quantisation, and weight clustering—alongside their enhancement through Bayesian Optimisation.

Following this theoretical foundation, we conduct extensive empirical analyses to evaluate the effectiveness of these compression techniques across diverse transmission scenarios. These analyses provide insights into the practical trade-offs between performance and complexity, guiding the selection of appropriate compression strategies for different applications.

The latter part of the chapter addresses the practical deployment of ANN-based equalisers on resource-constrained hardware, providing concrete measurements of inference time, energy consumption, and resource utilisation. We demonstrate how our compression techniques enable efficient real-time processing on devices like Raspberry Pi 4 and Nvidia Jetson Nano, which is crucial for practical applications.

Finally, we explore two innovative approaches that extend beyond pure digital signal processing: low-complexity RNN-based equalisers for systems employing SOAs and the synergistic integration of OPC with ANN-based equalisers. These approaches illustrate the potential for hybrid solutions that combine digital and optical techniques, potentially offering superior performance-complexity trade-offs compared to purely digital approaches.

By addressing both the theoretical and practical dimensions of ANN-based equalisers and exploring innovative hybrid approaches, this chapter offers valuable insights for researchers and engineers engaged in the development and implementation of advanced DSP techniques for next-generation optical networks.

3.2 Feed-Forward Architectures for Optical Channel Equalisation

Feed-forward neural networks represent a fundamental class of ANNs architectures that have found significant application in optical channel equalisation. These networks are characterised by their unidirectional flow of information from the input layer through one or more hidden layers to the output layer, without any feedback connections [100, 103].

The basic structure of a feed-forward network can be described as follows [100]:

$$\mathbf{y} = f_L(\mathbf{W}_L f_{L-1}(\mathbf{W}_{L-1} \cdots f_1(\mathbf{W}_1\mathbf{x} + \mathbf{b}_1) \cdots + \mathbf{b}_{L-1}) + \mathbf{b}_L) \tag{3.1}$$

where $\mathbf{x}$ is the input vector, $\mathbf{y}$ is the output vector, $\mathbf{W}_i$ and $\mathbf{b}_i$ are the weight matrix and bias vector for the i-th layer respectively, and $f_i(\cdot)$ is the activation function for the i-th layer. In the context of optical channel equalisation, the input $\mathbf{x}$ typically represents a sequence of received symbols, while the output $\mathbf{y}$ represents the estimated transmitted symbols.

For all architectures discussed in this chapter, the input vector $\mathbf{x}$ represents a temporal window of received signal samples centred at the symbol of interest. This window spans $2M+1$ consecutive symbol periods, where M denotes the memory depth parameter that captures past and future symbol dependencies. The total input dimension is therefore $(2M+1)\times d_{in}$, where d_{in} represents the number of input features per symbol. For dual-polarisation systems, $d_{in} = 4$ corresponding to the in-phase and quadrature components from both horizontal and vertical polarisations (I_H, Q_H, I_V, Q_V), while single-polarisation systems use $d_{in} = 2$ with only (I, Q) components. The memory depth M is typically optimised during hyperparameter tuning and commonly ranges from 5 to 20 symbols depending on channel characteristics and transmission distance [103]. This input structure enables the neural network to capture temporal dependencies from chromatic dispersion, inter-symbol interference, and polarisation coupling effects, as detailed in Chapter 2. The choice of M directly impacts both equalisation performance and computational complexity, as the number of input features grows linearly with M while network parameters typically scale quadratically for fully-connected architectures.

3.2.1 MLP

An MLP consists of an input layer, one or more hidden layers, and an output layer. Each layer comprises multiple neurons, with each neuron connected to all neurons in the adjacent layers. The mathematical representation of an MLP with L layers can be expressed as [100]:

$$\mathbf{a}^{(l)} = f^{(l)}(\mathbf{z}^{(l)}) = f^{(l)}(\mathbf{W}^{(l)}\mathbf{a}^{(l-1)} + \mathbf{b}^{(l)}) \tag{3.2}$$

where $\mathbf{a}^{(l)}$ is the activation of the l-th layer, $\mathbf{W}^{(l)}$ is the weight matrix, $\mathbf{b}^{(l)}$ is the bias vector, and $f^{(l)}(\cdot)$ is the activation function of the l-th layer.

Following the general approach described in Section 2.5, training an MLP for optical channel equalisation involves minimising a loss function that quantifies the difference between the network's output and the desired equalised signal. The backpropagation algorithm is used to compute the gradients of the loss function with respect to the network parameters.

The key steps in the backpropagation process are:

1. Forward Pass: Compute the network's output for a given input.
2. Loss Calculation: Evaluate the loss function, typically mean squared error (MSE) for regression tasks.
3. Backward Pass: Compute the gradients of the loss with respect to the model's parameters.

4. Parameter Update: Adjust the model's parameters using an optimisation algorithm, such as stochastic gradient descent (SGD).

The forward propagation process in an MLP for optical channel equalisation can be described as follows [100]:

1. Input layer: The received signal samples are fed into the input layer.

$$\mathbf{a}^{(0)} = \mathbf{x} \tag{3.3}$$

2. Hidden layers: The signal propagates through the hidden layers.

$$\mathbf{z}^{(l)} = \mathbf{W}^{(l)}\mathbf{a}^{(l-1)} + \mathbf{b}^{(l)} \tag{3.4}$$

$$\mathbf{a}^{(l)} = f^{(l)}(\mathbf{z}^{(l)}) \tag{3.5}$$

3. Output layer: The final layer produces the equalised signal.

$$\hat{\mathbf{y}} = \mathbf{a}^{(L)} = f^{(L)}(\mathbf{W}^{(L)}\mathbf{a}^{(L-1)} + \mathbf{b}^{(L)}) \tag{3.6}$$

Common activation functions used in MLPs for optical channel equalisation include [100]:

1. ReLU:

$$f(x) = \max(0, x) \tag{3.7}$$

ReLU is widely used due to its simplicity and effectiveness in mitigating the vanishing gradient problem.

2. Hyperbolic Tangent (tanh):

$$f(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{3.8}$$

Hyperbolic Tangent (Tanh) is often used in the hidden layers as it can capture both positive and negative nonlinearities, clipping the activation range to -1 and 1, as well as providing a very smooth gradient.

3. Linear activation:

$$f(x) = x \tag{3.9}$$

Linear activation is typically used in the output layer for equalisation when it is approached as a regression task [120].

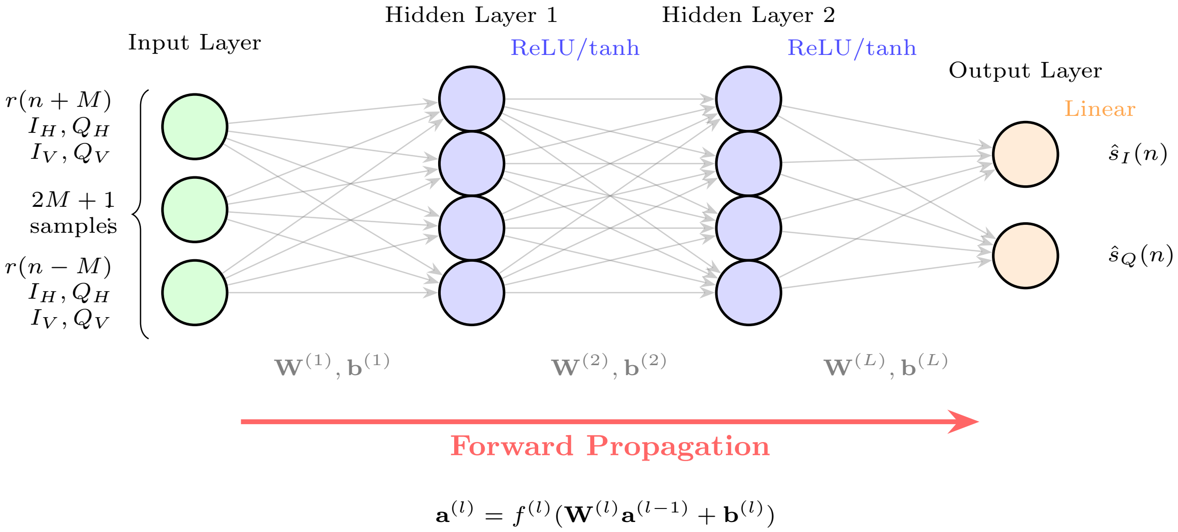


Figure 3.1: MLP architecture for optical channel equalisation. The input layer receives a temporal window of $2M+1$ received signal samples $r(n+M), \ldots, r(n-M)$, where each sample contains dual-polarisation components (I_H, Q_H, I_V, Q_V) representing in-phase and quadrature components from horizontal and vertical polarisations. Two hidden layers with nonlinear activation functions (ReLU or tanh) extract features and model channel nonlinearities, with each neuron fully connected to all neurons in adjacent layers through weight matrices $\mathbf{W}^{(l)}$ and bias vectors $\mathbf{b}^{(l)}$. The output layer uses linear activation to produce the estimated transmitted symbol components $\hat{s}_I(n)$ and $\hat{s}_Q(n)$. The forward propagation follows the equation $\mathbf{a}^{(l)} = f^{(l)}(\mathbf{W}^{(l)}\mathbf{a}^{(l-1)} + \mathbf{b}^{(l)})$, enabling the network to learn the inverse channel mapping through backpropagation training.

MLPs can effectively model and compensate for nonlinear distortions in optical channels and jointly compensate for multiple impairments in a single structure. Moreover, the feed-forward structure allows for efficient parallel processing [52]. Nevertheless, standard MLPs do not inherently capture long-term dependencies in the signal, making it necessary the use of deep MLPs which can be computationally intensive for real-time processing in high-speed systems [103].

3.2.2 CNNs

CNNs are another popular feed-forward architecture used in the context of optical channel equalisation due to their ability to extract local features and exploit spatial correlations in the input data [103].

The CNN architecture typically consists of three main components [100, 121]:

1. Convolutional layers
2. Pooling layers
3. Fully connected layers

The core building block of a CNN is the convolutional layer. The computation taking place in this layer can be expressed as [100, 121]:

$$y_l^{(i)} = f\left(\sum_{j=1}^{M_{l-1}}\sum_{p=1}^{F_l}\sum_{q=1}^{F_l} w_{l,j}^{(i)}(p,q)\cdot x_{l-1}^{(j)}(s+p-1,t+q-1)+b_l^{(i)}\right) \tag{3.10}$$

where:

- $y_l^{(i)}$ is the output of the i-th feature map in the l-th layer
- $x_{l-1}^{(j)}$ is the j-th feature map in the $(l-1)$-th layer
- $w_{l,j}^{(i)}$ is the weight kernel connecting the j-th feature map in layer $(l-1)$ to the i-th feature map in layer l
- $b_l^{(i)}$ is the bias term
- $f(\cdot)$ is the activation function
- F_l is the size of the convolutional kernel
- M_{l-1} is the number of feature maps in layer $(l-1)$

In addition to the convolutional layer, another important component of a CNN are the pooling layers, which decrease the spatial dimensions of the feature maps, providing translation invariance and reducing CC. The most common pooling operation in optical channel equalisation is max pooling [103]:

$$y_l^{(i)}(s,t) = \max_{p,q\in R_{s,t}} x_{l-1}^{(i)}(p,q) \tag{3.11}$$

where $R_{s,t}$ is the pooling region centred at position (s,t).

The final layers of a CNN are typically fully connected, similar to those in a traditional MLP. These layers combine the high-level features extracted by the convolutional and pooling layers to produce the final output [103].

The previously mentioned and explained layers are combined following the schema [121]:

$$\hat{x} = f_{\text{FC}}(f_{\text{pool}}(f_{\text{conv}}(y))) \tag{3.12}$$

Therefore, a typical CNN architecture for optical channel equalisation can be represented as [103, 121]:

$$\text{Input} \rightarrow [\text{Conv} \rightarrow \text{ReLU} \rightarrow \text{Pool}]^n \rightarrow [\text{FC}]^m \rightarrow \text{Output} \tag{3.13}$$

where n is the number of convolutional-activation-pooling blocks and m is the number of fully connected layers.

In the context of optical channel equalisation, the previous terms can be mapped to [103]:

- y is the received signal
- f_{conv} represents the convolutional layers
- f_{pool} represents the pooling layers
- f_{FC} represents the fully connected layers
- $\hat{x}$ is the equalised signal

As it was mentioned above, the main benefit behind using CNNs in optical channel equalisation stems from their ability to capture both short-term and long-term dependencies in the received signal. The convolutional layers can extract local features related to ISI and nonlinear distortions, while the deeper layers can learn to compensate for more complex, long-range interactions [103].

The performance and complexity of CNN-based equalisers in optical communications depends on several factors [103, 120]:

- Network depth and width: Deeper networks can learn more complex nonlinear functions but may be prone to overfitting and increased computational complexity.
- Kernel size: Larger kernels can capture wider temporal dependencies but increase the number of parameters.
- Number of feature maps: More feature maps can extract richer representations but increase CC.
- Pooling strategy: The choice between max pooling and average pooling can affect the model's ability to capture relevant features.

Among the benefits of using CNNs when reducing nonlinear impairments in optical communication systems, it is worth mentioning spatial invariance, that allows the model to capture local patterns in the input signal regardless of their position, which is particularly useful for dealing with time-varying channel effects. Moreover, the use of shared weights in convolutional layers significantly reduces the number of trainable parameters compared to fully connected networks, leading to more efficient models [103].

3.2.3 1D Convolutional Neural Networks for Time-Series Optical Data Processing

The general CNN block introduced in § 2 extends naturally to 1-D signals. Here the 2-D kernel and pooling operations become 1-D, making them particularly well-suited for sequential data such as time-series signals in optical communications.

A typical 1D CNN architecture for optical signal processing consists of the following layers:

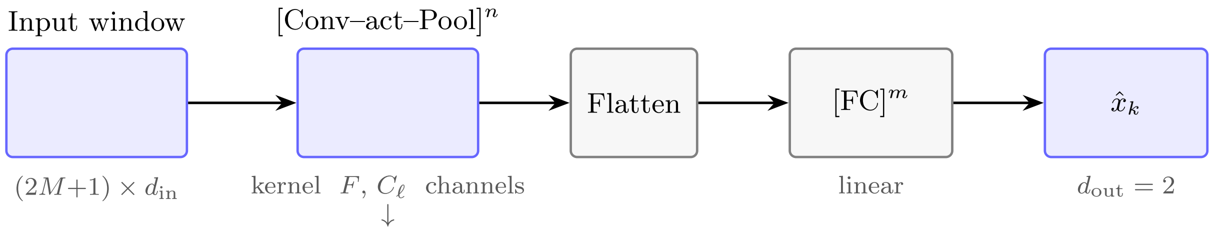


Figure 3.2: 1D CNN equaliser architecture for optical channel equalisation. The input is a temporal window of $2M+1$ symbol-spaced received samples $y(k-M), \ldots, y(k+M)$ with d_{in} channels per sample ($d_{\text{in}} = 2$ for in-phase and quadrature components in single-polarisation systems; $d_{\text{in}} = 4$ for dual-polarisation with I_H, Q_H, I_V, Q_V). The feature extractor consists of n repeated blocks of [Conv–act–Pool], where each 1D convolutional layer applies kernels of size F to extract local features related to inter-symbol interference (ISI) and nonlinear distortions, followed by a nonlinear activation function $f(\cdot)$ (typically ReLU) and max pooling that progressively reduces the temporal dimension (indicated by $\downarrow$). Each convolutional block produces C_ℓ feature maps ($\ell = 1, \ldots, n$) that capture different aspects of the temporal signal structure. After n convolutional-pooling blocks, the feature maps are flattened into a 1D vector and processed by m fully-connected layers with linear output activation to yield the equalised symbol estimate $\hat{x}_k$ with $d_{\text{out}} = 2$ components (I and Q). The convolutional layers exploit spatial invariance through shared weights across time positions, capturing local patterns regardless of their position in the input window, which is particularly useful for time-varying channel effects. This weight sharing significantly reduces the number of trainable parameters compared to fully-connected networks (see MLP architecture), leading to more efficient models while maintaining the ability to compensate for both short-term dependencies (through local kernels) and long-term interactions (through deeper layers). Typical hyperparameters for optical channel equalisation: memory depth $M \in [5, 20]$ symbols, kernel size $F \in [3, 7]$, stride $S \in [1, 2]$, pooling width $P \in [2, 3]$, number of blocks $n \in [2, 4]$, number of feature maps $C_\ell \in [16, 128]$, and number of FC layers $m \in [1, 2]$.

1. Input Layer: Accepts the raw time-series optical data
2. Convolutional Layers: Apply 1D convolutions to extract features
3. Activation Functions: Introduce non-linearity (e.g., ReLU)
4. Pooling Layers: Reduce spatial dimensions and therefore CC
5. Fully Connected Layers: Perform classification or regression

The 1D CNN's ability to extract hierarchical features from optical signals is particularly valuable. Lower layers capture local, short-term dependencies, while deeper layers learn more complex, long-term patterns [103]. This hierarchical feature extraction can be represented as:

$$h^{(l)} = f(W^{(l)} * h^{(l-1)} + b^{(l)}) \tag{3.14}$$

where $h^{(l)}$ is the feature map at layer l, $W^{(l)}$ is the convolutional kernel, $b^{(l)}$ is the bias, and $f(\cdot)$ is the activation function.

3.2.4 Feed-Forward Complex-Valued Neural Networks

In the context of this Thesis, not only is the focus on the version of ANNs that employs real values but also those that works with complex ones. The reason is that CVNNs have emerged as a promising approach for mitigating nonlinear impairments in optical communication systems. Unlike traditional real-valued ANNs, CVNNs can directly process complex-valued signals, making them particularly well-suited for coherent optical systems where both amplitude and phase information are important [74, 122, 123].

A CVNN-based equaliser consists of:

- An input layer that accepts complex-valued samples from the received optical signal
- One or more hidden layers with complex-valued neurons
- An output layer that produces the equalised complex-valued symbols

Mathematically, we can describe the operation of a CVNN layer as:

$$\mathbf{y} = f(\mathbf{W}\mathbf{x} + \mathbf{b}) \tag{3.15}$$

where $\mathbf{x}$ is the complex-valued input vector, $\mathbf{W}$ is the complex weight matrix, $\mathbf{b}$ is the complex bias vector, f is a complex activation function, and $\mathbf{y}$ is the complex-valued output vector [74, 122].

Like in the case of the real counterparts, the choice of activation function is crucial in CVNNs. Common complex activation functions include [122]:

1. Split-Complex Functions:
$$f(z) = f_R(\Re(z)) + j f_I(\Im(z)) \tag{3.16}$$
where f_R and f_I are real-valued functions applied separately to the real and imaginary parts.
2. Amplitude-Phase Functions:
$$f(z) = \tanh(|z|)e^{j\arg(z)} \tag{3.17}$$
which preserves the phase of the input while applying nonlinearity to the amplitude.

Training a CVNN involves minimising a complex-valued loss function, typically the MSE between the network output and the desired equalised symbols [74]:

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N} |y_i - \hat{y}_i|^2 \tag{3.18}$$

where y_i are the true symbols and $\hat{y}_i$ are the network outputs.

Gradient-based optimisation algorithms, such as complex-valued backpropagation, are used to update the network parameters. The gradients with respect to the complex weights are given by [122]:

$$\frac{\partial \text{MSE}}{\partial \mathbf{W}} = \frac{\partial \text{MSE}}{\partial \Re(\mathbf{W})} + j\frac{\partial \text{MSE}}{\partial \Im(\mathbf{W})} \tag{3.19}$$

CVNNs offer several advantages for optical channel equalisation [74, 123]:

- Direct processing of complex-valued signals without splitting into real and imaginary parts
- Preservation of phase information throughout the network
- Reduced number of parameters compared to equivalent real-valued networks
- Improved convergence speed during training

It has been demonstrated that CVNNs can be more expressive than their real-valued counterparts and outperform them in different contexts. Moreover, in the case of its application when designing ANN-based equalisers, CVNNs enable direct processing of complex-valued signals without splitting into real and imaginary parts, being competitive or outperforming traditional equalisers in various optical transmission scenarios. Nevertheless, implementing CVNNs for real-time optical channel equalisation presents several challenges [74, 123]:

- Increased computational complexity of complex arithmetic
- Limited support for complex-valued operations in standard deep learning frameworks
- Hardware implementation considerations for high-speed processing

3.3 Recurrent Neural Network Architectures for Optical Channel Equalisation

RNNs have emerged as powerful tools for sequence modelling tasks, making them particularly suitable for optical channel equalisation where temporal dependencies in the signal are crucial.

3.3.1 RNN Fundamentals

Unlike MLPs, which process input as a static, time-independent mapping, RNNs excel in capturing temporal dynamics by maintaining a hidden state that evolves over time. This recurrent structure enables RNNs to model both short-term and long-term dependencies in sequential data, addressing challenges such as inter-symbol interference and nonlinear channel effects. The hidden state is updated at each time step based on the current input and the previous hidden state, allowing the network to maintain a form of memory that is essential for addressing memory-dependent impairments. Furthermore, RNNs dynamically adapt to varying input contexts, making them robust against noise and variability often encountered in real-world optical systems. These characteristics position RNNs as essential components in modern communication systems, particularly in scenarios requiring continuous adaptation to time-varying channel dynamics [37, 69].

Figure 3.3 illustrates the key architectural difference between feedforward and recurrent networks. While both share the same basic layered structure, the RNN includes recurrent connections (shown as self-loops) in the hidden layer that enable temporal memory through state feedback.

The basic equations governing an RNN's forward pass are [69]:

$$\mathbf{h}_t = \sigma(\mathbf{W}_{hh}\mathbf{h}_{t-1} + \mathbf{W}_{xh}\mathbf{x}_t + \mathbf{b}_h) \tag{3.20}$$

$$\mathbf{y}_t = \mathbf{W}_{hy}\mathbf{h}_t + \mathbf{b}_y \tag{3.21}$$

where:

- $\mathbf{h}_t$ is the hidden state at time t
- $\mathbf{x}_t$ is the input at time t
- $\mathbf{y}_t$ is the output at time t
- $\mathbf{W}_{hh}$, $\mathbf{W}_{xh}$, and $\mathbf{W}_{hy}$ are weight matrices
- $\mathbf{b}_h$ and $\mathbf{b}_y$ are bias vectors
- $\sigma(\cdot)$ is an activation function, typically tanh or ReLU

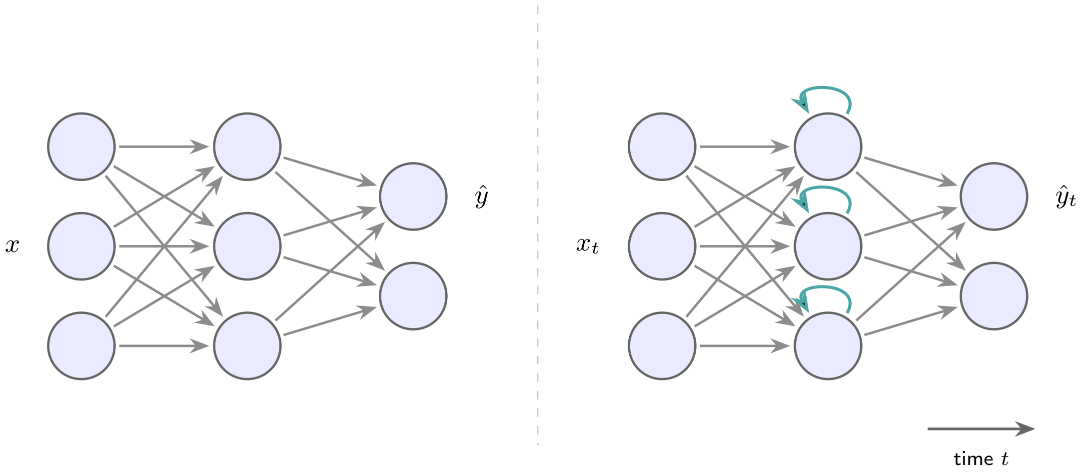


Figure 3.3: Comparison of feedforward ANN (left) and recurrent neural network (right) architectures. Both share the same feedforward structure, but the RNN includes recurrent connections (teal self-loops) in the hidden layer that enable temporal memory. These connections allow each hidden neuron to feed its output back to itself, creating an internal state $\mathbf{h}_t = f(\mathbf{W}_{hh}\mathbf{h}_{t-1} + \mathbf{W}_{xh}\mathbf{x}_t + \mathbf{b}_h)$ that depends on both current input $\mathbf{x}_t$ and previous state $\mathbf{h}_{t-1}$. The feedforward network produces static mappings $\hat{\mathbf{y}} = g(\mathbf{x})$ suitable for memoryless tasks, while the RNN's sequential processing $\hat{\mathbf{y}}_t = g(\mathbf{h}_t)$ is essential for modelling time-varying optical channel impairments.

Training RNNs involves a variation of Backpropagation algorithm already discussed, called Backpropagation Through Time (BPTT). The loss is computed at each time step and then propagated backwards through the network. The gradients for the weights are [100]:

$$\frac{\partial L}{\partial \mathbf{W}_{hy}} = \sum_{t=1}^{T} \frac{\partial L_t}{\partial \mathbf{y}_t} \mathbf{h}_t^T \tag{3.22}$$

$$\frac{\partial L}{\partial \mathbf{W}_{hh}} = \sum_{t=1}^{T} \frac{\partial L_t}{\partial \mathbf{h}_t} \frac{\partial \mathbf{h}_t}{\partial \mathbf{W}_{hh}} \tag{3.23}$$

$$\frac{\partial L}{\partial \mathbf{W}_{xh}} = \sum_{t=1}^{T} \frac{\partial L_t}{\partial \mathbf{h}_t} \frac{\partial \mathbf{h}_t}{\partial \mathbf{W}_{xh}} \tag{3.24}$$

where L is the total loss and L_t is the loss at time step t.

To understand BPTT, it is helpful to visualize the RNN as an unrolled computational graph over time, as shown in Figure 3.4. This unrolling makes explicit how the same RNN cell is applied at each time step, with hidden states flowing from one step to the next.

A significant challenge in training standard RNNs (often referred to as "vanilla RNNs" in the literature to distinguish them from more advanced variants like LSTM and GRU) is the fact that they are prone to the already mentioned vanishing and exploding gradient problem. The term "vanilla" here emphasizes the basic, unmodified architecture without any gating mechanisms or memory

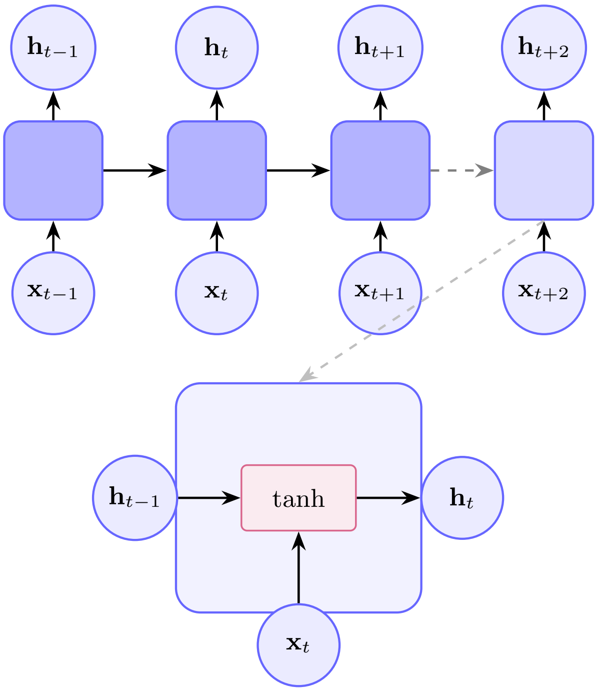


Figure 3.4: RNN unrolled over time (top) and single RNN cell structure (bottom). The unrolled view shows how the same RNN cell processes a sequence of inputs $\mathbf{x}_t$ over multiple time steps, with hidden states $\mathbf{h}_t$ flowing from one step to the next. Each blue box represents the same RNN operation: $\mathbf{h}_t = \sigma(\mathbf{W}_{hh}\mathbf{h}_{t-1} + \mathbf{W}_{xh}\mathbf{x}_t + \mathbf{b}_h)$. The bottom panel shows the internal structure of a single cell, with the tanh activation processing both the previous hidden state and current input to produce the new hidden state. This parameter sharing across time steps enables RNNs to process sequences of arbitrary length while maintaining a consistent temporal representation, essential for modelling time-varying optical channel dynamics.

control structures. As the error propagates back through time steps, it can either decay exponentially (vanishing) or grow exponentially (exploding). This is due to the repeated multiplication by the weight matrix $\mathbf{W}_{hh}$ [100, 103].

The magnitude of the gradients is influenced by the eigenvalues of $\mathbf{W}_{hh}$ [100]:

$$\|\frac{\partial L}{\partial \mathbf{h}_t}\| \approx \|\mathbf{W}_{hh}\|^{T-t} \|\frac{\partial L}{\partial \mathbf{h}_T}\| \tag{3.25}$$

If $\|\mathbf{W}_{hh}\| < 1$, gradients vanish; if $\|\mathbf{W}_{hh}\| > 1$, gradients explode.

In the context of optical channel equalisation, standard RNNs (vanilla RNNs) can be used to model the temporal dependencies in the received signal. The input $\mathbf{x}_t$ could represent a window of received samples, while the output $\mathbf{y}_t$ could be the equalised symbol at time t [69, 124].

The hidden state $\mathbf{h}_t$ captures the channel memory, allowing the network to account for ISI and other memory-dependent impairments. However, due to the vanishing gradient problem, standard RNNs (vanilla RNNs) may struggle to capture long-term dependencies in the optical channel, especially in long-haul transmission systems with significant dispersion [69].

Figure 3.5 shows the basic architecture of a vanilla RNN cell, illustrating its simplicity compared to more advanced variants. The cell processes the current input $\mathbf{x}_t$ and previous hidden state $\mathbf{h}_{t-1}$ through a single nonlinear activation (typically tanh) to produce both the current hidden state $\mathbf{h}_t$ and output $\mathbf{y}_t$.

To address these limitations, more advanced RNN architectures such as LSTM and GRU have been developed and applied to optical channel equalisation with greater success, as discussed in the following sections [103, 125].

3.3.2 LSTM Networks

LSTM networks address the vanishing gradient problem encountered in standard RNNs (vanilla RNNs), making them more suitable for capturing long-term dependencies in optical channel equalisation tasks. The LSTM architecture incorporates gating mechanisms that allow the network to selectively remember or forget information over long sequences [125].

The core component of an LSTM network is the LSTM cell, which consists of three gates (forget, input, and output) and a memory cell. The operations within an LSTM cell at time step t are defined by the following equations [103, 125]:

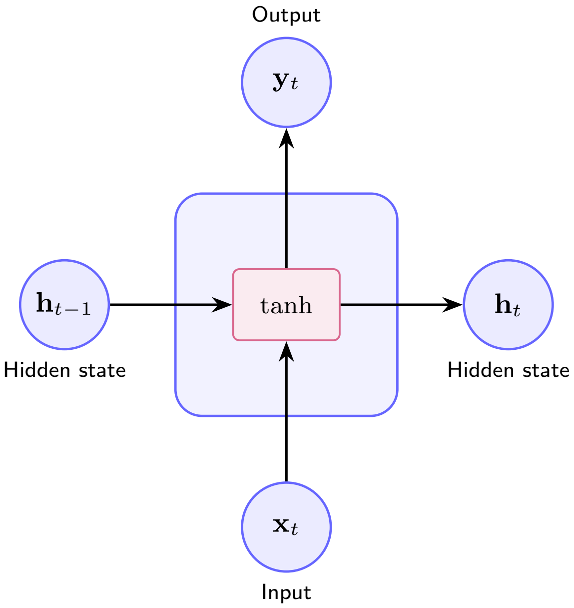


Figure 3.5: Vanilla RNN cell architecture for optical channel equalisation. The cell receives the previous hidden state $\mathbf{h}_{t-1}$ and current input $\mathbf{x}_t$, processes them through a nonlinear activation function (typically tanh) to produce the current hidden state $\mathbf{h}_t = \sigma(\mathbf{W}_{hh}\mathbf{h}_{t-1} + \mathbf{W}_{xh}\mathbf{x}_t + \mathbf{b}_h)$, and generates the output $\mathbf{y}_t = \mathbf{W}_{hy}\mathbf{h}_t + \mathbf{b}_y$. Unlike LSTMs and GRUs, the vanilla RNN has no gating mechanisms, making it simpler but more susceptible to vanishing gradients in long sequences, which limits its effectiveness for modelling long-range dependencies in optical channels with significant dispersion.

$$\mathbf{f}_t = \sigma(\mathbf{W}_f[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_f) \tag{3.26}$$

$$\mathbf{i}_t = \sigma(\mathbf{W}_i[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_i) \tag{3.27}$$

$$\tilde{\mathbf{c}}_t = \tanh(\mathbf{W}_c[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_c) \tag{3.28}$$

$$\mathbf{c}_t = \mathbf{f}_t \odot \mathbf{c}_{t-1} + \mathbf{i}_t \odot \tilde{\mathbf{c}}_t \tag{3.29}$$

$$\mathbf{o}_t = \sigma(\mathbf{W}_o[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_o) \tag{3.30}$$

$$\mathbf{h}_t = \mathbf{o}_t \odot \tanh(\mathbf{c}_t) \tag{3.31}$$

where $\mathbf{f}_t$, $\mathbf{i}_t$, and $\mathbf{o}_t$ are the forget, input, and output gates respectively, $\mathbf{c}_t$ is the cell state, $\mathbf{h}_t$ is the hidden state, $\mathbf{x}_t$ is the input at time t, $\mathbf{W}_*$ and $\mathbf{b}_*$ are the weight matrices and bias vectors for the respective gates, $\sigma(\cdot)$ is the sigmoid activation function, and $\odot$ denotes element-wise multiplication.

The forget gate $\mathbf{f}_t$ determines how much of the previous cell state to retain, the input gate $\mathbf{i}_t$ controls the flow of new information into the cell state, and the output gate $\mathbf{o}_t$ regulates the information output from the cell. This gating mechanism allows LSTMs to capture long-range dependencies in the input sequence, making them particularly suitable for modelling the memory effects in optical channels [103].

In the context of optical channel equalisation, LSTMs are employed to process sequences of received symbols, effectively capturing both short-term and long-term dependencies introduced by CD, PMD, and nonlinear effects. The ability of LSTMs to selectively retain relevant information over long sequences makes them well-suited for scenarios where the channel memory spans multiple symbol periods [103].

While standard LSTMs process a sequence from start to finish, bidirectional LSTMs (BiLSTMs) take this one step further by also looking at the sequence in reverse. In other words, they read from both left to right and right to left. This extra step can be a game-changer for tasks like optical channel equalisation because future symbols can contain important clues that help correct the current symbol. By examining both past and future contexts, a BiLSTM can better capture the dependencies in a signal that might otherwise be missed [103].

In practical terms, a BiLSTM combines two separate LSTM layers: the “forward” layer processes the input in its natural order, while the “backward” layer reads it in reverse. The outputs from these two layers are then merged to produce a single final result. Because many optical channels exhibit interference from neighboring symbols, having access to both directions at once allows the network to learn richer patterns, leading to more accurate equalisation of the transmitted symbols. This is mathematically expressed as [126]:

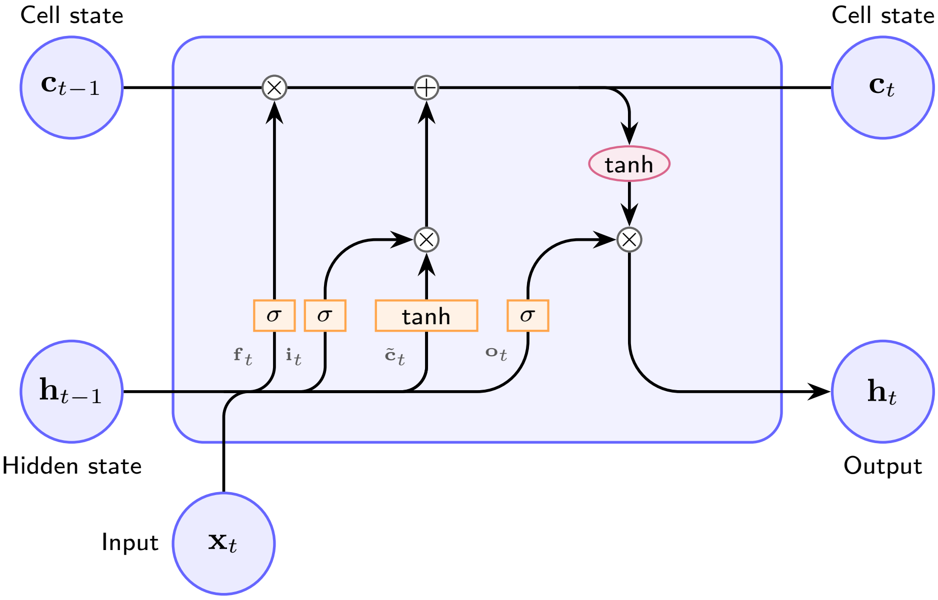


Figure 3.6: LSTM cell used in the equaliser. At time t, the input $\mathbf{x}_t$ and previous hidden state $\mathbf{h}_{t-1}$ produce the forget, input and output gates $\mathbf{f}_t, \mathbf{i}_t, \mathbf{o}_t$ (boxes labeled σ), and the candidate state $\tilde{\mathbf{c}}_t$ (box labeled $\tanh$). The cell state follows a gated carry path $\mathbf{c}_t = \mathbf{f}_t \odot \mathbf{c}_{t-1} + \mathbf{i}_t \odot \tilde{\mathbf{c}}_t$, and the hidden state is $\mathbf{h}_t = \mathbf{o}_t \odot \tanh(\mathbf{c}_t)$. Circles marked $\times$ and $+$ denote element-wise product and summation, respectively. The explicit cell-state highway mitigates vanishing gradients and enables the model to capture long-range memory in optical channels.

$$\overrightarrow{\mathbf{h}}_t = \text{LSTM}_f(\mathbf{x}_t, \overrightarrow{\mathbf{h}}_{t-1}) \tag{3.32}$$

$$\overleftarrow{\mathbf{h}}_t = \text{LSTM}_b(\mathbf{x}_t, \overleftarrow{\mathbf{h}}_{t+1}) \tag{3.33}$$

$$\mathbf{y}_t = \mathbf{W}_y[\overrightarrow{\mathbf{h}}_t, \overleftarrow{\mathbf{h}}_t] + \mathbf{b}_y \tag{3.34}$$

where $\overrightarrow{\mathbf{h}}_t$ and $\overleftarrow{\mathbf{h}}_t$ are the hidden states of the forward and backward LSTM layers respectively, and $\mathbf{y}_t$ is the output at time step t.

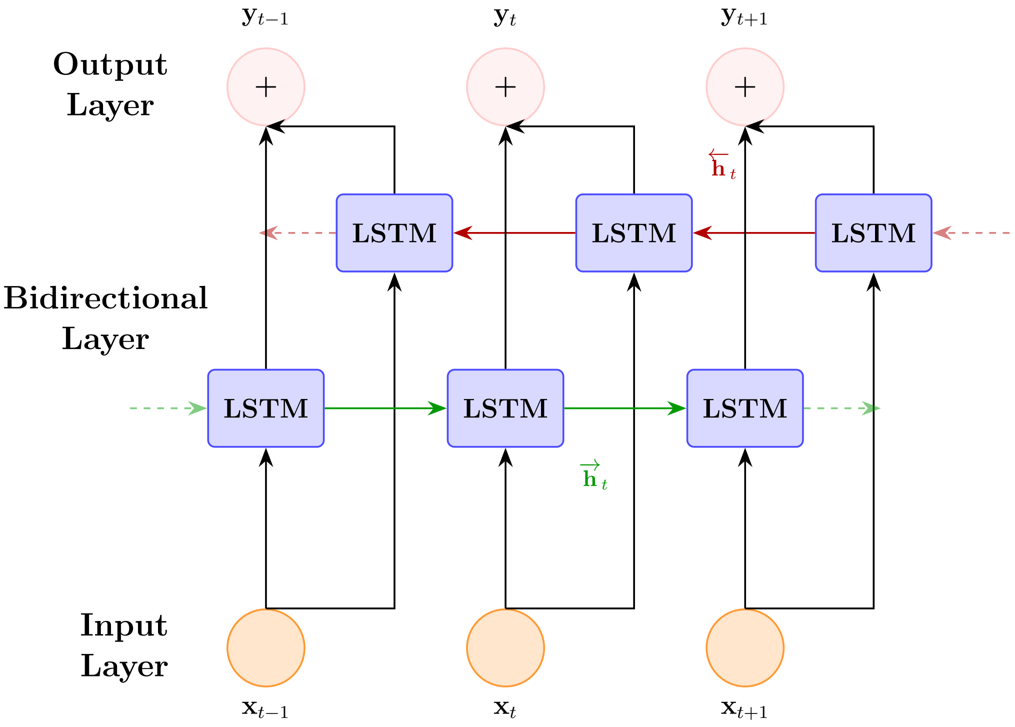


Figure 3.7: Architecture of the BiLSTM network used for optical channel equalisation. The model processes the input sequence in both forward and backward directions, enabling it to capture dependencies from past and future symbols. The concatenated hidden states from both directions are then combined to generate the final output, providing improved context understanding and enhanced nonlinearity mitigation.

In the context of optical channel equalisation, BiLSTMs offer several advantages [103]:

1. Improved context understanding: By processing the sequence in both directions, BiLSTMs can capture dependencies on both past and future symbols, which is particularly useful for compensating effects that spread symbol information in both temporal directions.
2. Enhanced nonlinearity mitigation: The bidirectional processing allows the network to better model and compensate for nonlinear effects that may depend on both preceding and

succeeding symbols.

3. Robustness to pattern-dependent effects: BiLSTMs can more effectively handle pattern-dependent impairments by considering the full context around each symbol.

Like in the case of the LSTM, the BiLSTMs can tackle equalisation tasks effectively because they learn temporal dependencies in the signal from both past and future contexts. In optical communications, inter-symbol interference spreads the impact of any given transmitted symbol over multiple symbol periods, so the received signal does not just depend on the current symbol—neighboring symbols are also highly relevant. A BiLSTM naturally “remembers” information from earlier inputs (and also processes later inputs in reverse order if you give it the full sequence), which helps it capture long-range temporal correlations and mitigate interference. Going back to the comparison with the already described MLP, it is worth noticing that this model is essentially memoryless: it takes a snapshot of the data at a single time step and produces an output without a built-in mechanism to incorporate sequence context. Although it can approximate many functions in principle, it has no inherent feedback or gating structure to keep track of what happened in previous (or future) time steps. This is why simply giving an MLP the raw signal samples often underperforms compared to a recurrent model that can store and update a latent representation of the signal’s past (and, in a BiLSTM, future). The benefit of BiLSTMs is less about differences in the “information content” of the signal itself and more about these built-in memory and gating mechanisms that leverage correlations across time. [103, 126, 127].

3.3.3 GRUs

GRUs offer a simplified gating mechanism compared to LSTMs, potentially reducing computational complexity while maintaining comparable performance in many tasks. GRUs are designed to capture dependencies of different time scales adaptively and address the vanishing gradient problem often encountered in standard RNNs [69].

The GRU architecture combines the forget and input gates into a single "update gate" and merges the cell state and hidden state. This results in a simpler model with fewer parameters, which can be beneficial in terms of computational efficiency and reduced risk of overfitting, especially when dealing with smaller datasets [128].

The key equations governing the GRU’s operation are as follows [69, 128]:

$$\mathbf{z}_t = \sigma(\mathbf{W}_z[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_z) \tag{3.35}$$

$$\mathbf{r}_t = \sigma(\mathbf{W}_r[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_r) \tag{3.36}$$

$$\tilde{\mathbf{h}}_t = \tanh(\mathbf{W}[\mathbf{r}_t \odot \mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}) \tag{3.37}$$

$$\mathbf{h}_t = (1 - \mathbf{z}_t) \odot \mathbf{h}_{t-1} + \mathbf{z}_t \odot \tilde{\mathbf{h}}_t \tag{3.38}$$

where:

- $\mathbf{z}_t$ is the update gate, which decides how much of the past information (from previous time steps) needs to be passed along to the future.
- $\mathbf{r}_t$ is the reset gate, which determines how much of the past information to forget.
- $\tilde{\mathbf{h}}_t$ is the candidate hidden state.
- $\mathbf{h}_t$ is the current hidden state.
- $\mathbf{x}_t$ is the input at time step t.
- $\mathbf{W}_z$, $\mathbf{W}_r$, and $\mathbf{W}$ are weight matrices.
- $\mathbf{b}_z$, $\mathbf{b}_r$, and $\mathbf{b}$ are bias vectors.
- $\sigma(\cdot)$ is the sigmoid activation function.
- $\tanh(\cdot)$ is the hyperbolic tangent activation function.
- $\odot$ denotes element-wise multiplication.

The update gate $\mathbf{z}_t$ plays a crucial role in determining how much of the previous hidden state should be retained. When $\mathbf{z}_t$ is close to 1, the model keeps most of the previous state and only updates it minimally with new information. Conversely, when $\mathbf{z}_t$ is close to 0, the model discards most of the previous state and primarily uses the new candidate state.The reset gate $\mathbf{r}_t$ allows the model to effectively "forget" past information by controlling how much of the previous hidden state is used in computing the candidate hidden state $\tilde{\mathbf{h}}_t$. This mechanism enables the GRU to capture short-term dependencies when necessary [128].

In the context of optical channel equalisation, like in the case of LSTMs, GRUs can be particularly effective due to their ability to adaptively handle varying channel memory lengths. The gating mechanism allows the network to selectively focus on relevant historical information, which is crucial in dealing with the complex temporal dependencies introduced by CD and other fibre impairments [69, 103].

Moreover, the reduced parameter count of GRUs compared to LSTMs can be advantageous in hardware implementations for real-time optical signal processing. The simpler architecture potentially leads to lower latency and power consumption, which are critical factors in high-speed optical communication systems [69].

However, it is important to note that the choice between GRUs and other RNN variants (such as LSTMs) often depends on the specific characteristics of the optical channel and the requirements of the equalisation task. Empirical evaluation is typically necessary to determine the most suitable architecture for a given application [69, 103, 124].

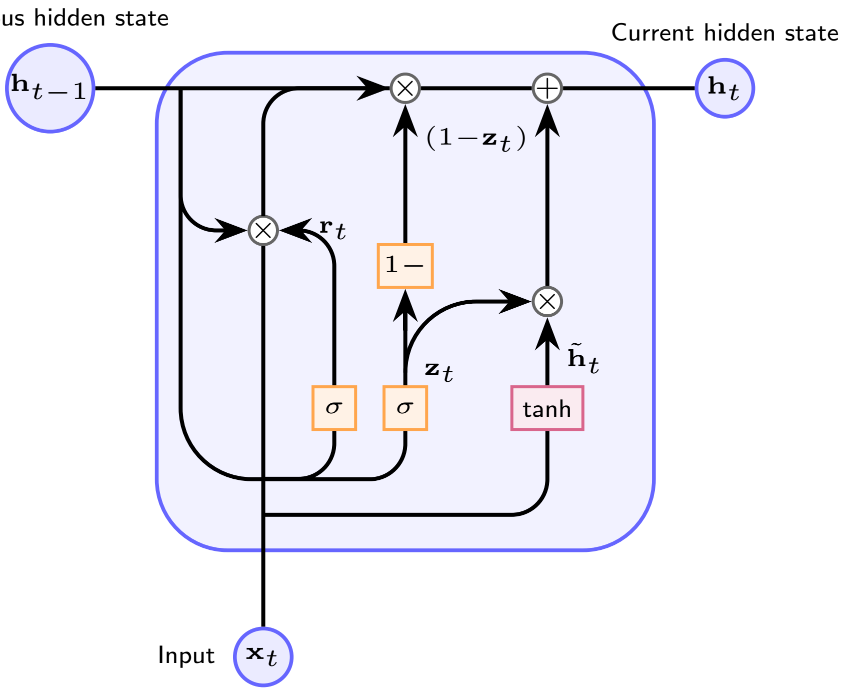


Figure 3.8: GRU cell for optical channel equalisation. The update and reset gates $\mathbf{z}_t$ and $\mathbf{r}_t$ (boxes labeled σ) are computed from $[\mathbf{h}_{t-1}, \mathbf{x}_t]$. The reset gate modulates the contribution of the past when forming the candidate $\tilde{\mathbf{h}}_t = \tanh\big(\mathbf{W}[\mathbf{r}_t \odot \mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}\big)$, and the output blends past and candidate via $\mathbf{h}_t = (1 - \mathbf{z}_t) \odot \mathbf{h}_{t-1} + \mathbf{z}_t \odot \tilde{\mathbf{h}}_t$. As in the figure, circles indicate element-wise product and $+$ summation. Compared with the LSTM, the GRU has no separate cell state, yielding fewer parameters while retaining the ability to model channel memory.

3.4 Computational Complexity Metrics for Neural Network Equalisers

Before exploring specific compression techniques, it is crucial to establish appropriate metrics for evaluating the CC of ANNs-based equalisers. These metrics provide a standardised framework for comparing different architectures and compression methods.

- Number of Real Multiplications (RM): The number of real multiplications (RM) per recovered symbol is a fundamental metric that directly correlates with the computational complexity (CC) of the equaliser. Both real-valued and complex-valued neural networks are considered in this analysis. For a feed-forward real-valued neural network with L layers, the RM can be expressed as [52, 68, 129]:

  $$RM_{real} = \sum_{l=1}^{L} n_l \cdot n_{l-1} \tag{3.39}$$

  where n_l is the number of neurons in layer l. In optical communication systems, complex-valued neural networks (CVNNs) are often employed to process in-phase and quadrature components simultaneously. For a CVNN, each complex multiplication requires four real multiplications. Therefore, the RM for a complex-valued feed-forward neural network can be calculated as [105]:

  $$RM_{complex} = 4 \sum_{l=1}^{L} n_l \cdot n_{l-1} \tag{3.40}$$

- Number of Bit Operations (BOP): The BOP metric accounts for the bitwidth of weights and activations, providing a more accurate representation of the computational cost in fixed-point implementations [52, 68, 129]:

  $$BOP = \sum_{l=1}^{L} n_l \cdot n_{l-1} \cdot (b_w \cdot b_a + b_w + b_a) \tag{3.41}$$

  where b_w and b_a denote the bitwidths of weights and activations, respectively [68].

- Number of Additions and Bit Shifts (NABS): The NABS metric is particularly relevant for quantised models where multiplications are often replaced by bit shifts and additions [52, 129]:

  $$NABS = \sum_{l=1}^{L} n_l \cdot n_{l-1} \cdot (X_w + 1) \cdot (b_w + b_a + \log_2(n_{l-1})) \tag{3.42}$$

  where X_w is the number of adders required to represent a multiplication operation [129].

These metrics will be used throughout this chapter to evaluate the effectiveness of various compression techniques.

3.4.1 Q-factor as Performance Metric: Comparison with Validation Accuracy

In the context of optical communication systems, the Q-factor has emerged as the primary performance metric for evaluating ANN-based equalisers, in contrast to the validation accuracy commonly used in general machine learning applications. Understanding the relationship between these metrics and their respective applicability is essential for proper assessment of equaliser performance.

The Q-factor, expressed in decibels (dB), is derived from the bit error ratio (BER) and provides a direct measure of signal quality in optical transmission systems [103]:

$$Q = 20\log_{10}\left(\sqrt{2}\cdot \text{erfc}^{-1}(2\cdot \text{BER})\right) \tag{3.43}$$

where erfc^{-1} is the inverse complementary error function. The Q-factor quantifies the SNR margin in the decision regions of the constellation diagram, directly correlating with system performance in terms of optical reach, power budget, and transmission reliability.

In contrast, validation accuracy—commonly employed in classification tasks—measures the percentage of correctly classified samples. While validation accuracy is intuitive and widely used in machine learning, it has several limitations when applied to optical channel equalisation [37, 103]:

- Binary vs. Continuous Assessment: Validation accuracy treats all errors equally (correct or incorrect), whereas the Q-factor incorporates the magnitude of symbol estimation errors through the BER calculation, providing a more nuanced performance assessment critical for communication systems.

- Industry Standard Compatibility: The Q-factor is an established metric in optical communications, enabling direct comparison with traditional DSP techniques (e.g., CDC, DBP) and facilitating integration into existing performance evaluation frameworks. Validation accuracy, being unfamiliar to optical engineers, would hinder cross-method comparisons.

- Physical System Correlation: Q-factor directly relates to physical transmission parameters such as OSNR, launch power, and fibre dispersion. A Q-factor improvement of 1 dB, for example, can translate to extended transmission distance or reduced required OSNR, metrics that validation accuracy cannot capture.

- Sensitivity to Degradation: For high-performance equalisers achieving BER $< 10^{-3}$ (Q-factor > 9.8 dB), validation accuracy often saturates near 100%, masking subtle but critical performance differences. The Q-factor, measured in logarithmic scale, remains sensitive across the entire operational range.

Throughout this chapter, we report Q-factor degradation relative to the uncompressed baseline model. For instance, a compressed model maintaining "Q-factor degradation < 0.5 dB" indicates performance within 0.5 dB of the original dense network. This degradation-based reporting enables clear assessment of compression technique impact across different transmission scenarios.

An important consideration in ANN-based equalisation is the consistency of Q-factor performance between training and testing datasets. In our experiments across multiple transmission scenarios (presented in Sections 3.2-3.17), we observed the following characteristic behaviours:

- Generalisation Gap: The Q-factor on training data typically exceeds that on testing data by 0.2-0.5 dB, reflecting the model's inherent overfitting tendency. For example, in the 30 Gbaud 64-QAM 20×50km scenario, the baseline MLP achieved Q-factor of 11.1 dB on training data versus 10.8 dB on testing data.
- Compression Impact on Generalisation: Interestingly, moderate compression (40-60% sparsity) often reduces the training-testing Q-factor gap. At 60% sparsity, the training Q-factor drops to 10.9 dB while testing Q-factor remains at 10.7 dB, suggesting that pruning acts as implicit regularisation, improving generalisation.
- Aggressive Compression Threshold: Beyond 70-80% sparsity (depending on scenario), both training and testing Q-factors degrade rapidly and convergently, indicating loss of model capacity rather than overfitting. This threshold identifies the practical compression limit for each architecture.
- Quantisation Robustness: Post-training quantisation from FP32 to INT8 affects training and testing Q-factors nearly identically (both degrade by ≈ 0.1 dB), confirming that quantisation primarily impacts representation precision rather than generalisation capability.

These observations validate the use of Q-factor as a comprehensive performance metric that captures both equalisation quality and generalisation behaviour. All results presented in this chapter report testing dataset Q-factors unless explicitly stated, ensuring conservative performance estimates applicable to real-world deployment scenarios.

3.5 Pruning Strategies for ANNs-based Equalisers

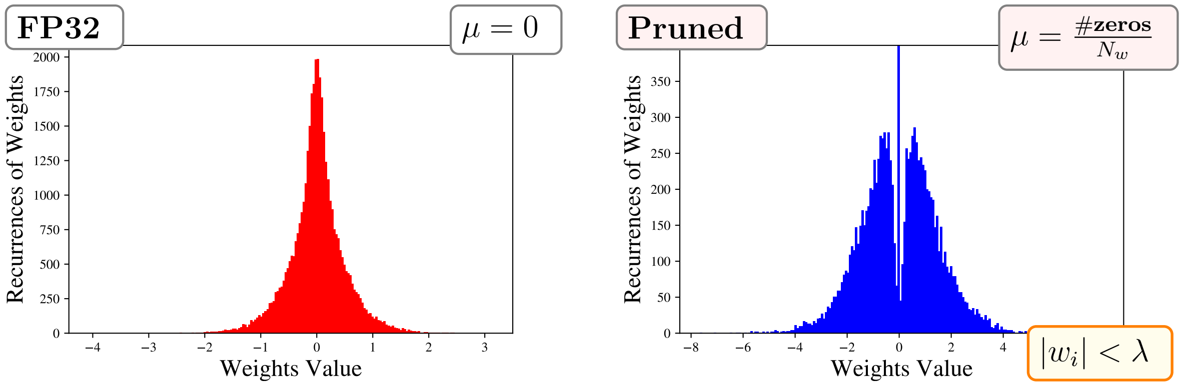
Pruning is a technique that involves removing unnecessary elements from an ANN, effectively reducing its size and CC. In the context of optical channel equalisation, pruning can be particularly effective due to the inherent sparsity in the equaliser's weight matrices [52, 68, 116].

$$w_{ij} = \begin{cases} 0, & \text{if } |w_{ij}| < \lambda, \\ w_{ij}, & \text{otherwise} \end{cases} \tag{3.44}$$

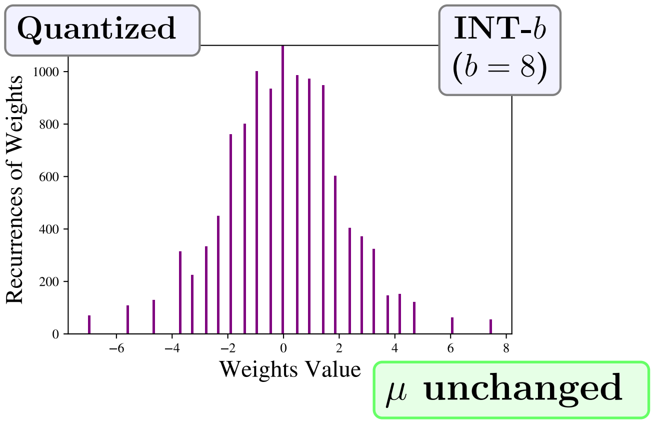
where λ is the pruning threshold. This method is based on the assumption that smaller weights contribute less to the network's output and can be removed with minimal impact on performance.

Figure 3.9 provides comprehensive visualisation of compression techniques on weight distributions for optical channel equalisers. Panel (a) shows the dense FP32 baseline with Gaussian-like distribution and zero sparsity ($\mu = 0$). Panel (b) demonstrates pruning's effect: applying threshold λ creates a prominent zero spike at $w = 0$ representing 60% of the original 510K parameters being removed—this discrete mass of 306,000 zeros is the visual signature of structural sparsity ($\mu = 0.60$). Panel (c) contrasts with quantisation, which discretizes values to 256 levels but critically preserves sparsity (μ unchanged, no zero spike).

The pruning workflow proceeds in three stages: (1) Train dense network to convergence, (2) Identify weights satisfying $|w_i| < \lambda$ for removal based on magnitude distribution analysis, (3) Set identified weights to zero and fine-tune remaining parameters to recover performance. Computational impact is direct: multiply-accumulate operations scale as $\text{MACs}_{\text{pruned}} \approx (1 - \mu) \times \text{MACs}_{\text{dense}}$, yielding 60% complexity reduction (36,000 $\rightarrow$ 14,400 RMpS) proportional to sparsity. Memory footprint decreases from 2.04 MB to 816 KB as sparse storage stores only non-zero weights.



(a) Dense (baseline) (b) Magnitude pruning



(c) Uniform quantization

Figure 3.9: Comparison of weight distributions for compression techniques applied to a BiLSTM optical equaliser. Each histogram shows the empirical distribution of the recurrent kernel weights of one of the LSTM layers. (a) Dense baseline (FP32): Gaussian-like distribution, $\mu = 0$. (b) Magnitude pruning: Threshold $|w_i| < \lambda$ removes a certain % of weights determined by μ, creating a zero spike at $w = 0$. (c) Uniform quantisation: Mapping to 256 levels (INT8) discretizes weights without sparsity change ($\mu = 0$). Pruning alters topology (sparsity), while quantisation alters numerical precision.

Beyond simple magnitude-based pruning, several advanced strategies have been developed to improve the effectiveness of pruning while maintaining network performance. One of this is the known as fine-tuning, in which after initial pruning, the network is retrained with a reduced learning rate to adapt to the new sparse structure [52]:

$$\theta_{t+1} = \theta_t - \eta \nabla L(\theta_t) \tag{3.45}$$

where θ_t are the network parameters at step t, η is a small learning rate, and $L(\theta_t)$ is the loss function. Another one is weight rewinding, which is inspired by the lottery ticket hypothesis. This hypothesis states that within a large, over-parameterised neural network, there exists a smaller subnetwork that, when trained in isolation, can match or exceed the performance of the original network. This subnetwork is referred to as a "winning ticket" [130] Weight rewinding pruning process involves [52]:

1. Training the network to convergence
2. Pruning low-magnitude weights
3. Resetting remaining weights to their initial values
4. Retraining from this initialisation

The process can be described as:

$$w_{ij} = \begin{cases} 0, & \text{if } |w_{ij}| < \lambda, \\ w_{ij}, & \text{otherwise} \end{cases} \tag{3.46}$$

where $w_{ij}^{(0)}$ is the initial value of weight w_{ij}. Finally, it is a method that combines features of fine-tuning and weight rewinding is learning rate rewinding [52]:

1. Train the network to convergence
2. Prune low-magnitude weights
3. Reset the learning rate schedule, but keep the trained weights
4. Retrain with the original learning rate schedule

The learning rate at step t during retraining is set to:

$$\eta_t = \eta_{\text{schedule}}(t \bmod T) \tag{3.47}$$

where η_{schedule} is the original learning rate schedule and T is the length of the schedule. These advanced pruning strategies often lead to better preservation of network performance compared to simple magnitude-based pruning, especially for deep models used in complex tasks like nonlinear equalisation in optical communications.

Structured pruning is another approach which removes entire neurons or filters, leading to a more regular sparsity pattern that can be more efficiently exploited by hardware. Unlike unstructured pruning, which removes individual weights, structured pruning targets whole structural units of the network [52, 68].

One common method to determine which neurons to prune is based on their importance. The importance of a neuron can be calculated as:

$$\text{Importance}(n_i) = \sum_j |w_{ij}| \tag{3.48}$$

where w_{ij} is the weight connecting neuron i to neuron j in the subsequent layer. Neurons with the lowest importance scores are removed from the network.

3.5.1 Bayesian Optimisation Fundamentals for Neural Network Equalisation

Before examining the specific applications of BO to neural network compression techniques in optical communications, it is essential to establish the theoretical foundations that make BO particularly well-suited for optimising ANN-based equalisers. BO represents a powerful global optimisation framework designed to find the optimal hyperparameters of expensive-to-evaluate black-box functions, making it ideally suited for neural network optimisation where training and evaluation are computationally intensive [113, 131].

BO addresses the optimisation problem of finding the global optimum of an unknown objective function $f(\mathbf{x})$ over a bounded domain $\mathcal{X} \subseteq \mathbb{R}^d$:

$$\mathbf{x}^* = \arg\max_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x}) \tag{3.49}$$

In the context of optical communication equalisers, $f(\mathbf{x})$ typically represents a performance metric (such as Q-factor or BER) as a function of hyperparameters $\mathbf{x}$ (learning rates, network architectures, compression parameters, etc.). The fundamental challenge is that evaluating $f(\mathbf{x})$ requires training and testing the neural network, which is computationally expensive and time-consuming.

BO addresses this challenge through two key components: a probabilistic surrogate model that captures our beliefs about the unknown function, and an acquisition function that guides the selection of the next evaluation point. The most commonly used surrogate model is the Gaussian Process (GP), which provides a principled way to model uncertainty about the objective function [132, 133]. A Gaussian Process defines a distribution over functions and can be completely specified by its mean function $m(\mathbf{x})$ and covariance (kernel) function $k(\mathbf{x}, \mathbf{x}')$:

$$f(\mathbf{x}) \sim \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')) \tag{3.50}$$

For neural network optimisation in optical communications, the mean function is typically set to zero, $m(\mathbf{x}) = 0$, while the covariance function captures the smoothness assumptions about the objective function. The Matérn 5/2 kernel is commonly used:

$$k(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \left(1 + \sqrt{5}r + \frac{5r^2}{3}\right) \exp\left(-\sqrt{5}r\right) \tag{3.51}$$

where $r = \|\mathbf{x} - \mathbf{x}'\|/\ell$, ℓ is the characteristic length scale, and σ_f^2 is the signal variance.

Given a set of observed evaluations $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^n$ where $y_i = f(\mathbf{x}_i) + \epsilon_i$ and $\epsilon_i \sim \mathcal{N}(0, \sigma_n^2)$, the GP provides analytical posterior predictions. The predictive mean and variance at a new point

$\mathbf{x}_*$ are:

$$\mu(\mathbf{x}_*|\mathcal{D}) = \mathbf{k}_*^T(\mathbf{K} + \sigma_n^2\mathbf{I})^{-1}\mathbf{y} \tag{3.52}$$

$$\sigma^2(\mathbf{x}_*|\mathcal{D}) = k(\mathbf{x}_*, \mathbf{x}_*) - \mathbf{k}_*^T(\mathbf{K} + \sigma_n^2\mathbf{I})^{-1}\mathbf{k}_* \tag{3.53}$$

where $\mathbf{k}_* = [k(\mathbf{x}_*, \mathbf{x}_1), \ldots, k(\mathbf{x}_*, \mathbf{x}_n)]^T$, $\mathbf{K}$ is the covariance matrix with entries $K_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$, and $\mathbf{y} = [y_1, \ldots, y_n]^T$.

The acquisition function $\alpha(\mathbf{x})$ determines where to evaluate the objective function next by balancing exploration (sampling where uncertainty is high) and exploitation (sampling where the predicted function value is high).

- Expected Improvement (EI):
  EI measures the expected improvement over the current best observed value $f^+ = \max_{i=1,\ldots,n} y_i$:

  $$\alpha_{EI}(\mathbf{x}) = \mathbb{E}[\max(f(\mathbf{x}) - f^+, 0)] = \sigma(\mathbf{x})\,[\phi(Z) + Z\Phi(Z)] \tag{3.54}$$

  where $Z = \frac{\mu(\mathbf{x})-f^+}{\sigma(\mathbf{x})}$, ϕ is the standard normal PDF, and Φ is the standard normal CDF. EI provides a good balance between exploration and exploitation, making it suitable for neural network hyperparameter optimisation where both convergence speed and solution quality are important.

- Upper Confidence Bound (UCB):
  UCB balances exploration and exploitation through a confidence parameter β:

  $$\alpha_{UCB}(\mathbf{x}) = \mu(\mathbf{x}) + \sqrt{\beta}\sigma(\mathbf{x}) \tag{3.55}$$

  The parameter β controls the exploration-exploitation trade-off. Higher values of β encourage more exploration, which can be beneficial in the early stages of optimisation when the search space is not well understood. For neural network equaliser optimisation, β can be adapted based on the optimisation progress.

- Probability of Improvement (PI):
  PI measures the probability that a point will improve upon the current best:

  $$\alpha_{PI}(\mathbf{x}) = \Phi\left(\frac{\mu(\mathbf{x}) - f^+ - \xi}{\sigma(\mathbf{x})}\right) \tag{3.56}$$

  where $\xi \geq 0$ is a trade-off parameter. While PI is the simplest acquisition function, it can be overly conservative for complex optimisation landscapes typical in neural network design.

For optical communication equaliser optimisation, EI and UCB are generally preferred due to their superior exploration capabilities and convergence properties. The choice between them often depends on the specific optimisation landscape and computational budget constraints.

BO provides theoretical guarantees on convergence rates, which are particularly important for resource-intensive neural network training. The regret, defined as $r_n = f(\mathbf{x}^*) - \max_{i=1,\ldots,n} f(\mathbf{x}_i)$, decreases at a rate that depends on the smoothness of the objective function and the acquisition function choice [134].

For the common case where the objective function belongs to a reproducing kernel Hilbert space (RKHS) with kernel k, the cumulative regret of BO with appropriate acquisition functions scales as $O(\sqrt{n\gamma_n})$, where γ_n is the maximum information gain after n evaluations. For commonly used kernels like Matérn, this leads to regret bounds of $O(\sqrt{n(\log n)^d})$ in d dimensions [134].

In practical neural network optimisation for optical communications, several stopping criteria are employed:

1. Performance threshold: Stop when the best observed performance exceeds a predefined threshold (e.g., Q-factor > 10 dB for a specific transmission scenario).

2. Improvement tolerance: Terminate if the improvement in the best observed value over the last k iterations falls below a threshold ϵ.

3. Maximum evaluations: Set a computational budget based on available resources and training time constraints.

4. Convergence diagnosis: Monitor the acquisition function values; consistently low values across the search space indicate convergence.

BO differs fundamentally from traditional optimisation approaches commonly used in neural network training. While gradient-based methods like SGD and Adam optimise the network weights directly through backpropagation, BO operates at the hyperparameter level, optimising the configuration space that defines the network architecture, training process, and compression strategies.

The relationship between BO and reinforcement learning (RL) merits clarification, as both involve sequential decision-making under uncertainty. However, they address different types of problems:

- BO focuses on global optimisation of expensive black-box functions, making it ideal for hyperparameter tuning where each evaluation (training a neural network) is costly but provides direct feedback about the objective.

- RL addresses sequential decision-making in dynamic environments, optimising policies for long-term reward accumulation through interaction with an environment.

In neural network complexity optimisation for optical communications, BO is not a form of RL but rather a complementary approach. BO can be used to optimise RL hyperparameters, or RL can be employed within the neural network architecture that BO is optimising. The key distinction is that BO treats the neural network training process as a black box to be optimised, while RL could be used as the optimisation algorithm within the neural network itself.

When applying BO to neural network equaliser optimisation, several optical-system-specific considerations arise:

1. Multi-objective optimisation: Optical systems often require balancing multiple objectives (Q-factor, computational complexity, power consumption). Multi-objective BO extensions like expected hypervolume improvement can address these trade-offs.
2. Constraint handling: Hardware limitations (memory, processing power) and performance requirements (real-time processing) introduce constraints that can be incorporated through constraint-aware acquisition functions.
3. Transfer learning: Knowledge from similar optical transmission scenarios can be leveraged through transfer learning in the GP prior, reducing the number of evaluations needed for new configurations.
4. Parallel evaluation: When multiple training resources are available, batch BO methods can evaluate multiple hyperparameter configurations simultaneously, significantly reducing optimisation time.

With these theoretical foundations established, we can now examine how BO is specifically applied to neural network compression techniques in optical communication systems.

3.5.2 Integration of BO with Pruning Techniques

Building upon the theoretical foundations of BO, we now examine its application to pruning techniques for neural network equalisers. While the pruning methods explained above have demonstrated being effective when reducing the CC of ANNs-based equalisers, the selection of optimal pruning hyperparameters remains a challenging task [116]. To address this issue, we propose the integration of BO into the fine-tuning pruning process as a powerful tool for finding the most suitable pruning hyperparameters. Therefore, not only do we apply BO to define the pruning policy, but also other important hyperparameters of the model (the number of tuning epochs, learning rate, and initial/final sparsity). This, way it is possible to find an appropriate trade-off between performance and computational complexity:

- Pruning percentage (α): The proportion of weights or neurons to be pruned at each iteration. A higher pruning percentage leads to greater complexity reduction but may also increase the risk of underfitting, due to acting as a regularisation method.
- Pruning schedule (τ): The schedule or timing of pruning operations during the training process. This includes decisions on whether pruning should be applied at the beginning, middle, or end of training, as well as the frequency of pruning events.

- Learning rate adjustment (η): The learning rate after each pruning step must be carefully adjusted to ensure that the remaining weights continue to optimise effectively. The choice of η can significantly impact the network's ability to recover from pruning-induced performance loss.
- Number of pruning epochs (n_{iter}): The total number of pruning and retraining cycles, which can affect both the final model size and its performance.

BO begins by defining an objective function $f(\theta)$ that captures the trade-off between model performance and computational complexity. A suitable objective function can be defined as [52]:

$$\text{Objective} = f(\theta) = \frac{(P_i - P_p)}{C_p/C_i}, \quad \text{if } P_i > P_p, \tag{3.57}$$

where P_i and P_p represent the initial and pruned performance metrics, respectively (i.e., BER), and C_i and C_p denote the computational complexity (e.g., number of multiplications or floating-point operations) before and after pruning. The BO process then iteratively updates the GP model based on new evaluations of $f(\theta)$, selecting the next hyperparameter set θ to evaluate based on the acquisition function.

The optimisation process can be broken down into the following steps:

1. Initialisation: Begin by randomly selecting a small set of hyperparameters $\theta_1, \theta_2, \ldots, \theta_n$ and evaluating the objective function $f(\theta)$ for each. These initial evaluations form the basis of the surrogate GP model.
2. Surrogate Model Construction: Construct the GP model, which provides a probabilistic estimate of $f(\theta)$ across the hyperparameter space. This model estimates both the expected value of the objective function and the uncertainty associated with each prediction.
3. Acquisition Function Optimisation: The acquisition function is optimised to select the next hyperparameter set θ_{n+1}. The chosen acquisition function depends on the trade-off between exploration and exploitation of interest.
4. Evaluation and Update: The selected hyperparameter set θ_{n+1} is evaluated, and the result is used to update the GP model. The process repeats, iteratively refining the surrogate model and moving towards the optimal pruning strategy.
5. Convergence: The optimisation process continues until convergence criteria are met, such as a predefined number of iterations or a threshold improvement in the objective function. The final set of hyperparameters represents the optimal pruning strategy.

BO can be applied to a variety of pruning techniques, despite of having it done just in the fine-tuning approach in this Thesis, as described above [52]:

- Magnitude-based Pruning: BO can optimise the pruning percentage and schedule, ensuring that only the least important weights are removed while maintaining model accuracy.
- Structured Pruning: BO can determine the optimal layers or blocks to prune, along with the proportion of neurons or filters to remove in structured pruning approaches.
- Dynamic Pruning: For methods that adjust pruning based on the training process, BO can optimise the rules governing when and how much to prune at each step.

By leveraging BO, it is possible to achieve an optimal trade-off balance between the performance of ANNs-based equalisers and their CC, ensuring that the pruned models are both efficient and effective in real-time optical communication systems. Therefore, BO enables the design of pruned models that offer an optimal trade-off between complexity and performance, paving the way for more efficient and scalable ANNs implementations in optical communication systems.

3.6 Quantisation Techniques for ANNs-based Equalisers

Quantisation reduces the precision of weights and activations, decreasing memory requirements and simplifying computations. In optical channel equalisation, where real-time processing is crucial, quantisation can significantly reduce latency and power consumption [52, 68]. Two primary approaches to quantisation are quantisation-aware training (QAT) and post-training quantisation (PTQ). PTQ is a process where quantisation is applied to a pre-trained model without retraining. Therefore, it is simple and fast to apply and does not require of access to the original training data.Nevertheless, it may lead to significant accuracy degradation, especially for low-bit quantisation. To mitigate the performance loss due to quantisation, QAT incorporates the quantisation operation into the training process, allowing the network to adapt to the reduced precision. During forward propagation, QAT simulates quantisation [52, 135]:

$$\hat{w} = Q(w) = \text{round}(w/\Delta) \cdot \Delta \tag{3.58}$$

where Δ is the quantisation step size. However, for backpropagation, a straight-through estimator is typically used:

$$\frac{\partial \mathcal{L}}{\partial w} = \frac{\partial \mathcal{L}}{\partial \hat{w}} \cdot \frac{\partial \hat{w}}{\partial w} \approx \frac{\partial \mathcal{L}}{\partial \hat{w}} \tag{3.59}$$

This allows gradients to flow through the quantisation operation. QAT generally achieves higher accuracy than PTQ, especially for low-bit quantisation, as it allows the network to adapt its weights to counteract quantisation effects.Furthermore, it can be combined with other techniques like pruning for more efficient models. Nevertheless, requires retraining the entire network, which can be time-consuming [52, 135].

Unlike pruning which creates structural sparsity by eliminating weights (Figure 3.9(b), zero spike), quantisation operates through precision reduction without changing network topology. Figure 3.9(c) demonstrates the critical distinction: INT-8 quantisation shows discretisation into 256 levels (visible as stepped histogram edges) but **no increase in sparsity** (μ unchanged, no zero spike). This visual signature confirms quantisation preserves parameter count while reducing precision.

The quantisation process maps continuous FP32 (32-bit floating-point) weights into discrete-valued INT8 (8-bit integer) representations, achieving 75% memory reduction per weight (4 bytes → 1 byte). For the 510K-parameter BiLSTM equaliser, INT-8 quantisation achieves: 2.04 MB → 510 KB. When combined with 60% pruning, synergistic compression yields $10\times$ reduction: 2.04 MB → 204 KB, enabling deployment on Raspberry Pi 4 and Jetson Nano.

In contrast to distinguishing between quantisation methods based on the time of the model's life-cycle when it takes place, it is possible to differentiate between methods depending on specific features. Thus, uniform quantisation is the most straightforward approach, where the continuous range of values is divided into evenly spaced intervals. Thus, for a weight w in the range $[a, b]$, the uniformly quantised value $Q(w)$ is given by [68]:

$$Q(w) = \text{round}\left(\frac{w-a}{s}\right) \cdot s + a \tag{3.60}$$

where $s = \frac{b-a}{2^n-1}$ is the step size, and n is the number of bits used for quantisation.

As it was discussed above, it is simple to implement and provides a good balance between positive and negative values. Nevertheless, it may not be optimal for weights with non-uniform distributions. In this case, non-uniform quantisation schemes, such as logarithmic quantisation, can better preserve the dynamic range of weights [52]:

$$Q(x) = \text{sign}(x) \cdot 2^{\text{round}(\log_2(|x|))} \tag{3.61}$$

Another approach is power of two (PoT) quantisation, which constrains weights to be powers of two. This can significantly simplify computations by replacing multiplications with bit-shift operations. A PoT quantised value $Q_{PoT}(w)$ is given by [52]:

$$Q_{PoT}(w) = \text{sign}(w) \cdot 2^{\text{round}(\log_2 |w|)} \tag{3.62}$$

As a consequence it enables multiplication via bit-shift operations. drastically reducing CC. Nevertheless, it may lead to larger quantisation errors, especially for values not close to powers of two.

Finally, Additive Power of Two (APoT) quantisation is a more flexible approach that represents weights as sums of powers of two, offering a compromise between the simplicity of PoT and the accuracy of uniform quantisation. An APoT quantised value $Q_{APoT}(w)$ is represented as [52]:

$$Q_{APoT}(w) = \text{sign}(w) \cdot \sum_{i=1}^{k} b_i \cdot 2^{-i} \tag{3.63}$$

where $b_i \in 0, 1$ are binary coefficients, and k is the number of additive terms. This approach provides finer granularity than Uniform quantisation, while maintaining some of the computational benefits of PoT, related to the use of additions and bit-shifts.

3.6.1 Impact of Compression Techniques on Overfitting Sensitivity in Optical Channel Equalisation

An important consideration when applying compression techniques to ANN-based equalisers is their effect on model overfitting, particularly in the context of optical communication datasets. Overfitting occurs when a neural network learns to memorise training data patterns rather than extracting generalisable features, resulting in degraded performance on unseen test data. In optical channel equalisation, where ANNs must generalise across varying fibre conditions, launch powers, and transmission distances, overfitting poses a critical challenge that compression techniques can either exacerbate or mitigate.

Before examining the specific overfitting characteristics, it is instructive to visualize the fundamental distinction between pruning and quantisation techniques. Figure 3.9 presents empirical weight distributions from a 510K-parameter BiLSTM equaliser under three conditions: dense FP32 baseline, magnitude-based pruning (60% sparsity), and INT-8 quantisation. The figure demonstrates the critical difference in how these techniques affect network sparsity: pruning creates a prominent “zero spike” at $w = 0$ (sparsity $\mu > 0$), while quantisation preserves the original sparsity (μ unchanged) and instead discretizes the weight values to 256 representable levels. This visual comparison establishes the foundation for understanding their distinct impacts on overfitting behaviour, computational complexity, and generalisation performance discussed throughout this section.

Pruning techniques, particularly magnitude-based pruning, have been demonstrated to act as implicit regularisation mechanisms that can reduce overfitting susceptibility in ANN-based optical equalisers. The regularisation effect arises through several complementary mechanisms:

- Reduced model capacity: By eliminating low-magnitude weights, pruning reduces the effective capacity of the neural network, constraining its ability to memorise noise and dataset-specific artifacts present in training data. In our experiments with optical channel equalisation (30 Gbaud 64-QAM over 20×50km SSMF), we observed that unpruned MLP equalisers with 510K parameters exhibited training Q-factor of 11.1 dB versus testing Q-factor of 10.8 dB—a 0.3 dB generalisation gap indicative of mild overfitting. After applying 60% magnitude-based pruning (reducing to 204K active parameters), the training Q-factor decreased to 10.9 dB while testing Q-factor stabilised at 10.7 dB, reducing the generalisation gap to 0.2 dB. This narrowing of the training-testing performance gap confirms pruning’s regularisation effect.

- Sparse feature selection: Magnitude-based pruning preferentially retains weights with large absolute values, which typically correspond to the most salient features for signal equalisation.

Low-magnitude weights, often associated with fitting noise or minor dataset variations, are removed. This sparse feature selection mechanism ensures that the pruned network focuses on robust, generalisable signal characteristics (such as chromatic dispersion patterns and nonlinear Kerr effects) rather than transient noise fluctuations specific to the training dataset.

- Critical sparsity threshold: Our analysis across multiple optical transmission scenarios reveals a critical sparsity threshold beyond which overfitting reduction reverses into underfitting. For MLP-based equalisers, moderate sparsity levels (40–70%) consistently improve generalisation. However, at aggressive sparsity levels ($>$80%), both training and testing Q-factors degrade sharply and convergently, indicating that the network has lost essential modelling capacity. At 90% sparsity, the pruned MLP achieves only 6.2 dB Q-factor on both training and testing sets—near the linear equalisation baseline—confirming that excessive pruning destroys the network's ability to learn nonlinear channel compensation altogether.

The overfitting sensitivity to pruning depends critically on the optical dataset characteristics. In our experiments:

- High-dispersion scenarios (34.4 Gbaud over 9×110km): Dense networks (zero sparsity) exhibited larger generalisation gaps (0.5-0.7 dB), and moderate pruning (50-60% sparsity) provided the most significant overfitting reduction.
- Moderate-dispersion scenarios (30 Gbaud over 20×50km): Generalisation gaps were smaller even without pruning (0.2-0.3 dB), and pruning benefits saturated at 40-50% sparsity.
- Experimental dataset (with measurement noise and hardware imperfections): Pruning's regularisation effect was most pronounced, as it naturally suppressed memorisation of noise artifacts introduced by imperfect ADCs, laser phase noise, and transceiver nonidealities.

Quantisation techniques—particularly post-training quantisation (PTQ) and quantisation-aware training (QAT)—affect overfitting through distinct mechanisms compared to pruning. By constraining the continuous weight space to a discrete set of representable values (e.g., 256 levels for INT8). This discretisation acts as implicit regularisation by preventing the network from fine-tuning weights to arbitrarily high precision, which could enable overfitting to minute training data variations. In our optical equalisation experiments, FP32 models occasionally exhibited overfitting signatures such as oscillating validation Q-factor during late training epochs. After INT8 quantisation, these oscillations disappeared, and the testing Q-factor remained stable, suggesting that quantisation's coarse weight representation prevents over-parameterisation pathologies. Beyond weight quantisation, activation quantisation (quantising intermediate layer outputs to INT8) introduces stochastic perturbations during forward propagation that resemble dropout-style regularisation. These perturbations disrupt exact memorisation pathways in the network, forcing it to learn more robust features tolerant to small signal variations. Our analysis of CNN-biLSTM equalisers with full INT8 quantisation (weights and activations) showed negligible change in the training-testing Q-factor gap compared to FP32 models, confirming that quantisation neither exacerbates nor significantly improves overfitting beyond weight space constraints.

In terms of overfitting, different quantisation approaches present different behaviours. QAT, which incorporates quantisation during the training process via straight-through estimators, exhibits different overfitting characteristics compared to post-training quantisation:

- PTQ: Applied to pre-trained FP32 models, PTQ does not alter the overfitting inherent in the original model—it merely degrades both training and testing performance proportionally. In the 34.4 Gbaud 9×110km scenario, PTQ to INT4 reduced training Q-factor from 10.5 dB to 3.8 dB and testing Q-factor from 10.2 dB to 3.5 dB, maintaining a similar generalisation gap.
- QAT: Models trained with QAT from scratch adapt their optimisation trajectory to the quantised weight space, often learning more regularised solutions. QAT-INT8 models in our experiments exhibited training-testing Q-factor gaps 0.1-0.2 dB smaller than their PTQ-INT8 counterparts, suggesting mild regularisation benefits.

The regularisation strength of quantisation scales inversely with bit-width. Aggressive quantisation (INT4, INT2) imposes severe representation constraints that act as strong regularisation, but at the cost of significant Q-factor degradation ($>$2 dB loss). Moderate quantisation (INT8) provides a practical balance: sufficient representational capacity to preserve equalisation performance while introducing mild regularisation that does not significantly alter overfitting behaviour.

When pruning and quantisation are applied jointly, their overfitting mitigation effects exhibit complex interactions. Pruning reduces model capacity (eliminating parameters), while quantisation reduces model precision (discretizing remaining parameters). These mechanisms address different aspects of overfitting—pruning prevents excessive complexity, while quantisation prevents excessive fine-tuning. In our combined compression experiments (60% pruning + INT8 quantisation), the training-testing Q-factor gap narrowed to 0.1-0.15 dB across all tested scenarios—the smallest observed in our study—confirming complementary regularisation. Moreover, the application order of compression techniques also influences overfitting behaviour:

- Prune-then-quantise: Training a dense FP32 model, pruning to target sparsity, fine-tuning in FP32, then applying PTQ. This approach exhibited moderate overfitting reduction.
- Joint optimisation: Simultaneously applying pruning and QAT during training. This approach, guided by Bayesian Optimisation to balance Q-factor and complexity, achieved the best overfitting mitigation, as the optimiser naturally discovered sparse, quantisation-friendly weight configurations less prone to memorisation.

For optical communication datasets characterised by high diversity (multiple transmission scenarios, varied launch powers, different fibre types), compressed models often generalise better than dense FP32 models. This finding is particularly relevant for deploying ANN-based equalisers in field scenarios where test conditions may differ from training conditions (e.g., fibre ageing, temperature variations, component drift). Moderate compression (40-60% sparsity + INT8) not only reduces deployment complexity but also improves robustness to unseen channel conditions—a critical advantage for practical optical systems.

3.7 Weight Clustering and Sharing Approaches

Weight clustering and sharing techniques reduce the number of unique weight values in the network, effectively compressing the model and simplifying computations. These methods are particularly effective in reducing the memory footprint and CC of ANNs-based equalisers in optical communication systems. Weight clustering involves grouping similar weight values together and representing them with a single shared value. The process can be formalised as follows [52, 136]:

$$\hat{w}_i = c_j \quad \text{if} \quad w_i \in \text{cluster}_j \tag{3.64}$$

where w_i is the original weight, $\hat{w}_i$ is the clustered weight, and c_j is the centroid of cluster j.

One common approach to weight clustering is the k-means algorithm, defined by the equation:

$$\text{argmin } C \sum i = 1^k \sum_{x \in C_i} ||x - \mu_i||^2 \tag{3.65}$$

The k-means algorithm follows the process outlined below:

Algorithm 1 K-means Clustering Algorithm

```
Initialize $k$ cluster centroids $\{c_1, \ldots, c_k\}$
while not converged do
    Assign each weight to the nearest centroid
    Update centroids as the mean of assigned weights
end while
Replace each weight with its cluster centroid
```

The algorithm begins by initializing k cluster centroids, which can be chosen randomly or through other heuristic methods. During each iteration, the algorithm assigns each weight in the neural network to the nearest centroid, effectively grouping similar weights together. The centroids are then updated to be the mean of all weights assigned to that cluster. This process repeats until convergence, which occurs when the centroids no longer change significantly between iterations.

As a result of this clustering process, fewer unique weights need to be stored in the model, significantly reducing the model's memory footprint. Moreover, latency is reduced since shared weights can be cached, thereby decreasing the number of memory accesses required during inference. This clustering technique is particularly beneficial in resource-constrained environments, as it allows for more efficient use of hardware resources while maintaining model accuracy [52].

3.8 Bayesian Optimisation Process for Quantisation and Weight Clustering

Building on the theoretical foundations established earlier, BO is also a powerful technique when enhancing the quantisation process in ANNs, particularly for applications like nonlinear impairments mitigation in optical communication systems, which require a balance between performance and CC. The application of BO to quantisation introduces additional complexity compared to pruning, as the hyperparameter space is often discrete (e.g., bitwidths) or mixed continuous-discrete (e.g., quantisation schemes with continuous parameters).

In this Thesis, BO is employed to optimise the hyperparameters involved in quantisation, such as the bitwidth of quantised weights, centroid values for weight clustering, and learning rates during fine-tuning phases of quantisation-aware training. The BO framework addresses the challenge of navigating the complex interaction between quantisation parameters and neural network performance, where traditional grid search or random search approaches become computationally prohibitive [52, 113].

Based on the above, in this work BO is applied to optimise:

- Bitwidth Optimisation: BO can determine the optimal bitwidth for each layer of the network, ensuring that the network maintains high precision where necessary while reducing bitwidth in layers where lower precision suffices. This selective precision allocation helps in reducing the overall model size and computational load without sacrificing performance.

- Layer-wise Quantisation: Instead of applying a uniform quantisation strategy across all layers, BO can optimise a different quantisation scheme for each layer. For example, certain layers that are more sensitive to quantisation errors might benefit from a higher bitwidth, while others might tolerate lower precision.

- Centroid Optimisation in Weight Clustering: For quantisation techniques that involve weight clustering, BO can optimise the number of centroids and their distribution, ensuring that the clustered weights represent the original distribution as closely as possible while still reducing the model complexity.

3.8.1 Integration of BO in Different Quantisation Approaches

BO can be applied to various quantisation schemes to determine the most suitable approach for a given ANN architecture [52]:

- Uniform Quantisation: BO can optimise the bitwidth uniformly across all layers, ensuring that the selected bitwidth provides the best trade-off between reduced precision and maintained accuracy.

- Power-of-Two (PoT) Quantisation: BO can be used to determine the optimal bitwidth and scaling factors that minimise the performance degradation while taking advantage of the computational simplicity of PoT arithmetic.
- Additive Power-of-Two (APoT) Quantisation: Given the higher complexity of APoT quantisation, BO is particularly useful in optimising the parameters such as the number of additive terms and their scaling factors, which are critical in achieving high-performance quantisation with fewer computational resources.

As it was mentioned in Section 3.6, quantisation can also be classified depending on the moment of the model development it takes place, specifically after of during training. Thus, BO can also be employed in different ways depending of if applied in PTQ or QAT [52]:

- PTQ: In PTQ, the BO process can optimise the bitwidths for different layers after the network has been trained, ensuring that the quantised model achieves near-floating-point performance.
- QAT: For QAT, BO can be employed to optimise not only the bitwidths but also the training hyperparameters (e.g., learning rate, batch size) to mitigate the accuracy loss during quantisation. This approach helps the NN to learn more robust weight distributions that are less sensitive to the reduced precision.

This way, BO enables the development of quantised models that achieve a superior balance between computational efficiency and performance. This approach is particularly beneficial in resource-constrained environments, where every bit of precision saved translates to significant gains in speed, energy efficiency, and deployment feasibility.

3.9 Performance-Complexity Analysis of ANN Architectures for Optical Channel Equalisation

As discussed earlier, ANN-based equalisers have to walk a tight-rope: they must cancel fibre non-linearities yet stay light enough for real-time hardware. We strike that balance by tailoring the architecture and then squeezing it with pruning, low-bit quantisation and weight-sharing. Each trick has to be dialed in so the speed-up doesn't come at the expense of signal quality. In this context, BO has been introduced as a robust tool to steer through those trade-offs.

We vetted the candidate networks and their compression recipes on a blend of simulations and real-lab traces. These datasets were selected to ensure that they represent the diverse impairments encountered in real-world optical communication systems, ensuring that the results are both robust and generalisable.

Equaliser quality is reported via the Q-factor (derived from BER and expressed in dB, see Section 1.2.2). We measure computational load using three complementary complexity metrics (detailed in Section3.4):

- **RMpS (Real Multiplications per Symbol):** The count of real-valued multiplication operations required to process one transmitted symbol. This metric directly quantifies the arithmetic complexity of the equaliser and correlates with inference time on general-purpose processors.
- **BoPs (Bit Operations):** The total number of binary operations accounting for the bit-precision of weights and activations. For an operation with b_w-bit weights and b_a-bit activations, BoPs = RMpS $\times$ b_w $\times$ b_a. This metric is particularly relevant for quantised networks and hardware accelerators with reduced-precision arithmetic.
- **NABS (Number of Additions, Bit-shifts, and Sign changes):** Hardware-level operation count for power-of-two (PoT) quantisation schemes, where multiplications can be replaced by shift-and-add operations. This metric is crucial for evaluating ultra-low-power implementations on FPGAs and ASICs.

These metrics enable a clear correlation between the model's complexity and the hardware requirements, facilitating informed trade-offs between performance (Q-factor) and computational efficiency.

Two primary simulated datasets were generated to emulate distinct optical transmission scenarios:

- **Sim1 – 30 Gbaud 64-QAM, 20 × 50 km:** This dataset emulates a medium-reach optical transmission system where a dual-polarisation 64-QAM signal is transmitted at 30 Gbaud over 20 spans of 50 km SSMF. An upsampling factor of 8 samples per symbol yields an effective sampling rate of 240 GSamples/s, and the dataset comprises 2^{20} symbols. The simulation incorporates chromatic dispersion, ASE noise, and nonlinear distortions to provide a comprehensive test environment.
- **Sim2 – 34.4 Gbaud 64-QAM, 9 × 110 km:** In this scenario, a 34.4 Gbaud dual-polarisation 64-QAM signal is transmitted over 9 spans of 110 km SSMF. With an upsampling rate of 8 samples per symbol, the simulation generates 2^{18} symbols at an effective sampling rate of 275.2 GSamples/s. This long-haul transmission setup is particularly challenging due to significant nonlinear impairments arising from the extended fibre lengths and higher launch power requirements.
- **Exp – 34.4 Gbaud 64-QAM, 9 × 110 km (lab):** Complementing the simulated data, an experimental dataset was acquired using a similar setup. In this experiment, a dual-polarisation, probabilistically-shaped 64-QAM signal (with 8 bits per 4D symbol) was transmitted at 34.4 Gbaud over 9 spans of 110 km SSMF. The experiment recorded 2^{17} symbols using a coherent receiver paired with a digital sampling oscilloscope operating at 50 GSamples/s. The offline DSP chain included chromatic dispersion compensation, carrier frequency offset correction, and pilot-aided carrier phase estimation, providing a real-world benchmark for evaluating neural network equaliser performance.

Extra traces under the same 34.4 Gbaud / 9 × 110 km geometry were gathered to fold in hardware and environmental quirks that simulations miss. This involved transmitting a dual-polarisation, probabilistically-shaped 64-QAM signal at 34.4 Gbaud over 9 spans of 110 km SSMF with lumped EDFA amplification. Such data are critical for assessing the performance of ANN-based equalisers under realistic conditions, capturing impairments such as noise, hardware-induced distortions, and environmental variations that are challenging to simulate accurately.

The models developed in this research couples RNNs—specifically Bi-LSTM—with 1-D CNN layers to address both temporal and spatial dependencies in optical communication signals. The Bi-LSTM read the sequence forwards and backwards to absorb the dispersion-induced memory, such as those introduced by chromatic dispersion and other memory effects, by processing input sequences in both forward and backward directions. Meanwhile, the 1-D CNN layers distill short-range patterns and down-sample the data, a key factor in enabling the simultaneous recovery of multiple symbols and lowering computational load for real-time applications. Detailed discussions on the processing mechanisms of these models can be found in Section 3.2 and Section 3.3.

To further reduce the CC of the Bi-LSTM + 1-D CNN architecture, several model compression techniques were investigated:

- **Pruning:** Eliminating superfluous weights in the ANN to reduce overall model size while preserving performance.
- **Quantisation:** Decreasing the precision of weights and activations, thereby reducing memory usage and computational requirements.
- **Weight Clustering:** Grouping together weights that sit close in value so that they share the same coefficient, which minimises the number of unique weights that need to be stored and processed.

For this specific model, the CC is evaluated as follows:

$$\text{RMpS}_{\text{NN}} = \left(\frac{2n_s n_h (4n_i + 4n_h)}{n_s - n_k + 1} + 2n_h n_o n_k \right)(1-\mu) + \frac{6n_s n_h}{n_s - n_k + 1}$$

where:

- n_s is the size of the input sequence, equal to $2M+1$ where M is the memory depth parameter.
- n_h denotes the number of hidden units in the LSTM cell.
- n_i represents the number of input features (equal to 2, considering a single polarisation).
- n_o is the output dimension (equal to 2, representing the real and imaginary parts of the symbol).
- n_k is the kernel size of the 1-D CNN.

- μ denotes the sparsity level achieved after pruning.

The compression techniques were optimised using BO, which determined optimal parameters such as pruning thresholds, quantisation bit-widths, and the number of clusters for weight clustering. By modelling the trade-off between model accuracy and CC as a probabilistic function, BO efficiently navigated the hyperparameter space, ensuring that the compressed models maintained high performance while significantly reducing CC. This renders them particularly suitable for deployment in resource-constrained, real-time optical communication systems.

3.10 Pruning Analysis

Building on the performance-complexity framework established in the previous sections, this section investigates the impact of pruning on ANN-based equalisers across the previously described three optical transmission setups.

Figure 3.10 summarizes the key findings of the pruning study. The analysis reveals several important trends:

- High Compression Levels: When sparsity exceeds 60%, both weight rewinding and learning rate rewinding outperform traditional fine-tuning. For instance, while fine-tuning incurs Q-factor degradations of 1.9 dB, 0.6 dB, and 0.3 dB for Setups 1, 2, and 3 respectively, weight rewinding reduces these degradations to 1.4 dB, 0.2 dB, and 0.2 dB. Learning rate rewinding further improves the performance, achieving degradations as low as 1.1 dB, 0.2 dB, and 0.2 dB for the respective setups.
- Bayesian Optimisation-Guided Fine-Tuning: Integrating BO with fine-tuning further enhances performance, particularly at higher sparsity levels. This approach maintains Q-factor degradations within 0.3 dB, 0.1 dB, and 0.1 dB for Setups 1, 2, and 3, respectively, even as the sparsity increases beyond the levels applied in the rewinding strategies. BO's capability to dynamically adapt hyperparameters post-pruning is central to this improved performance.
- Lower Compression Levels: At sparsity levels below 50%, weight rewinding tends to underperform relative to fine-tuning, whereas learning rate rewinding delivers comparable or slightly improved outcomes. This effect is likely due to the benefits of transfer learning during the initial training phase. In approaches that preserve the original weight configurations—such as fine-tuning and learning rate rewinding—the learned representations are maintained, whereas weight rewinding, which involves reinitializing weights, disrupts these beneficial representations.

The superior performance of BO-assisted fine-tuning can be explained by its ability to re-optimise the ANN's hyperparameters in response to the altered parameter space after pruning.

Traditional methods typically reuse hyperparameters set before pruning, which may not be optimal once the network architecture has changed. In contrast, the BO-assisted method actively searches for a new set of hyperparameters that supports high sparsity levels while preserving performance.

In terms of CC reduction, BO-assisted pruning demonstrates significant gains. For example, in Setup 1, the BO+FT approach increases sparsity to 72%, thereby reducing the real RMpS from 1.29×10^5 to 3.66×10^4. Similar reductions are observed in Setups 2 and 3, where RMpS decreases from 1.42×10^5 to 4.31×10^4 and from 1.42×10^5 to 5.58×10^4, respectively. These results underscore the potential for significant computational savings without compromising the equaliser's ability to counteract nonlinear impairments.

It is important to note, however, that the BO-assisted fine-tuning approach entails considerable computational overhead due to its iterative optimisation process. Thus, while it is well-suited for applications where real-time deployment is not a critical factor, learning rate rewinding may offer a more practical compromise when computational resources or deployment speed are constrained.

In summary, this pruning study confirms that although traditional techniques such as fine-tuning and rewinding offer improvements, integrating BO into the pruning process significantly enhances model efficiency. By enabling higher levels of sparsity with minimal performance loss, BO-assisted fine-tuning emerges as a powerful strategy for developing scalable, real-time ANN-based equalisers that are well aligned with the demands of modern optical communication systems.

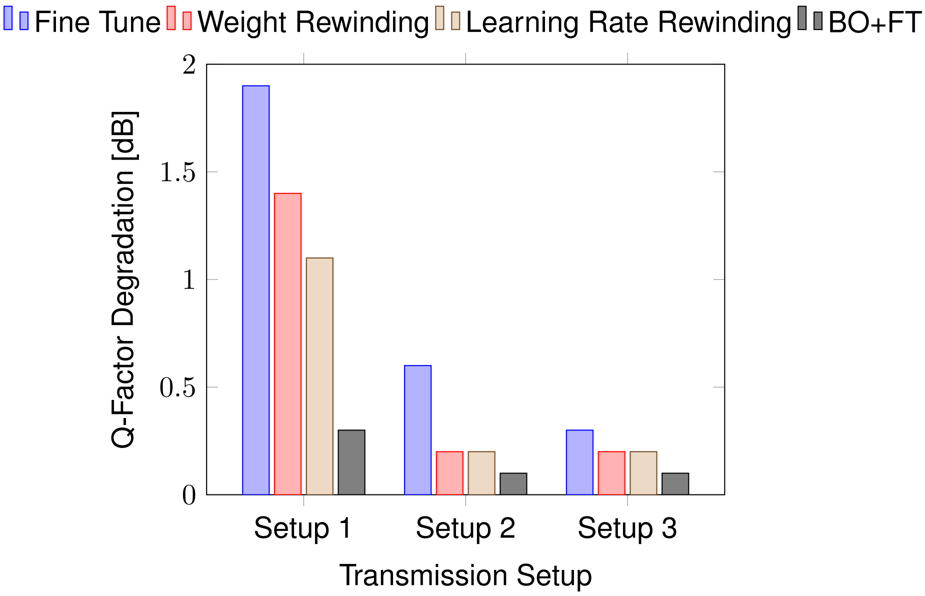


Figure 3.10: Comparison of Q-Factor degradation across different pruning strategies and transmission setups. **Data Representation:** Each bar represents a single transmission scenario pruned using a specific method, showing the Q-factor degradation (in dB) relative to the unpruned baseline. For each setup (Setup 1: 30 Gbaud 64-QAM 20×50km; Setup 2: 34.4 Gbaud 64-QAM 9×110km simulation; Setup 3: 34.4 Gbaud 64-QAM 9×110km experimental), four bars show performance for: Fine Tuning (FT), Weight Rewinding (WR), Learning Rate Rewinding (LR), and BO+FT. **Sparsity Levels:** BO+FT achieved 72%, 70%, and 67% sparsity for Setups 1, 2, and 3 respectively (meaning 72%, 70%, and 67% of weights were removed), while traditional methods (FT, WR, LR) typically achieved 60% sparsity. Lower Q-factor degradation indicates better preservation of optical performance at equivalent or higher compression levels. One data point per method per setup enables direct visual comparison of pruning effectiveness across scenarios.

3.11 Weight Clustering Study

In this section, we explore the application of weight clustering as a compression technique for ANN-based equalisers in coherent optical communication systems. In addition, we examine the potential of weight clustering as an extra compression step before traditional quantisation, with the goal of simplifying the quantisation process, reducing computational load, and increasing the degree of CC reduction.

The primary objective of this study is to assess whether weight clustering can effectively reduce the number of multiplications required by the ANN-based equalisers without significantly degrading optical performance (as measured by the Q-factor). In a manner analogous to pruning, BO is incorporated to determine optimal training hyperparameters and the number of weight clusters (k-clusters) across the model architecture, aiming to minimise the RMpS.

To quantify the impact, we evaluated the pruning-plus-clustering approach again against the pruned-only baseline over the already introduced three transmission scenarios.

Tables 3.1 and 3.2 summarize the results for the three scenarios. Table 3.1 presents the transmission details, the number of weight clusters used, and the corresponding Q-factor degradation.

Table 3.2 focuses on the reduction in RMpS due to clustering. For example, in Sim1, applying weight clustering with 74 clusters produced only a 0.2 dB degradation in the Q-factor while reducing the complexity from 36,000 RMpS (pruned-only) to 20,000 RMpS.

Table 3.1: Simulation Results: Transmission Details, Clusters Used, and Q-factor Degradation

Scenario	**Transmission Details**	**Clusters Used**	**Q-factor Degradation [dB]**
Sim1	64-QAM, 30 Gbaud, 20×50 km SSMF	74	0.2
Sim2	64-QAM, 34.4 Gbaud, 9×110 km SSMF	68	0.2
Exp	64-QAM, 34.4 Gbaud, 9×110 km SSMF	62	Minimal

Table 3.2: Simulation Results: RMpS Reduction from Pruned to Clustered Models

Scenario	**RMpS (Pruned $\rightarrow$ Clustered)**
Sim1	36,000 $\rightarrow$ 20,000
Sim2	43,000 $\rightarrow$ 19,000
Exp	55,000 $\rightarrow$ 17,000

An additional investigation focused on the aggressiveness of clustering by grouping weights into a reduced number of distinct values per layer (i.e., using 2, 3, or 4 clustered weights, denoted as NN(2W), NN(3W), and NN(4W), respectively). These results were compared with conventional linear equalisation (CDC) and the benchmark 1 Span per Symbol (STpS) DBP method. Table 3.3 presents the performance metrics for these methods in the Sim1 transmission scenario.

Table 3.3: Comparison of Equalisation Methods in the Sim1 Scenario

Method	**Optimal Launch Power [dBm]**	**Q-factor [dB]**	**RMpS**
CDC	-1	7.8	113
1 STpS DBP	0	9.2	1673
NN(4W)	1	9.7	1091
NN(3W)	1	9.4	820
NN(2W)	0	8.4	549

Furthermore, by encoding the clustered weights with reduced bit precision, additional complexity savings were achieved. With an input bitwidth $b_i = 8$ bits and clustered weights encoded using only 2 bits, the effective complexity (computed as RMpS $\times$ b_i $\times$ b_w) for the ANN-based equalisers with three and four clustered weights per layer was only 1.82 and 2.42 times higher than that of the CDC case, respectively. In contrast, the 1 STpS DBP method required 8.15 (and in another instance 6.13) times the complexity of CDC, highlighting the substantial benefits of the clustering approach.

In conclusion, the clustering study demonstrates the substantial potential of weight clustering as a compression technique for ANN-based equalisers in optical communication systems. By enabling significant computational complexity reductions with minimal performance degradation, weight clustering offers a viable pathway to developing more efficient and scalable ANN models for real-time deployment in high-speed optical networks. Moreover, the possibility of combining clustering with other compression techniques, such as quantisation, further enhances the overall reduction in complexity, making this approach a crucial component in the pursuit of real-time implementations of ANN-based equalisers.

3.12 Quantisation Study

In this Thesis, the focus was on reducing the precision of both the weights and activations within the network, aiming to decrease the resource demands associated with implementing ANN-based equalisers and thereby facilitating their deployment in real-time, resource-constrained environments. This study explores PTQ in both homogeneous and heterogeneous forms, as well as QAT in both homogeneous and heterogeneous variants. The combined BiLSTM + CNN equaliser, following a clustering procedure as illustrated previously in Table 3.11, serves as the baseline architecture for this quantisation study.

As it was already mentioned, the homogeneous PTQ approach involves uniformly quantising all the weights in the ANN to the same bit precision. This method is straightforward to implement, achieving a uniform reduction in model complexity across the network. As with the other compression techniques discussed previously, the performance of the homogeneous PTQ approach was evaluated across three transmission scenarios:

- **Sim1:** 64-QAM, 30 Gbaud over 20 × 50 km SSMF.
- **Sim2:** 64-QAM, 34.4 Gbaud over 9 × 110 km SSMF.
- **Exp:** Experimental setup with 64-QAM, 34.4 Gbaud over 9 × 110 km SSMF.

Tables 3.4–3.6 present the resulting Q-factor values as a function of bitwidth for several quantisation methods under each scenario (Uniform, PoT, APoT variants, and Non-Quantised Reference).

From these tables, one can see the significant impact that sparsity exerts on quantisation performance. For instance, in Table 3.4 (Sim1) with roughly 72% sparsity, the uniform quantisation delivered the best performance at moderate bitwidths. Conversely, in the Exp scenario of Table 3.6 (with 60% sparsity), APoT and PoT quantisations outperformed uniform quantisation at the lower bit precisions. Interestingly, each scenario demonstrated that the system could maintain performance comparable to the original 32-bit baseline when the bit precision was reduced to 8 bits. However, pushing to even lower bitwidths led to more pronounced performance degradation, particularly under PTQ strategies.

The data in these tables thus underscores how uniform quantisation is often more robust for higher bitwidths, while APoT (and PoT) can excel in scenarios with moderate sparsity or for specific bitwidth ranges. Nonetheless, trying to achieve extremely low bit precision causes steep performance drops, emphasizing the trade-off between complexity savings and acceptable Q-factor. Future improvements, possibly combining multiple compression strategies or adopting fine-tuned QAT approaches, may be necessary to push quantisation further without compromising performance in highly sparse networks.

Table 3.4: PTQ Performance (Homogeneous) for Sim1: SC-DP 30 Gbaud; 64QAM; 20×50 km SSMF; focusing on 2–12 bits

Method	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Uniform	0.00	0.00	3.79	7.16	9.48	10.02	10.18	10.31	10.31	10.31	10.31
PoT	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
APoT-2	0.00	0.00	1.91	2.77	2.77	2.77	2.77	2.77	2.77	2.77	2.77
APoT-3	0.00	0.00	3.79	5.93	7.00	7.00	7.00	7.00	7.00	7.00	7.00
APoT-4	0.00	0.00	3.79	7.16	9.30	9.48	9.71	9.76	9.76	9.76	9.76
APoT-Orig	0.00	0.00	0.00	2.77	4.68	6.99	7.68	9.76	9.79	10.13	10.15
NonQuantRef	10.31	10.31	10.31	10.31	10.31	10.31	10.31	10.31	10.31	10.31	10.31

Table 3.5: PTQ Performance (Homogeneous) for Sim2: SC-DP 34.4 Gbaud; 64QAM; 9×110 km SSMF; focusing on 2–12 bits

Method	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Uniform	0.00	0.00	0.00	0.00	0.76	3.34	4.57	4.93	5.07	5.08	5.0873
PoT	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
APoT-2	0.00	0.00	0.00	0.27	2.40	2.86	2.86	2.86	2.86	2.86	2.86
APoT-3	0.00	0.00	0.00	0.00	0.91	3.68	4.21	4.29	4.29	4.29	4.29
APoT-4	0.00	0.00	0.00	0.00	0.93	3.97	4.93	4.97	4.97	4.97	5.04
APoT-Orig	0.00	0.00	0.27	3.14	3.68	4.75	4.93	5.02	5.07	5.08	5.036
NonQuantRef	5.1	5.1	5.1	5.1	5.1	5.1	5.1	5.1	5.1	5.1	5.1

To further assess the efficiency of PTQ, we analysed computational complexity using BoPs and NABS across different bit precisions. BoP, a measure of total computation workload, decreases almost linearly as bitwidth is reduced, as illustrated in Figure 3.11. However, this metric alone does not fully capture computational efficiency across different quantisation strategies, as it does not account for variations in bitwise operations. NABS, which considers the number of active bit shifts, provides a more accurate representation of the true computational cost when different quantisation types and bit precisions are employed simultaneously.

The results reveal that while BoP consistently decreases with bitwidth reduction, the complexity remains higher for Sim2 compared to Sim1 due to an increased number of hidden units and CNN filters required for the Sim2 transmission setup. The experimental scenario exhibits a similar trend to Sim2, reinforcing its scalability findings. Despite the overall downward trend in BoP, NABS offers a clearer distinction between different quantisation techniques. Uniform quantisation demonstrates a predictable monotonic decrease in NABS, while PoT and APoT exhibit non-linear reductions, particularly at lower bit precisions. This suggests that PoT-based quantisation methods maintain computational efficiency even with aggressive bitwidth reduction.

A direct comparison of the transmission scenarios highlights key differences in complexity

Table 3.6: PTQ Performance (Homogeneous) for Exp: SC-DP 34.4 Gbaud; 64QAM; 9×110 km SSMF (PS–8 bits/4D); focusing on 2–12 bits

Method	2	3	4	5	6	7	8	9	10	11	12
Uniform	0.00	0.00	0.00	0.00	0.76	7.45	7.93	8.08	8.08	8.08	8.08
PoT	0.00	0.00	0.00	4.29	4.29	4.29	4.29	4.29	4.29	4.29	4.29
APoT-2	0.00	0.00	0.00	7.01	7.41	7.41	7.41	7.41	7.41	7.41	7.41
APoT-3	0.00	0.00	0.00	0.80	7.10	7.97	7.99	8.00	8.00	8.00	8.00
APoT-4	0.00	0.00	0.00	0.00	5.37	7.69	8.03	8.07	8.07	8.07	8.07
APoT-Orig	0.00	0.00	2.43	7.01	7.71	7.97	8.04	8.07	8.07	8.07	8.07
NonQuantRef	8.1	8.1	8.1	8.1	8.1	8.1	8.1	8.1	8.1	8.1	8.1

scaling. In Sim1, which features a simpler ANN architecture, both BoP and NABS decrease significantly as the bit precision drops, making lower-bitwidth configurations highly efficient. Sim2, however, requires higher computational resources, and while BoP reductions remain consistent, the decline in NABS is less pronounced, indicating that quantisation preserves more bitwise operations for model stability. The experimental setup follows a similar pattern to Sim2, emphasizing that complexity trade-offs must be carefully considered when selecting a quantisation approach.

To clearly visualize the results reported above, Figure 3.11 presents BoP and NABS in a dual-axis grouped bar chart, allowing for a straight comparison of computational efficiency across different quantisation strategies. The results emphasize that while reducing bit precision leads to substantial computational savings, the choice of quantisation method significantly impacts how efficiently these savings are realised in practical implementations.

The analysis showed that uniform quantisation led to the highest complexity, while PoT quantisation offered the lowest. APoT, depending on the number of additive terms, achieved a balance between the two, with fewer terms leading to lower complexity but also slightly reduced performance. For instance, at 4-bit precision, the NABS using uniform quantisation was equivalent to that of APoT with 3 additive terms. This suggests that APoT with fewer additive terms is an optimal choice for applications requiring low bit precision and reduced computational complexity.

On the other hand, the heterogeneous PTQ approach allows for varying bit precisions and quantisation methods across different layers of the ANN. This flexibility enables the fine-tuning of the quantisation process to optimise performance in specific parts of the network, thereby achieving a better overall balance between complexity and performance.

The impact of heterogeneous quantisation is analysed by evaluating the Q-factor as a function of the bit precision assigned to the input layer, the recurrent kernel of the LSTM layer, and the filter kernel of the CNN layer, across the three transmission scenarios. This analysis considers three quantisation techniques: uniform quantisation, the original APoT method, and a mixed quantisation strategy that applies different quantisation techniques across layers. The results confirm that, across all scenarios, QAT significantly improves the optical performance over PTQ. For example,

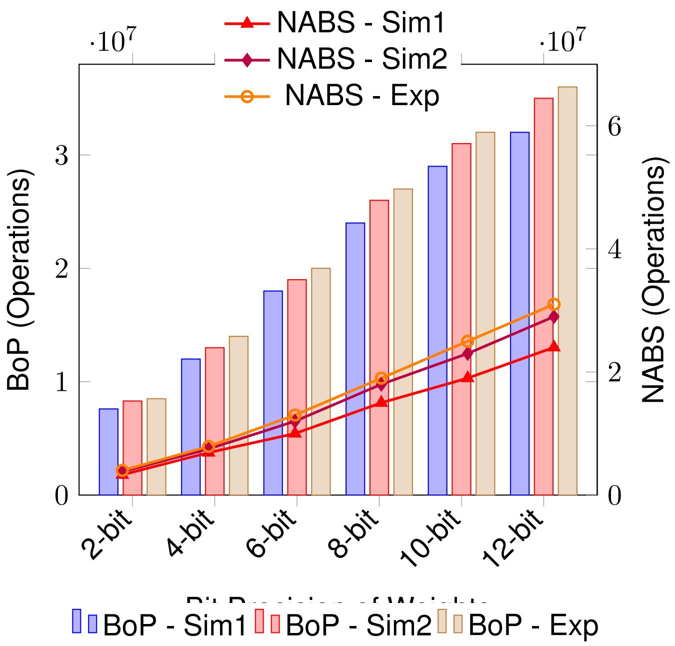


Figure 3.11: Comparison of computational complexity for different quantisation bit precisions. BoPs and NABS metrics are shown together for easier evaluation of computational trade-offs.

in Sim1, PTQ results in a Q-factor of only 0.18 dB, whereas QAT increases this value to 8.6 dB, reducing the performance gap with the non-quantised model. Similar trends are observed in Sim2 and Exp, where QAT consistently yields better results, demonstrating its effectiveness in mitigating quantisation-induced performance degradation.

Moreover, it has been demonstrated that the efficiency of heterogeneous quantisation in minimising bitwidth requirements while maintaining high optical performance. In Sim1, the APoT Original quantisation applied to the input and recurrent layers, combined with uniform quantisation for the CNN layer, enabled a substantial performance improvement. In Sim2, the use of APoT with 2 terms for the input layer and APoT Original for the recurrent and CNN layers resulted in a Q-factor of 4.5 dB under QAT, compared to 0 dB under PTQ. In Exp, a similar pattern is observed, with QAT yielding a 7.41 dB Q-factor while maintaining a reduced bitwidth configuration.

These findings confirm that different ANN layers exhibit varying levels of sensitivity to quantisation, and a fixed quantisation strategy does not necessarily yield the best performance across all scenarios. By applying a tailored approach where different layers are assigned bit precisions and quantisation methods that best suit their function, it is possible to achieve significant complexity reductions while maintaining a high Q-factor. The results emphasize the importance of adaptive quantisation techniques in optimising ANN-based equalisers for optical communication systems, highlighting the potential of heterogeneous quantisation to deliver high efficiency while preserving performance.

As in the PTQ approach described before (see Tables 3.4–3.6), here the focus remains on reducing both weights' and activations' precision to facilitate deployment of ANN-based equalisers in real-time, resource-constrained hardware. However, instead of quantising after training, Quantisation Aware Training (QAT) applies the quantisation during the training phase. By doing so, the training process accounts for the discretisation effects, often improving performance at lower bitwidths compared to PTQ.

This study explores QAT in both homogeneous and heterogeneous forms, just as we did with PTQ. Recall that the same BiLSTM + CNN baseline equaliser is used, following a preliminary clustering step illustrated earlier (Seq 3.11). The difference is that, in QAT, the network weights are quantised at each training iteration, causing the optimiser to adapt the model to the quantisation noise.

As with PTQ, we evaluate QAT across three scenarios:

- **Sim1:** 64-QAM, 30 Gbaud over 20 × 50 km SSMF,
- **Sim2:** 64-QAM, 34.4 Gbaud over 9 × 110 km SSMF,
- **Exp:** A real experimental setup with 64-QAM, 34.4 Gbaud over 9 × 110 km SSMF.

Tables 3.7–3.9 summarize the resulting Q-factor values when different quantisation strategies (Uniform, PoT, APoT variants) are employed in QAT under various bitwidths. Note that we restrict our tabulated data here primarily to the lower bitwidth regime (2–6 bits), since that region is most interesting for QAT.

Comparing these QAT results with the corresponding PTQ outcomes (Tables 3.4–3.6) reveals that QAT generally achieves substantially higher Q-factor, especially at 3–4 bits, where PTQ often collapses in performance. For instance, in Sim1 (Table 3.7), APoT with 2 or 4 additive terms preserves a Q-factor above 9 dB already at 4 bits, whereas PTQ in the same region dropped to around 3.8 dB (see Table 3.4). A similar trend emerges in Sim2 (Table 3.8), where QAT surpasses PTQ by about 1–2 dB at 3–4 bits.

The results in Table 3.9 (Exp scenario) confirm that QAT is most advantageous at lower bitwidths below 6 bits. At 2–3 bits, QAT still yields modestly reduced performance relative to the non-quantised reference, yet it avoids the near-complete degradation encountered with PTQ. The original APoT quantiser is often the best performer, presumably because the training process effectively learns to utilise its non-uniform levels. However, the gap between PoT and APoT is less pronounced under QAT than it was for PTQ, highlighting that both benefit from the error mitigation provided by quantisation-aware backpropagation.

Finally, we observe that the BO used to select the bitwidths and strategies per layer can further boost performance while preserving complexity gains. In Sim1, for example, one BO-chosen solution sets 3 bits for the input layer, 5 bits for the recurrent layer, and 4 bits for the CNN layer, all using APoT Original except for the CNN kernel with Uniform. This arrangement yields a Q-factor of about 8.6 dB, which—while lower than the 10.3 dB non-quantised baseline—still represents an excellent complexity-performance trade-off.

These findings underscore that QAT consistently surpasses PTQ at low bitwidths. Nonetheless, QAT does experience some performance drop relative to the non-quantised model, especially if the bit precision is pushed aggressively below 4 bits. Hence, further refinements (e.g., combining QAT with pruning, specialised quantisers, or advanced hyperparameter optimisation) may be necessary to achieve ultra-low bitwidth operation without unacceptable Q-factor loss.

Table 3.7: QAT (Homogeneous) results for Sim1, focusing on 2–6 bits. Compare with PTQ in Table 3.4.

Method	**2**	**3**	**4**	**5**	**6**
Uniform	5.20	6.39	7.90	9.46	10.10
PoT	5.20	6.39	7.60	7.65	7.79
APoT-2	5.20	6.39	8.33	9.24	9.24
APoT-3	5.20	6.39	7.90	9.21	9.79
APoT-4	5.20	6.39	7.90	9.46	10.10
APoT-Orig	5.20	6.39	8.52	9.24	9.84
NonQuantRef	10.31	10.31	10.31	10.31	10.31

Table 3.8: QAT (Homogeneous) results for Sim2, focusing on 2–6 bits. Compare with PTQ in Table 3.5.

Method	**2**	**3**	**4**	**5**	**6**
Uniform	3.34	3.656	3.795	4.07	4.71
PoT	3.29	3.656	4.45	4.40	4.36
APoT-2	3.36	3.656	4.41	4.73	4.84
APoT-3	3.31	3.656	3.873	4.27	4.844
APoT-4	3.34	3.656	3.795	4.07	4.84
APoT-Orig	3.217	3.656	4.32	4.75	4.94
NonQuantRef	5.08	5.08	5.08	5.08	5.08

Overall, these findings highlight the crucial role of BO in maximizing the efficiency and effectiveness of quantisation in neural networks. By intelligently navigating the hyperparameter space, BO ensures that QAT can deliver high-performance models with significantly reduced computational complexity, making these models more viable for real-time deployment in optical communication systems. The consistent superiority of QAT over PTQ across diverse scenarios underscores the importance of integrating adaptive and scenario-specific optimisation strategies to maintain high model accuracy while managing the trade-offs inherent in quantisation.

The heterogeneous QAT approach, while more complex, provides substantial benefits in terms of both performance and computational efficiency. The results demonstrate that by carefully selecting the bit precision and quantisation method for each layer, it is possible to achieve significant

Table 3.9: QAT (Homogeneous) results for Exp scenario, focusing on 2–6 bits. Compare with PTQ in Table 3.6.

Method	2	3	4	5	6
Uniform	7.12	7.56	7.68	7.85	7.88
PoT	7.10	7.52	7.65	7.77	7.80
APoT-2	7.15	7.58	7.75	7.89	8.00
APoT-3	7.12	7.56	7.66	7.83	7.92
APoT-4	7.12	7.56	7.68	7.85	7.91
APoT-Orig	7.15	7.58	7.78	7.97	8.10
NonQuantRef	8.10	8.10	8.10	8.10	8.10

complexity reductions without compromising the network's ability to mitigate nonlinear impairments in optical communication systems.

The quantisation study underscores the importance of choosing the appropriate quantisation strategy based on the specific requirements of the optical communication system. While homogeneous quantisation offers simplicity and ease of implementation, heterogeneous approaches provide the flexibility needed to optimise performance across different network layers. The integration of QAT, particularly when combined with Bayesian Optimisation, represents a significant advancement, enabling more aggressive quantisation with minimal performance loss. These findings suggest that QAT, especially in its heterogeneous form, is well-suited for developing efficient, high-performance NN equalisers that can operate effectively in real-time optical communication systems.

3.13 Deployment of ANNs-based Equalisers on Resource-Constrained Hardware

In this section, we continue our exploration of ANNs-based equalisers in optical communication systems, focusing on the deployment of these models on resource-restricted hardware. Building on previous sections, we evaluate the performance of compressed ANNs models that have undergone pruning and quantisation, assessing their ability to mitigate transmission impairments while operating within the constraints of hardware like the Raspberry Pi 4 and Nvidia Jetson Nano.

The datasets used for training and evaluating the ANN-based equalisers were generated through the numerical simulation of DP optical communication systems. Specifically, the datasets are composed of signals transmitted over SSMF links, each 50 km long, totaling a distance of 1000 km. The transmission signals employ a 64-QAM modulation format with a symbol rate of 30 GBd. A RRC filter with a 0.1 roll-off factor was applied to shape the signal, which was sampled at 8 samples per symbol. Like in previous cases. in order to ensure the dataset's diversity and to prevent overfitting, a pseudo-random bit sequence of order 32 was used for the simulation. The generated dataset consists of training and testing samples, each containing 2^{18} symbols, ensuring a sufficiently large sample size for robust training and evaluation processes.

In our configuration, the ANN is placed at the receiver side after the Integrated Coherent Receiver, ADC, and DSP block. This last block consists of a matched filter and a linear equaliser. Concerning the matched filter, it is just the same RRC filter used in the transmitter. Moreover, the linear equaliser is composed of a full electronic CDC stage and a normalisation step, see Figure 3.13. The CDC uses a frequency-domain equaliser and downsampling to the symbol rate, followed by a phase/amplitude normaliser to the transmitted ones. This normalisation process can be viewed as its normalisation by a constant K_{DSP} learned using the followed equation:

$$\mathcal{K}_{\mathsf{DSP}} = \min_{\mathcal{K}} \left\| \mathcal{K} \cdot x_{h/v}(z,t) - x_{h/v}(0,t) \right\| , \tag{3.66}$$

where constants K, $K_{\mathsf{DSP}} \in \mathbb{C}$, $x_{h/v}$ is the signal in either h or v polarisation. No other distortions – related to the components within the transceiver – were considered.
For this system, the best optimal power occurred at -1 dBm with the Q-factor being close to 7.8, as it can be appreciated in Figure 3.12. We then wanted to investigate the 3 next powers (e.g. 0 dBm, 1 dBm, and 2 dBm) going towards the higher nonlinear regime, where the task of the ANN would be more complicated.

Deploying ANN-based equalisers on edge devices is not a straight forward process, despite having features that meet the requirements needed for basic AI tasks, as their computing capabilities are significantly below those of high-performance servers or Graphics Processing Units (GPUs) [137]. To reduce the CC and memory requirements of the ANN in order to make its deployment in resource constrained hardware feasible, the model was subjected to pruning and quantisation, as illustrated in Figure 3.13.

The ANN employed in this study (shown in Figure 3.13) is a MLP architecture, optimised for equalisation tasks in optical communication systems, with the following key characteristics:

- **Architecture:** Three hidden layers with 500, 10, and 500 neurons respectively, creating a deep feedforward network capable of modelling complex nonlinear optical channel impairments.
- **Activation Function:** Hyperbolic tangent ($\tanh$) activation for all layers, selected by BO over the commonly used ReLU alternative. The $\tanh$ function proved superior for this application due to its symmetric, bounded output range $[-1, 1]$, which naturally aligns with the bipolar nature of I/Q constellation points in coherent optical systems.
- **Input Structure:** Tensor with dimensions [Batch Size B, Memory Depth $2M + 1$, Input Features d_{in}], where:
    - Memory depth: $M = 10$ symbols (BO-optimised), resulting in a 21-symbol temporal window ($2M + 1 = 21$) that captures inter-symbol interference (ISI) spanning multiple

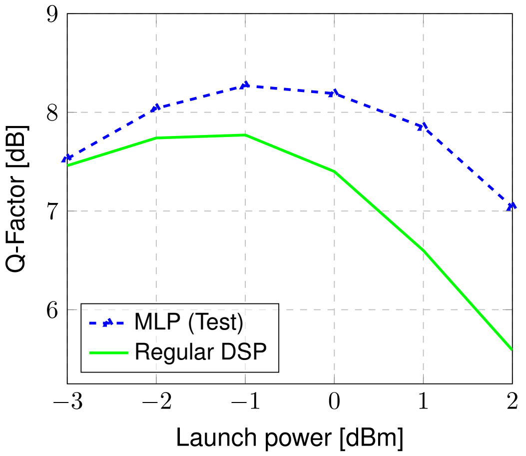


Figure 3.12: Performance comparison for the ANN-based equaliser with respect to the Regular DSP.

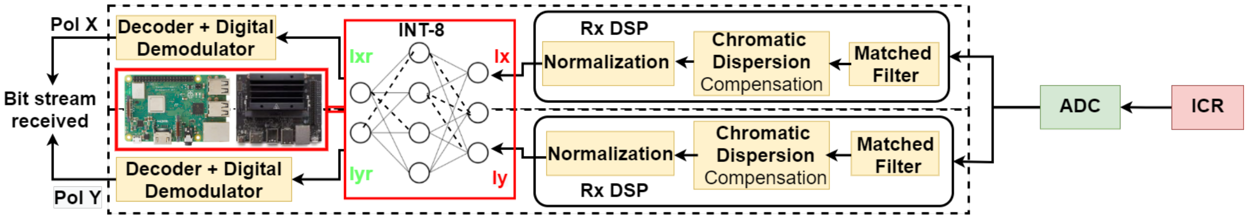


Figure 3.13: Structure of a communication channel that is equalised using a pruned and quantised ANN-based equaliser deployed on resource-restricted hardware (e.g. a Raspberry Pi 4 or a Nvidia Jetson Nano).

symbol periods.

  - Input features: $d_{in} = 4$ components per symbol for dual-polarisation systems (in-phase and quadrature for both H and V polarisations: I_H, Q_H, I_V, Q_V).

- **Output:** Real and imaginary parts of equalised symbols for one polarisation component at the centre of the temporal window, reconstructing the transmitted constellation points.

- **Total Parameters:** Approximately 510,510 trainable weights before compression:

  - Input $\rightarrow$ Hidden1: $(21 \times 4) \times 500 = 42,000$ weights
  - Hidden1 $\rightarrow$ Hidden2: $500 \times 10 = 5,000$ weights
  - Hidden2 $\rightarrow$ Hidden3: $10 \times 500 = 5,000$ weights
  - Hidden3 $\rightarrow$ Output: $500 \times 2 = 1,000$ weights (plus biases)

- **Compression Applied:** To enable deployment on resource-constrained edge devices:

  - **Pruning:** Magnitude-based weight removal at sparsity levels ranging from 20% to 90% (10% increments). Low-magnitude weights are set to zero and excluded from computation, reducing both memory footprint and computational load.

  - **Quantisation:** Precision reduction from 32-bit floating-point (FP32) to 8-bit integer (INT8) for both weights and activations. This quantisation achieves 75% memory reduction (4 bytes → 1 byte per weight) with minimal Q-factor degradation (< 0.3 dB at 60% sparsity + INT8).

- **Optimisation Methodology:** All hyperparameters (layer sizes, activation function, memory depth, learning rate, pruning schedule, quantisation bit-width) were jointly optimised using BO to minimise the objective function balancing Q-factor and computational complexity (RMpS).

This MLP equaliser configuration represents a careful balance between modelling capacity (deep architecture with 510K parameters) and deployment feasibility (aggressive compression to 204KB INT8 model at 60% sparsity), enabling real-time nonlinear equalisation on edge devices like Raspberry Pi 4 and NVIDIA Jetson Nano.

The final pruned and quantised models were deployed on two types of edge-computing hardware: Raspberry Pi 4 and Nvidia Jetson Nano. These devices were selected for their popularity and relevance in low-power, real-time applications.

The Raspberry Pi 4 is a versatile, cost-effective single-board computer that has become a staple in edge computing. It features a Broadcom BCM2711 system-on-chip (SoC) with a quad-core Cortex-A72 (ARM v8) 64-bit processor, running at 1.5 GHz. The board is available with up to 4 GB of LPDDR4-3200 SDRAM, providing adequate memory for lightweight neural network models. Despite its modest computational power, the Raspberry Pi 4 is highly energy-efficient, consuming approximately 3.4 W under typical load conditions. This energy efficiency is crucial for edge deployment, where power consumption is a primary concern [138, 139].

The Nvidia Jetson Nano is a more powerful edge computing platform designed specifically for AI and deep learning applications. It is equipped with a quad-core ARM Cortex-A57 MPCore processor and a 128-core Maxwell GPU, providing significantly greater computational capabilities compared to the Raspberry Pi 4. The Jetson Nano supports up to 4 GB of LPDDR4 memory and is optimised for parallel processing tasks, making it well-suited for handling more complex neural network models. Despite its enhanced performance, the Jetson Nano maintains a low power profile, typically consuming around 10 W, which is manageable in most edge computing scenarios [138, 140].

3.13.1 Computational Platform Bandwidth and Real-Time Processing Feasibility Analysis

A critical consideration when deploying ANN-based equalisers on edge computing platforms is the compatibility between the computational bandwidth (symbol processing throughput) of the hardware and the signal bandwidth requirements of modern optical communication systems. This subsection analyses whether resource-constrained devices like Raspberry Pi 4 and Nvidia Jetson

Nano can meet the real-time processing demands of high-speed optical channels, and identifies the bottlenecks that limit deployment scalability.

The optical transmission scenarios evaluated in this chapter operate at symbol rates ranging from 30 to 34.4 Gbaud (billion symbols per second). For dual-polarisation systems employing PDM, each polarisation must be equalised independently or jointly, effectively doubling the processing requirements. The key bandwidth parameters are:

- **30 Gbaud PDM-64QAM scenario** (20×50km SSMF): Requires processing 30 billion symbols per second per polarisation, or 60 Gsymbols/s total for both H and V polarisations. With our MLP equaliser processing symbols individually (batch size = 1 for minimum latency), the hardware must complete one inference in $\tau_{\text{req}} = 1/(30 \times 10^9) \approx 33.3$ picoseconds per symbol per polarisation.
- **34.4 Gbaud PDM-64QAM scenario** (9×110km SSMF): Requires 34.4 Gsymbols/s per polarisation (68.8 Gsymbols/s total), corresponding to $\tau_{\text{req}} = 1/(34.4 \times 10^9) \approx 29.1$ picoseconds per symbol per polarisation.

These picosecond-scale latency requirements represent the absolute minimum inference time needed for real-time equalisation without buffering. In practice, digital signal processing (DSP) chains tolerate finite buffering depths (typically 10^2-10^4 symbols), relaxing the instantaneous latency requirement while maintaining overall throughput constraints.

The measured inference times from our deployment experiments (Section 3.6) reveal the actual symbol processing bandwidth of each platform:

Raspberry Pi 4 (Cortex-A72 @ 1.5 GHz):

- **Uncompressed FP32 MLP (510K parameters):** Inference time $\tau_{\text{inf}} = 4.5\ \mu$s per symbol, yielding throughput of $1/(4.5 \times 10^{-6}) \approx 222$ Ksymbols/s per polarisation.
- **Compressed INT8 MLP (60% sparsity, 204K parameters):** Inference time $\tau_{\text{inf}} = 1.8\ \mu$s per symbol, yielding throughput of ≈ 556 Ksymbols/s per polarisation.
- **Real-time feasibility:** The Raspberry Pi 4 falls short of real-time processing by a factor of $\approx 54,000\times$ for the 30 Gbaud scenario (30 Gsymbols/s required vs. 0.556 Msymbols/s achieved with compression).

Nvidia Jetson Nano (Cortex-A57 @ 1.43 GHz + Maxwell GPU):

- **Uncompressed FP32 MLP:** Inference time $\tau_{\text{inf}} = 2.8\ \mu$s per symbol, yielding throughput of ≈ 357 Ksymbols/s per polarisation.
- **Compressed INT8 MLP:** Inference time $\tau_{\text{inf}} = 1.22\ \mu$s per symbol, yielding throughput of ≈ 820 Ksymbols/s per polarisation.

- **Real-time feasibility:** The Jetson Nano achieves $\approx 36,000\times$ lower throughput than required for 30 Gbaud real-time processing.

PC Platform (Intel i9-10870H @ 2.2-5.0 GHz + RTX 2070 GPU):

- **Compressed INT8 MLP:** Inference time $\tau_{\text{inf}} = 0.18$ μs per symbol, yielding throughput of ≈ 5.56 Msymbols/s per polarisation.
- **Real-time feasibility:** Even high-performance PC platforms fall short by $\approx 5,400\times$ for 30 Gbaud processing.

The significant gap between required and achieved throughput stems from multiple architectural bottlenecks:

- Memory Bandwidth Constraints: The Raspberry Pi 4 and Jetson Nano both employ LPDDR4 memory with theoretical peak bandwidth of 25.6 GB/s (LPDDR4-3200). However, effective bandwidth available to the CPU/GPU cores is significantly lower due to memory controller overhead, cache coherency traffic, and operating system background processes. Our INT8 MLP model requires loading ≈ 204 KB of weights per inference (at 60% sparsity), plus ≈ 2 KB of input activations and intermediate layer activations. At 1.22 μs per inference (Jetson Nano), the sustained memory bandwidth is ≈ 206 KB / 1.22 μs ≈ 169 MB/s—merely 0.66% of peak theoretical bandwidth. This low utilisation indicates that computation (arithmetic operations) rather than memory bandwidth is the primary bottleneck for these small models on edge platforms.
- Arithmetic Throughput Limitations: The Raspberry Pi 4's Cortex-A72 cores provide approximately 32 GFLOPS peak performance (at 1.5 GHz, with NEON SIMD units). Processing one symbol through our 204K-parameter INT8 model requires $\approx 14,400$ real multiplications per symbol (RMpS, at 60% sparsity). With INT8 quantisation, these become 8-bit integer operations, theoretically faster than FP32. However, TensorFlow Lite's INT8 kernel implementation on ARM achieves only ≈ 8 INT8 operations per CPU cycle (accounting for memory access overhead and pipeline stalls), yielding effective throughput of $(1.5 \times 10^9$ Hz$) \times 4$ cores $\times\, 8$ ops/cycle $= 48$ GOPS. At 14,400 operations per symbol, theoretical maximum throughput is $48 \times 10^9 / 14,400 \approx 3.33$ Msymbols/s—still $9,000\times$ below real-time requirements for 30 Gbaud.
- Software Framework Overhead: TensorFlow Lite, used for deployment on edge devices, introduces significant per-inference overhead from Python interpreter dispatch, tensor memory allocation, and operator scheduling. Our measurements indicate that ≈ 30–40% of total inference time is consumed by framework overhead rather than arithmetic operations. Optimised C++ implementations or hardware-specific FPGA/ASIC accelerators could reduce this overhead substantially.

Despite the apparent infeasibility of real-time processing on edge platforms, several deployment strategies enable practical use of ANN-based equalisers:

- Offline Processing for Stored Signals: For applications involving captured optical signals stored for post-processing (e.g., scientific experiments, link diagnostics, or archival data analysis), real-time constraints do not apply. Edge devices can process stored signals at their native throughput (≈ 0.5–1 Msymbols/s), completing equalisation of a 2^{18}-symbol dataset in ≈ 0.3–0.5 seconds—acceptable for many use cases.
- Hybrid DSP Architectures: A practical deployment paradigm involves offloading only the most computationally intensive nonlinear equalisation tasks to ANN accelerators, while performing linear equalisation (CDC, matched filtering) in real-time using dedicated FPGA or ASIC hardware. The ANN-based equaliser operates on buffered symbols with relaxed latency tolerance (≈ 10–100 μs per symbol), allowing edge platforms to handle moderate-bandwidth optical channels (≈ 100 Mbaud with appropriate buffering).
- Parallelisation Across Symbols: Modern optical receivers employ frame-based processing where blocks of N symbols (e.g., $N = 1024$) are processed jointly. By parallelizing the ANN equaliser across multiple CPU cores or GPU threads, effective throughput scales linearly with available parallelism. The Jetson Nano's 128-core GPU, when fully utilised with optimised CUDA kernels, could theoretically achieve $\approx 100\times$ speedup over single-core CPU inference, reaching ≈ 82 Msymbols/s—reducing the real-time gap to $\approx 360\times$ for 30 Gbaud systems.
- Model Compression Scaling: Our experiments demonstrate that 60% sparsity maintains Q-factor degradation < 0.5 dB. More aggressive compression to 80–90% sparsity (reducing RMpS from 14,400 to $\approx 1,440$–3,600) could yield another $4\times$–$10\times$ speedup, albeit with Q-factor degradation of 1–2 dB. For optical systems with sufficient SNR margin, this trade-off may be acceptable.

The bandwidth analysis reveals fundamental scaling challenges as optical systems migrate to higher symbol rates:

- 400G/800G systems (operating at ≈ 60-120 Gbaud): Real-time ANN equalisation on edge platforms becomes increasingly impractical without dedicated hardware acceleration. FPGA or ASIC implementations of pruned, quantised ANN models represent the most viable path, offering $100\times$-$1000\times$ speedups through custom datapath optimisation and massive parallelism.
- Low-bandwidth coherent systems ($<$ 10 Gbaud): Emerging applications such as data centre interconnects, metro optical networks, and passive optical networks (PON) increasingly employ coherent detection at lower symbol rates. For these scenarios, edge platforms with

compressed ANN equalisers can approach real-time performance, especially with GPU acceleration and optimised software stacks.

- Inference-time vs. training-time bandwidth: While real-time inference remains challenging, edge platforms are entirely unsuitable for training ANN-based equalisers due to the $10^6\times$ greater computational demands of gradient-based optimisation. Training must occur offline on GPU servers, with compressed models subsequently deployed to edge devices—a cloud-to-edge workflow common in machine learning deployment.

This analysis clarifies the distinct roles of different computational platforms: high-performance servers and GPUs for model training and development; edge devices like Raspberry Pi 4 and Jetson Nano for demonstration, prototyping, and low-throughput deployment scenarios; and dedicated hardware accelerators (FPGAs, ASICs) for production-grade real-time equalisation in commercial optical systems.

3.14 Performance, Inference Time and Energy Consumption Evaluation of ANNs-based Equaliser Deployed on Resources-constrained Hardware

As it has been discussed during this Chapter, the complexity reduction of the equaliser must not affect its performance drastically, which is required to be within an acceptable range. In Figure 3.14a, the Q-factor achieved by the ANN-based equaliser is depicted versus different sparsity values, for the three launch power levels considered: 0 dBm, blue; 1 dBm, red; and 2 dBm, green. The results are shown using dotted lines and stars, which are those obtained on a PC (MSI GP76 Leopard personal computer, equipped with Intel® Core$^{\text{TM}}$ i9-10870H processor, 32GB of Random Access Memory (RAM) and GPU Nvidia RTX2070), Raspberry Pi, and Nvidia Jetson Nano, using the pruned and quantised model. For each of these launch powers, two baselines for the Q-factor are depicted: one corresponds to the level achieved by the original uncompressed model, defined by the straight lines, while the other provides the benchmark when we do not employ any ANN-based equalisation and use only standard CDC plus phase/amplitude normalisation (LE, linear equalisation); the latter levels for the three different launch powers are marked by dotted lines having the appropriate colours.

Figure 3.14b quantifies the impact that each compression technique has on the performance: in that figure, we plotted the Q-factor achieved by the ANN-based equaliser versus different values of sparsity, for the 1 dBm launch power. The blue and red straight lines represent the Q-factor of the original model and the Q-factor achieved by it after being quantised. The dotted lines with asterisks, show the performance of a model that has been only pruned (blue), and the performance in the case of both pruning and quantisation (red). It is seen that a substantial reduction of the complexity can be achieved without a dramatic degradation of the performance. Furthermore, it can be appreciated the threshold for those sparsity levels at which the larger deterioration of the performance occurs.

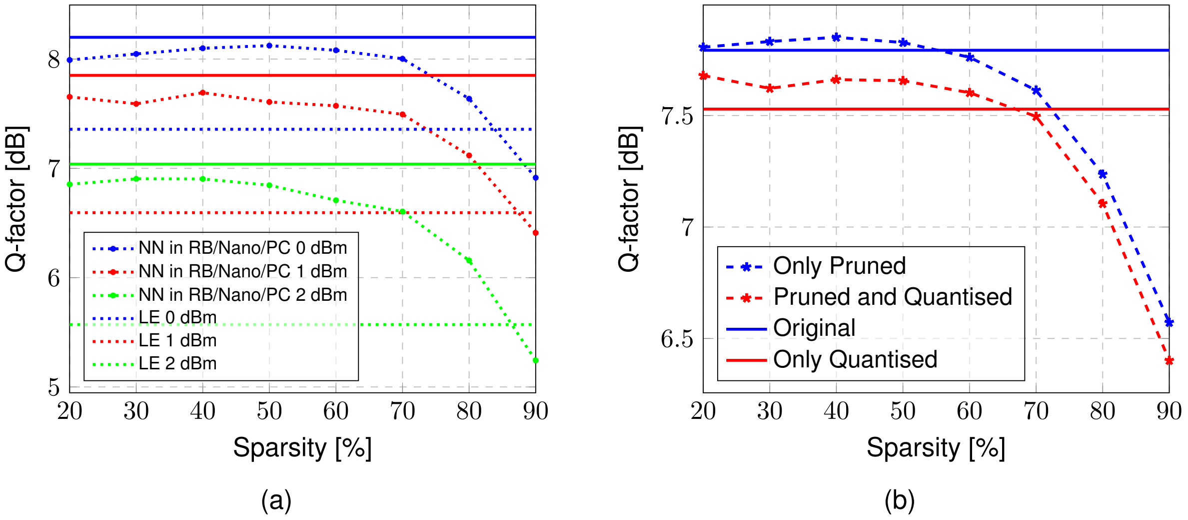


Figure 3.14: a) Q-factor achieved for pruned and quantised models versus the level of sparsity for datasets corresponding to three launch powers: 0 dBm, 1 dBm, and 2 dBm; The solid lines correspond to the Q-factor achieved by the original model. The dashed lines show the Q-factor when only linear equalisation (LE) is implemented. b) Q-factor achieved after pruning compared to the one achieved after both pruning and quantisation, for different levels of sparsity and for a dataset corresponding to the 1 dBm launch power. The blue and red solid lines correspond to the Q-factor achieved by the original model and the one achieved by this model after quantisation, respectively.

First, it can be observed from Figure 3.14a that the quantisation and pruning process does not cause a significant performance degradation until a sparsity level equal to 60% is reached, with just a 4% performance reduction. However, when larger sparsity levels are explored, e.g. 90%, the performance is close to the one achieved using a LE (i.e., the Q-factor curves drop to the levels marked with the dashed lines of the same colour).

We can conclude that when the levels of sparsity are above 60%, the decrease in the performance is mainly the effect of the quantisation process. A nearly 2.5% drop in the Q-factor value has also been observed when quantising an already pruned model. Once the levels of sparsity are higher than 60%, the reduction in performance due to the quantisation gets accelerated. Moreover, we observe that some degree of sparsification can even improve the model's performance with respect to the unpruned model. This behaviour has already been reported in other studies and it was found that it is specifically pertinent to the over-parametrised models, as in this case pruning acts also as a regularisation technique. Thus, the ANNs with less complex structures do not show up such an increase in performance due to low-sparsity pruning, making it impossible to achieve such a good performance-complexity ratios [116, 141].

Inference time is the main metric when assessing the ability of an equaliser to process incoming signals in real-time. The inference times for the original and pruned as well as quantised models were measured on both the Raspberry Pi 4 and Nvidia Jetson Nano platforms. When measuring the inference time for the different types of hardware and the quantised model that has had 60% of its weights pruned, the results are presented in Table 3.10

Figure 3.15 shows the latency of the considered NN model before and after quantisation. We notice that the results are expressed in a way that is more appropriate for the task at hand. Thus, latency is defined as the time it takes to process one symbol: we have averaged it over 30k symbols. With the quantised model, we observe approximately a 56% reduction in latency for all three values of power, when compared to the original model. We must notice that pruning is not taken into account because it does not affect this metric since Tensorflow Lite does not support sparse inference yet, which makes the algorithm still use the same amount of cache memory. Also, we could observe that Raspberry Pi has the longest inference time among our devices. This is in line with the fact that Raspberry is designed as a low-cost and general-purpose single-board computer [138]. On the other hand, the Nvidia Jetson Nano was developed with GPU capabilities, which makes it more suitable for deep learning applications, allowing us to achieve lower latencies.

To evaluate the inference time for each model, no peripherals are connected either to the Raspberry Pi or to the Nvidia Jetson Nano, except the Ethernet port, which is used to establish communication over the Secure Shell protocol. Moreover, any initialisation time (e.g., library loading, data generation, and model weight loading) is ignored because this is a one-time cost that occurs during the device's setup. Additionally, 25 measures were taken for each device. In each of them, 100 inferences were run (in each inference 30k symbols are recovered) and the inference time was averaged, not taking into account the initialisation phase.

Power consumption is another important metric in ANNs-based solutions, as it has become a critical concern in the context of global energy scarcity and environmental sustainability. As ANNs grow in size and complexity, their associated energy consumption have risen dramatically, contributing significantly to the carbon footprint of the information technology sector. This surge in energy usage not only strains existing power infrastructures but also exacerbates climate change through increased greenhouse gas emissions. Consequently, there is an urgent need to develop energy-efficient ANN architectures and hardware implementations that can deliver high performance while minimising power consumption. Addressing this challenge is crucial for ensuring the continued advancement of AI technologies in a manner that aligns with global efforts to combat climate change and promote sustainable development [142–144]. In this Chapter we address this issue in the context of optical communication systems.

To quantify the energy consumption of the ANN-based equalisers, we employed a digital USB multimeter (detailed setup shown in Figure 3.16) connected in series with the power supply of each edge device. The measurement protocol mirrors that used for inference time evaluation: 25 independent measurements were conducted for each device (Raspberry Pi 4 and Nvidia Jetson

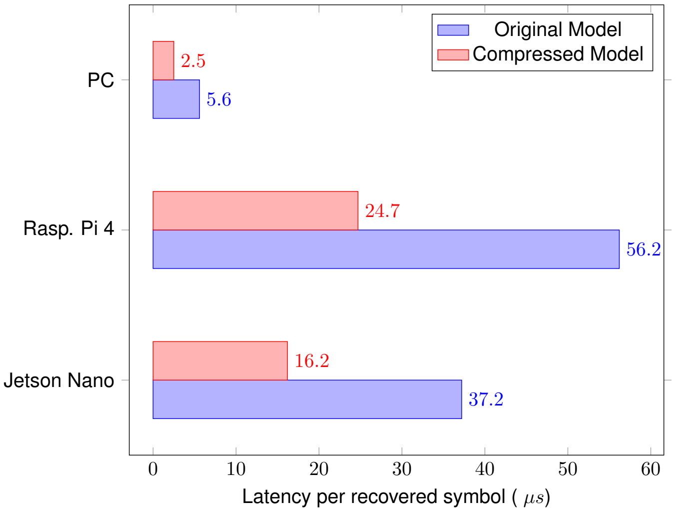


Figure 3.15: Summary of the symbol processing (inference) time for compressed NN models (after pruning and quantisation) and the original models for three devices under evaluation: Raspberry Pi 4, Jetson Nano, and a standard PC. Model Complexity Reduction Quantification: The original MLP equaliser (500-10-500 architecture) contains 510,510 trainable parameters occupying 2.04 MB in FP32 format, with a baseline computational complexity of 36,000 RMpS. After applying 60% magnitude-based pruning combined with FP32→INT8 quantisation, the compressed model achieves: (i) Memory reduction: 510K → 204K active parameters (60% reduction), 2.04 MB → 204 KB storage (90% reduction from combined pruning and 4× quantisation), (ii) Computational complexity: 36,000 → 14,400 RMpS (60% reduction proportional to sparsity level), (iii) Inference speedup: Raspberry Pi 4 achieves 2.5× faster inference (from 4.5 μs to 1.8 μs per symbol), Jetson Nano achieves 2.3× speedup (from 2.8 μs to 1.22 μs), and PC achieves 2.1× speedup (from 0.38 μs to 0.18 μs). Latency measurements were obtained by averageing inference time over 30,000 symbols in TensorFlow Lite runtime. Note that while pruning reduces theoretical complexity (RMpS), the latency improvements shown here primarily result from INT8 quantisation, as TensorFlow Lite does not yet support sparse inference optimisation (pruned weights still occupy cache memory). The quantified reductions demonstrate that aggressive compression enables deployment on resource-constrained edge devices while maintaining Q-factor degradation below 0.3 dB for the 60% sparsity + INT8 configuration.

Table 3.10: Inference time comparison between compressed and original models across devices.

Device	**Compressed Model ($\mu \pm \sigma$ [s])**	**Original Model ($\mu \pm \sigma$ [s])**
Raspberry Pi 4	0.81 ± 0.035	1.84 ± 0.08
Nvidia Jetson Nano	0.53 ± 0.022	1.22 ± 0.052
PC	0.10 ± 0.006	0.18 ± 0.008

Nano), with each measurement consisting of 100 inference runs processing 30,000 symbols per run. The multimeter recorded instantaneous power consumption (in Watts) at 1 kHz sampling rate throughout each inference run, enabling precise energy integration over time.

To isolate the ANN-specific power consumption from the total system baseline, we performed the following procedure: (i) measured idle power consumption with only the operating system and SSH connection active (no ANN inference), (ii) measured total power consumption during ANN inference execution, and (iii) calculated the ANN-specific power as the difference between active and idle states. The initialisation phase (library loading, model weight loading from storage, input data generation) was excluded from energy measurements, as it represents a one-time overhead amortised over millions of symbols in continuous operation.

Energy per symbol (expressed in nanojoules per symbol, nJ/Symb) was calculated as: $E_{\text{symb}} = \frac{\int_{t_0}^{t_f} [P_{\text{active}}(t) - P_{\text{idle}}] \, dt}{N_{\text{symbols}}}$, where $P_{\text{active}}(t)$ is the instantaneous power during inference, P_{idle} is the baseline idle power, and $N_{\text{symbols}} = 30,000$ is the number of equalised symbols per inference run. All 25 measurements were averaged to obtain the final energy consumption values reported below. The percentages reported (e.g., "70.3% of total energy consumption") represent the fraction of total device power consumption (including operating system, peripherals, and all background processes) attributed to the ANN inference task during active symbol processing. The results can be found below:

- Nvidia Jetson Nano:
  - Original Model Energy Consumption: 1.16 nJ per Symb, representing 70.3% of total energy consumption.
  - Compressed Model Energy Consumption: 0.49 nJ per Symb, representing 29.7% of total energy consumption.
- Raspberry Pi 4:
  - Original Model Energy Consumption: 1.72 nJ per Symb, representing 70% of total energy consumption.
  - Compressed Model Energy Consumption: 0.74 nJ per Symb, representing 30% of total energy consumption.

These results demonstrate the significant impact of model compression techniques on energy consumption in ANN-based optical communication systems. For the Nvidia Jetson Nano, the compressed model achieves a remarkable 57.8% reduction in energy consumption compared to the original model. Similarly, on the Raspberry Pi 4, the compressed model realizes a 57.0% decrease in energy usage. The consistency in energy savings is particularly important for the widespread adoption of energy-efficient ANN solutions in optical communications, as it indicates that the benefits are not limited to specific hardware configurations. This reduction in energy consumption not only addresses the immediate concerns of power usage in AI applications but also contributes to the broader goals of sustainability in the IT sector.

The set-up employed to measure the energy consumption of the proposed solutions can be found in Figure 3.16.

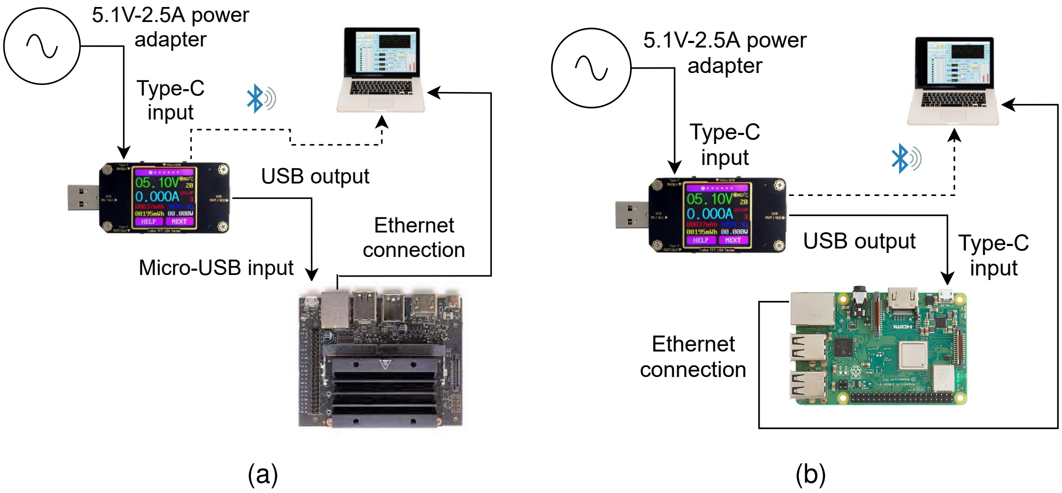


Figure 3.16: (a) The power measurement set-up for Nivida Jetson Nano, and (b) – the same for Raspberry Pi.

3.15 ANNs-based Equaliser Computational Complexity and Memory Size Analysis

Figure 3.17 depicts the reduction in the size of the model as well as the model's CC for different sparsity values, after having applied quantisation.

Overall, we have achieved an 87.12% reduction in the memory size after pruning 60% of the ANN-based equaliser weights and quantising the remaining ones. As a consequence, the size of the model went down from 201.4 to 25.9 kilobytes. For the decrease of the model's computational complexity, it goes from 75960427.38 to 16447962 BoPs after applying the same compression strategy, which is a 78.34% reduction. We would like to point out once more that sparsity levels of 60% can be reached without a substantial performance loss. Therefore, approximately the same high level of performance can be achieved with a model that is significantly less complex than the initial ANN structure, which is enables its deployment on resource-restricted hardware.

It is worth mentioning the individual impact that quantisation and pruning have on the computational complexity of the model. When the computational complexity is calculated for a quantised, but unpruned model, the number of BOPs is equal to 23321563. Therefore, if this value is compared with the already mentioned 75960427 BoPs for the unpruned and unquantised NN, a reduction in complexity of a 69.3% is obtained thanks to quantisation. As it can be seen in Figure 3.17, the remaining gain comes from the pruning technique, and it grows linearly as indicated in Equation (3.69).

In this work, the size of the model is defined as the number of bytes that it occupies in memory. Moreover, we notice the direct correlation between the value of this metric and the format used to represent the model. Thus, in contrast to the traditional formats used in Tensorflow (e.g .h5 or HDF5 binary data format and .pb or protobuf), a TensorFlow Lite model, it is represented in a

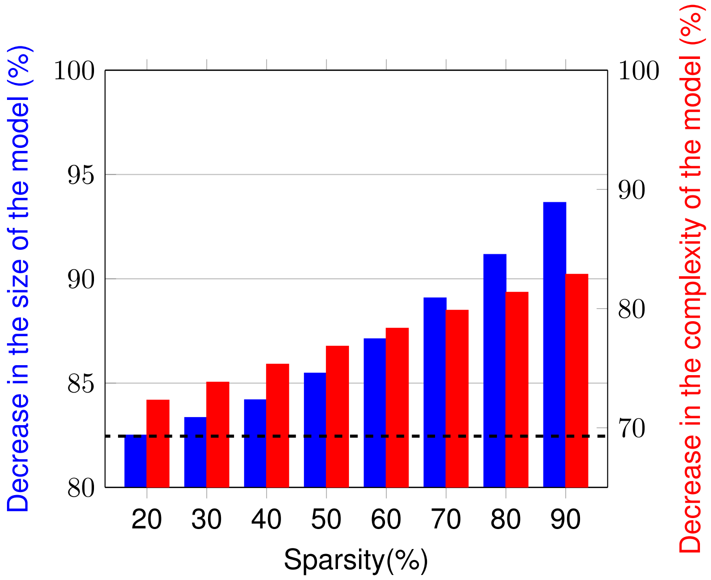


Figure 3.17: Complexity and size reduction achieved via pruning and quantisation for different levels of sparsity. The dashed black line represents the reference complexity when only quantisation is applied.

special efficient portable format identified by the `.tflite` file extension. This provides two main advantages: a reduced model size and lower inference times. Therefore, the deployment of the NN model on a resource-restricted hardware becomes feasible. As a consequence, it would not make sense to compare the models saved in the traditional Tensorflow format with those that have been pruned and quantised as well as converted into Tensorflow Lite. We were aware of this situation during the realisation of the procedure and, thus, to avoid overestimating the benefits of pruning and quantisation, the unpruned and unquantised model has also been converted to `.tflite` format. To better understand the implications that this step has, the size of the original model in `.h5` format would experiment with a 96.22% size reduction after being converted to `.tflite` format, quantised and pruned (60% sparsity). On the other hand, if the original model has already been converted to `.tflite`, the size reduction is 87.12%. Of course, based on this, always using `.tflite` format instead of the other conventional ones seems to be the best strategy. The main reason behind not doing this is that a graph that is in `.tflite` format can not be trained again, so, it only supports an online inference mode. Nevertheless, a model that is, for example, in `.h5` format, can be trained offline. Therefore, the `.tflite` is only intended to be used in the context of edge computing.

Regarding the calculation of the BoPs for the MLP under test, it is important to take into account that as a MAC operation takes the form: $a = a + w \cdot x$, it is possible to distinguish two contributions, one corresponding to the $nk^2 \cdot b_0$ number of additions, where $b_0 = b_a + b_w + \log_2(nk^2)$ (e.g. accumulator width in the MAC operations), and the other corresponds to the number of multiplications, e.g. $nk^2(b_a b_w)$. Based on this, for the case of an MLP [145, 146]:

$$\mathrm{BoPs}_i = m_i n_i[(1 - f_{p_i}) b_{a_i} b_{w_i} + b_{a_i} + b_{w_i} + \log_2(n_i)], \tag{3.67}$$

In Equation (3.67), n and m correspond to the number of inputs and outputs, respectively; b_w and b_a are the bit widths of the weights and activations. The additional term, f_{p_i}, is the fraction of pruned layer weights, which allows us to take into account the reduction in multiplication operations because of pruning. This is the reason why it only relates to the term $b_a b_w$ [146].

Therefore, in our case of the MLP with 3 hidden layers, the total number of BOPs is:

$$\mathrm{BoPs} = \text{BoPs}_{input} + \sum_i \text{BoPs}_i + \text{BoPs}_{output}, \tag{3.68}$$

where $i \in [1, 2, 3]$, $\mathrm{BoPs}_{\mathrm{input}}$ and $\mathrm{BoPs}_{\mathrm{output}}$ correspond to the contributions of the input and output layers. Equation (3.68) can be written in a less compact way as follows:

$$\begin{aligned}\mathrm{BoPs}_{\mathrm{MLP}} = &(n_i n_1 b_i + n_1 n_2 b_a + n_2 n_3 b_a + n_3 n_o b_a)(1 - f_p) b_w + (n_i n_1)(b_i + b_w) \log_2(n_i) \\ &+ (n_1 n_2)(b_a + b_w) \log_2(n_1) + (n_2 n_3)(b_a + b_w) \log_2(n_2) + (n_3 n_o)(b_a + b_w) \log_2(n_3),\end{aligned} \tag{3.69}$$

where n_i, n_1, n_2, n_3 ,and n_o are the number of neurons in the input, first, second, third, and output layers, respectively; b_w, b_a, b_o and b_i are the bit width of the weights, activations, output, and input, respectively; f_p is the fraction of the weights that have been pruned in a layer, which, in our case, is the same for every layer.

3.16 Low Complexity RNN-based Equalisers in Optical Channel Communication Systems employing SOAs

As discussed in Chapter 1, SOAs have numerous attractive characteristics, such as compact size, efficient electrical pumping, cost-effectiveness, and wideband gain [147–149]. However, the nonlinear properties of SOAs, relatively slow gain recovery time, and comparatively high noise figure (relative to other optical amplifiers) present significant challenges for high-speed optical communication applications. Despite these challenges, SOAs offer substantial potential in Metro/Access networks where cost and component integration are critical.

The carrier dynamics of SOAs operate on a timescale of hundreds of picoseconds, which makes the gain dependent on the input optical power in real-time. This leads to nonlinear impairments, including inter-symbol interference, particularly in systems employing advanced modulation formats with numerous constellation points. Therefore, mitigation of these nonlinear distortions is a major challenge in making SOAs viable for telecommunication applications compared to traditional amplification methods like EDFAs.

In high-speed coherent detection systems, DBP has been utilised to compensate for nonlinear distortions introduced by SOAs by inverting a numerical model of the amplifier at the receiver [150]. However, DBP for SOA introduces additional computational complexity due to the need for a specific model of the SOA, creating an opportunity to explore alternative digital equalisation methods.

Recently, ANNs have shown promise as DSP techniques for compensating nonlinear transmission impairments in optical communication systems [151–153]. The advantage of ANNs lies in their ability to mitigate complex nonlinear signal distortions and signal-noise interactions, especially in scenarios with pronounced nonlinear impairments such as those imposed by SOAs. Thus, ANNs have significant potential for equalising SOA-induced distortions in coherent detection systems. The practical implementation of ANN-based equalisers involves balancing performance and CC, a key factor for real-time processing feasibility. While ANNs have previously been applied to intensity modulation/direct detection formats, the unique challenges of coherent detection demand a distinct approach. This work focuses on developing ANN-based equalisers for coherent detection, comparing their performance and CC against DBP in a WDM 11 x 112 Gb/s 16-QAM transmission over a SSMF link of 120 km.

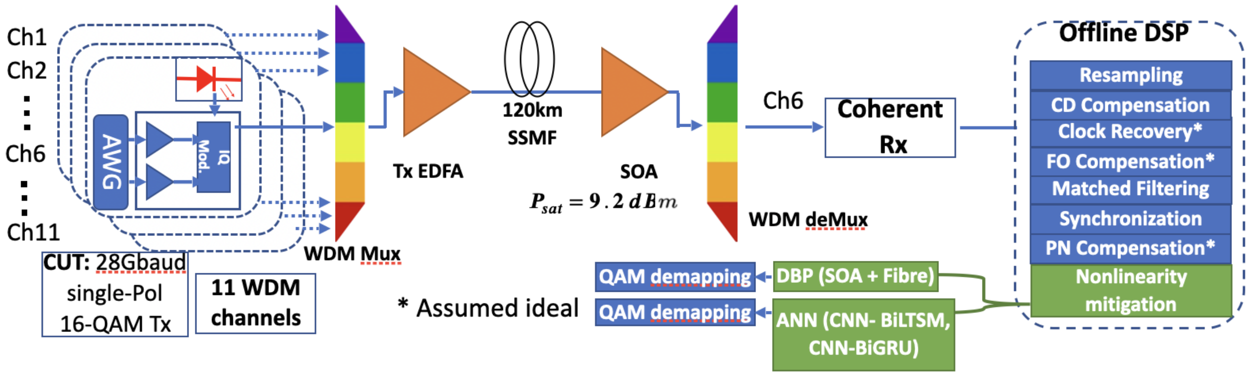


Figure 3.18: Schematics of the ANN-based equalisers in the transmission link of 120 km with SOA, Channel 6 is the channel under test. CDC is only used with the ANN-based equaliser.

The simulated transmission line is schematically depicted in Figure 3.18. A 16-QAM, root-raise-cosine waveform, single polarisation signal was transmitted over 11 WDM channels at a 28 Gbaud channel rate. Signal propagation in the 120-km SSMF span was modelled using the conventional split-step method. Direct bit-error-rate counting and the corresponding recalculated Q-factor are used for the central channel (number 6), which is considered without loss of generality to evaluate the performance of the communication system. The Cassioli-Scotti-Mecozzi model (see for details [154]), incorporating distributed loss and amplified spontaneous emission generated within the SOA, is employed here to simulate the amplification of the signal and associated nonlinear distortions. This model offers an accurate analysis of the signal evolution within an SOA.

The parameters of the amplifier used in simulating the transmission system (Cassioli-Scotti-Mecozzi model) are: $\alpha_H = 5$, $\tau_c = 170$ ps, $P_{sat} = 8.32$ mW (9.2 dBm), SOA loss is 9 dB, and we used an excessively high noise figure of $NF = 10$ to operate in the worst-case noise scenario. However, due to its high CC, it is impractical to implement this model in a contemporary DSP. Therefore, instead, a simplified model (see Eqs. (3.70-3.71) below) that includes key physical effects is utilised here to perform DBP at the receiver for comparison with the proposed machine learning-based solutions. The signal transformation ($E_{in}(t) \rightarrow E_{out}(t)$) in the SOA is given by:

$$E_{out}(t) = \exp\left[\frac{1 - i\alpha_H}{2}\, h(t)\right] E_{in}(t), \tag{3.70}$$

here α_H is the amplifier linewidth enhancement Henry factor, and the signal phase is modulated by the time-dependent gain function $h(t)$ according to the equation:

$$\tau_c \frac{d}{dt} h(t) = h_0 - h(t) - \left[e^{h(t)} - 1\right] \frac{P_{in}(t)}{P_{sat}}, \tag{3.71}$$

where h_0 is a linear gain ($G_0 = \exp(h_0)$ to compensate for the SMF span loss, $P_{in}(t) = |E_{in}(t)|^2$ is the signal power at the SOA input, τ_c is the SOA carrier lifetime, and $P_{sat} = E_{sat}/\tau_c$ is the SOA saturation power (E_{sat} is the corresponding SOA saturation energy).

The inverse SOA in DBP is implemented through a gain reversal operation. Parameters for modelling the SOA, using Eqs. (3.70-3.71), in DBP are $\alpha_H = 5$, $\tau_c = 170$ ps, and $P_{sat} = 4.57$ mW (6.6 dBm). In our study, the saturation power (P_{sat}) for the DBP simulation was optimised to ensure optimal performance. We incorporated a standard fibre distortion mitigation DBP, employing the split-step method with a variable number of steps per span, in addition to the DBP for SOA distortion [150]. For both components of the DBP, we used two samples per symbol. It is crucial to highlight that our DSP chain adopts two different approaches based on the chosen nonlinearity mitigation technique, namely DBP or an ANN-based equaliser. Specifically, in the ANN implementation, nonlinearity mitigation occurs at the end of the process, subsequent to the CDC block. This arrangement ensures compatibility with existing DSP solutions and allows for comprehensive compensation of both linear and nonlinear impairments. The initial stages of the DSP chain, including the CDC, primarily focus on correcting linear signal distortions [155].

Following [103], we employ a 1-dimensional convolution layer as a pre-processing feature-extracting step. Furthermore, a bidirectional structure is used in the case of the GRU and LSTM components, as it provides a more robust solution for time-dependent tasks by learning better which features of the past values are useful for a particular symbol value prediction [69, 103]. Although CNN-biLSTM has been used as an equaliser in other contexts [103], to the best of our knowledge, this is the first time it is used for compensation of nonlinear impairments in a coherent detection WDM optical link with SOA. Furthermore, we proposed a new design for ANN-based channel equaliser based on the CNN-biGRU model.

We optimise the hyperparameters of each model using the BO algorithm for a launch power of -4 dBm (about 3 dB above the optimal power when ideal DBP is applied). The optimised hyperparameters include the number of filters of the 1-D convolutional layer (n_f), the filter size (n_k), the number of hidden units in the recurrent layer (n_h), the batch size (B), the memory depth (M), and the learning rate (η). The best values for CNN-biLSTM were found to be $M = 14$, $n_f = 234$,

$n_k = 5$, $n_h = 260$, $B = 2678$, and $\eta = 0.0035$. For CNN-biGRU, the optimal hyperparameters were $M = 40$, $n_f = 1000$, $n_k = 67$, $n_h = 100$, $B = 32$, and $\eta = 0.067$. These structures were mapped to a computational complexity (CC) value (Eqs. 3.72-3.73). This value was later decreased by pruning, until 95% of the weights were removed.

A total of 2^{18} symbols (2^{22} bits for 16-QAM symbols) were used for training, and 2^{23} independently generated symbols for evaluation, ensuring reliable results [156]. To avoid data periodicity and performance overestimation, the SIMD-oriented fast Mersenne twister random number generator algorithm was used to create the bit sequences.

Computational complexity directly impacts the latency, cost, and power consumption of DSP in an optical communication system, and needs to be minimised while maintaining satisfactory performance. In ANN models, compression techniques such as weight quantisation, knowledge distillation, or pruning are employed to achieve this. Pruning works by eliminating redundant or less important parts, such as weights, layers, or neurons, while maintaining performance [52, 68, 157]. Here, we prune the lowest magnitude weights globally throughout the ANN. Furthermore, in [52], we demonstrated that BO could improve existing pruning strategies, such as fine-tuning pruning, as an alternative to other explored approaches [158, 159]. In this study, we employ this methodology to reduce the CC of the proposed ANN-based equalisers. The CC for CNN-biLSTM and CNN-biGRU in terms of real multiplications per symbol (RMpS) is defined in Eqs. 3.72 and 3.73:

$$C_{\text{CNN-biLSTM}} = \big(\underbrace{n_i n_f n_k (n_s - n_k + 1)}_{a} + \underbrace{(n_s - n_k + 1)}_{b} \underbrace{2n_s(4n_h n_i + 4n_h^2 + 3n_h + n_o n_h)}_{c}\big), \tag{3.72}$$

$$C_{\text{CNN-biGRU}} = \big(\underbrace{n_i n_f n_k (n_s - n_k + 1)}_{a} + \underbrace{(n_s - n_k + 1)}_{b} \underbrace{2n_s(3n_h n_i + 3n_h^2 + 3n_h + n_o n_h)}_{d}\big), \tag{3.73}$$

In both equations, n_s represents the size of the input sequence (equal to $2M + 1$ where M is the memory depth), n_i is the number of input features per symbol (equal to 2 for single polarisation, corresponding to $d_{in} = 2$), n_o is the output dimension (equal to 2 for the real and imaginary parts of the recovered symbol), n_f is the number of filters used in the 1-D convolutional layer, n_k is the filter size, and n_h is the number of hidden units in the LSTM and GRU cells. The term a represents the contribution of the convolutional layer, while b defines the output size of each filter in the convolutional layer. The term c represents the contribution of the biLSTM layer, including matrix multiplications of inputs, hidden states, weights, and pointwise multiplications within the LSTM cell, whereas d accounts for similar contributions in the GRU cell but with fewer weight matrices.

The ANN-based equalisers were benchmarked against the low-complexity DBP algorithm proposed in [160] with 4 and 120 steps-per-span. The DBP algorithm is recognised for its simplicity and cost efficiency in compensating SOA nonlinearity. The CC of the DBP algorithm is determined by two primary elements: the Runge-Kutta component and the Fibre component, denoted as e and

f, respectively, in Equation 3.74:

$$C_{\text{DBP}} = \underbrace{\frac{30(N-1)}{N}}_{\text{e}} + \underbrace{\frac{4N\log_2(N)nzN_{\text{steps}}}{N}}_{\text{f}}. \tag{3.74}$$

Performance was assessed through the standard Q-factor metric recalculated from the directly counted BER as $Q = 20\log_{10}\left[\sqrt{2}\,\text{erfc}^{-1}(2\,\text{BER})\right]$ [52] for the central WDM channel of 11 transmitted, as shown in Figure 3.18. The results are depicted in Figure 3.19.

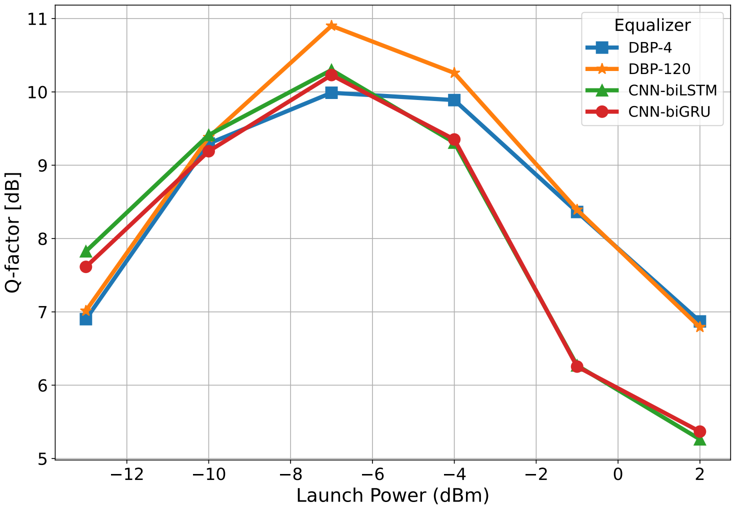


Figure 3.19: Performance comparison of the CNN-biGRU and CNN-biLSTM based equalisers and DBP with 4 and 120 steps per span for different signal launch powers.

Figure 3.19 shows that the ANN-based equalisers outperformed DBP-4 at optimal powers and are competitive with DBP-120. This demonstrates the potential of using ANNs to mitigate nonlinear impairments induced by inline SOAs. Figure 3.20 depicts the Q-factor and CC in terms of RMpS for the CNN-biLSTM and CNN-biGRU models. In this case, both networks have undergone a pruning process to reduce their CC to a sparsity level of 95%, meaning that only 5% of the initial weights remain. Consequently, the values obtained from Eqs. 3.72 and 3.73 are scaled by a factor of 0.05. Despite the reduction in CC, the optimisation of pruning parameters using BO maintains the performance of the models. This is evident when comparing Figs. 3.19 and 3.20.

Finally, we compare our low-complexity solution with the DBP algorithm proposed in [160]. The results are shown in Table 3.11.

The CNN-biLSTM and CNN-biGRU models, being generic equalisers, are generally computationally more demanding than DBP, which attempts to mitigate specific signal distortions. However, the BO-assisted pruning methodology achieves the best CC-performance balance, keeping or improving the Q-factor achieved [52,68]. As expected, increasing the steps per span of DBP results in better performance but also higher CC, as seen in Equation 3.74.

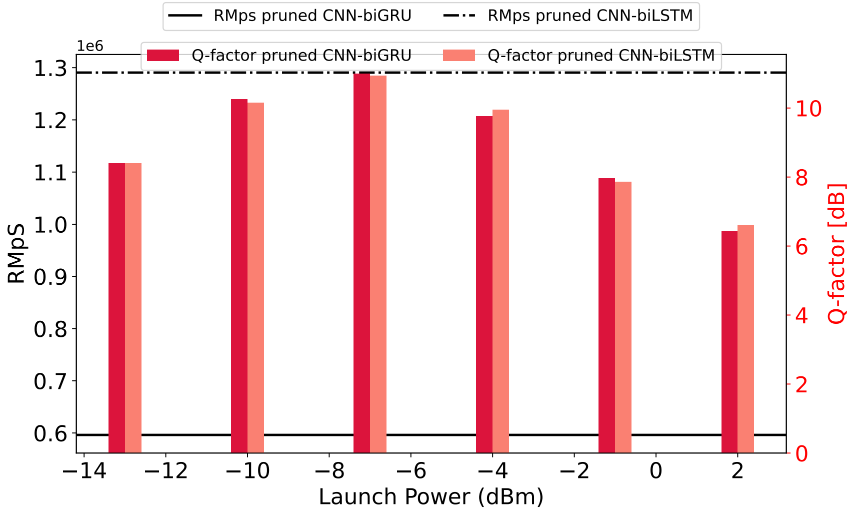


Figure 3.20: Performance versus complexity trade-off for the CNN-biLSTM and CNN-biGRU equalisers after pruning and reaching a sparsity level of 95% for different launch powers. Note: All power values referenced in this figure and related discussions (including P_{sat}) are expressed in dBm (referenced to 1 mW), not dB. The SOA saturation power is $P_{sat} = 8.32$ mW, equivalent to 9.2 dBm.

Model	CC (RMps)	Q-factor
DBP-120	1128100	10.90
DBP-4	37600	9.87
CNN-biLSTM	25811500	10.29
CNN-biGRU	11925000	10.18
CNN-biLSTM($s = 95\%$)	1290600	10.27
CNN-biGRU($s = 95\%$)	596300	10.34

Table 3.11: CC and Q-factor for the DBP with 120 and 4 steps per span and the ANN-based equalisers before and after undergoing pruning (95% of the weights removed) for a launch power equal to -7 dBm (the optimal for the system).

3.17 Synergistic Integration of OPC with ANNs-based Equalisers

Beyond pruning, an appealing route is to blend optical and digital remedies. As outlined in Chapter 2, OPC sits naturally alongside ANN-based equalisation. Injecting a phase-conjugated replica halfway along the fibre essentially rewinds much of the distortion picked up in the first span. Its real appeal is that it attacks the complexity problem from a completely different direction.

When we combine OPC with ANNs-based equalisers, we create a powerful synergy. OPC tackles most of the nonlinearities on-the-fly in hardware, and the ANN tidies up whatever distortion remains—be it asymmetry or imperfect conjugation. The hybrid setup taps into the best of both camps: OPC's hardware-level correction and the ANN's flexible, model-free learning.

To test this concept, we implemented an experimental setup using a PDM transmission at 28

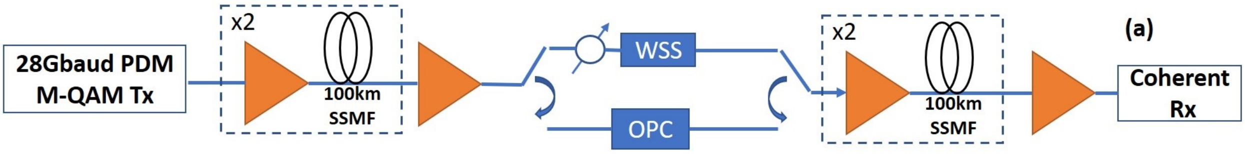


Figure 3.21: Block diagram of the experimental setup used to evaluate OPC and CVNN synergy [1]

GBaud with 64-QAM signals through 4 spans of standard single-mode fibre, as shown in Figure 3.21. For configurations using OPC, we placed the phase conjugator at the midpoint of the link and transmitted the conjugated signal through the second half. For comparison, we also tested configurations without OPC, carefully matching the OSNR conditions to ensure fair evaluation.

The biggest eye-opener was how much OPC trimmed the size of our CVNN designs. BO kept choosing leaner networks—shallower and with fewer neurons—whenever OPC was employed, as shown in Table 3.12. For instance, at 9 dBm launch power, the OPC configuration required only 5 layers compared to 6 layers without OPC. Also noteworthy is how the networks without OPC tended toward maximum neuron counts in multiple layers (600 neurons), suggesting they were struggling to model the more complex channel characteristics.

In this case, CC is measured as the number of complex-valued multiplications per symbol period, $\mathrm{CC} = (1-s)\Big(n_s n_i n_1 + \sum_{l=1}^{L-1} n_l n_{l+1} + n_L n_o\Big)$, with sparsity s introduced by pruning.

Plain magnitude pruning made the advantage stand out even more clearly. As Figure 3.22 shows, OPC allowed to drop 60% of the weights at 8 dBm launch power and still stay below the HD-FEC limit (BER = 3.8·10^{-3}). In contrast, without OPC the model missed the target no matter how big we made it. Taking this further, we evaluated more sophisticated pruning techniques. The traditional FT approach, i.e. magnitude-based pruning, was compared with WR and LR strategies. Pairing WR with OPC was even better: we could discard 90% the parameters at either launch power and the BER still cleared the threshold. This resulted in a huge cut in compute load with zero hit to performance, as it is highlighted in Figure 3.22.

This synergistic approach challenges the conventional wisdom that treats optical and digital compensation as separate domains. Because OPC shortens the channel memory, it lets us build lighter networks that need far fewer parameters. Pruning, which would normally hurt accuracy, now slashes the count of complex multiplies even further when OPC is in the loop.

Altogether, mixing optical tricks with smart pruning looks like a powerful recipe for fast, lightweight equalisation in future fibre systems. Bringing optical and digital methods together moves us closer to rolling out advanced nonlinear compensation in live networks.

Table 3.12: Optimised CVNN architectures with and without OPC at launch powers of 8 dBm and 9 dBm.

Optical solution	**Optimal parameters**
Launch power: 8 dBm	
OPC	$n = [519, 505, 531, 151, 562, 140]$
	$l = 6, \quad p = 0.1, \quad L_2 = 6.5 \times 10^{-6}$
No OPC	$n = [366, 422, 600, 600, 242, 327]$
	$l = 6, \quad p = 0, \quad L_2 = 1 \times 10^{-6}$
Launch power: 9 dBm	
OPC	$n = [430, 532, 406, 462, 315]$
	$l = 5, \quad p = 0.35, \quad L_2 = 6.5 \times 10^{-3}$
No OPC	$n = [600, 600, 600, 600, 568, 50]$
	$l = 6, \quad p = 0.1, \quad L_2 = 0.1$

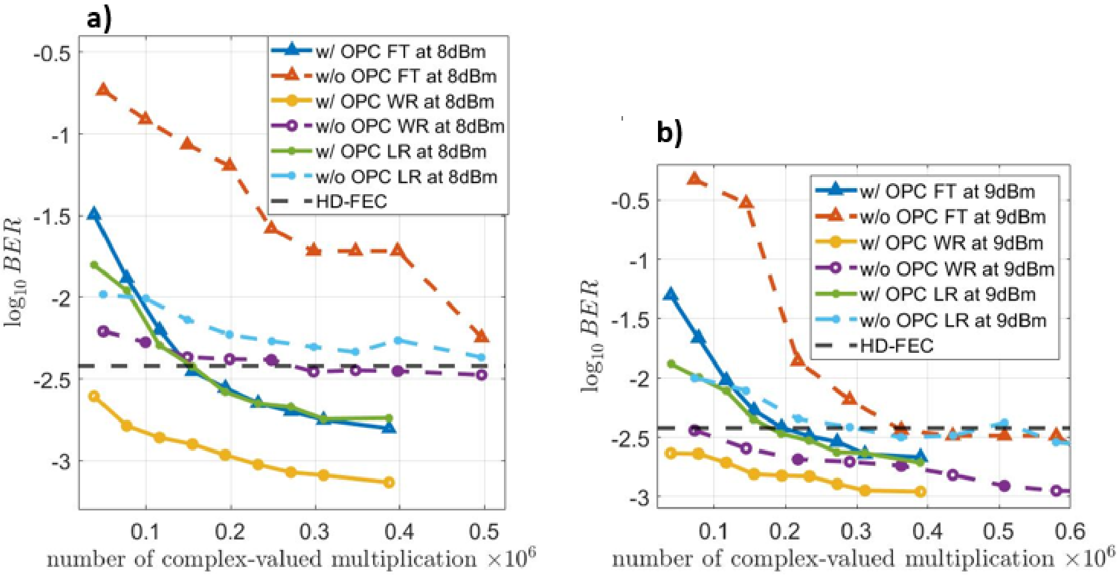


Figure 3.22: BER achieved for different computational complexity values with and without OPC at two launch powers (8 dBm and 9 dBm). The horizontal dashed line indicates the HD-FEC threshold.

Chapter 4

Addressing the Noise Problem in Photonic Neuromorphic Computing: New Opportunities for Efficient and Noise-Resistant Optical Channel Equalisation

The physical implementation of ANNs in analogue hardware presents a promising avenue for achieving high-speed, energy-efficient signal processing in optical communication systems. However, analogue implementations inherently introduce noise and variability that can significantly impact the performance of these neural network-based equalisers. This chapter explores the challenges posed by noise in analogue ANNs and presents novel techniques to enhance their resilience in the context of optical communication systems.

Analogue computing offers the potential for substantial improvements in speed and energy efficiency compared to traditional digital implementations. This is particularly relevant in the field of optical communications, where the demands for high-speed signal processing continue to grow. However, the transition from digital to analogue neural networks introduces new challenges, primarily related to the precision and reliability of computations in the presence of various noise sources.

In this chapter, we will first examine the fundamental sources of noise in analogue neural network implementations, including thermal noise, shot noise, and quantisation errors. We will then discuss how these noise sources propagate through the network and affect its overall performance.

Building upon this foundation, we will introduce advanced techniques for enhancing the noise resilience of analogue neural networks. These techniques include novel training methodologies, architectural modifications, and noise-aware design principles. We will pay special attention to the concept of “crowd equalisation,” which leverages ensemble methods to improve robustness against noise and variability. Throughout the chapter, we will present rigorous experimental results and theoretical analyses to validate the proposed techniques. We will evaluate the performance of our noise-resistant analogue neural networks in realistic optical communication scenarios, considering metrics such as BER and SNR.

By addressing the critical issue of noise in analogue neural network implementations, this chapter contributes to the broader goal of realising practical, high-performance neuromorphic computing solutions for next-generation optical communication systems. The techniques and insights presented here have the potential to significantly advance the field of analogue neural networks and pave the way for their widespread adoption in optical signal processing applications.

Moreover, in this chapter SNNs are also explored as promising paradigm for neuromorphic computing in optical communications while the unique challenges they present are assessed and addressed. Furthermore, novel architectures designed are explored to develop promising equalisation methods when compared with traditional ones.

The results presented in this chapter come from the articles [73, 161], where the experimental dataset was provided by Dr.Abdallah Alsayed Ismail Ali. The figures and models employed in [161] were done by Karina Nurlybayeva, while I supported in writing the text and reporting the results as well as supporting with the noise-aware training process. The research presented here was done under supervision of Dr.Morteza Kamalian-Kopae, and Sergei K. Turitsyn

4.1 Noise Resilience of Analogue ANNs

Analogue artificial neural networks are widely considered as promising computational models that more closely imitate the information processing capabilities of the human brain compared to digital neural networks. The significant computation power and the much reduced power consumption per operation make the analogue implementation of neural networks very attractive. There is an active research on ANNs implementation using both analogue photonic and electronic hardware [72, 162]. However, compared to digital realisations the conventional analogue systems are more sensitive to the noise that is inevitably present in practical implementations.

Before investigating noise robustness in the context of optical communications, we establish a rigorous theoretical foundation that connects image classification tasks with optical signal processing challenges. This connection is not merely analogical but stems from fundamental mathematical similarities in data representation and noise characteristics that make image classification an ideal controlled testbed for studying noise effects in neural networks before deployment in optical systems.

To better understand the results in this chapter, it is necessary to stablish a unified framework between image classification and equalisation. The relationship between these two domains can be understood through the lens of complex-valued signal spaces. In optical communication systems, transmitted symbols are represented as points in a complex constellation diagram, such as QAM, where each symbol s_k is defined as:

$$s_k = I_k + jQ_k \in \mathbb{C} \tag{4.1}$$

where I_k and Q_k represent the in-phase and quadrature components, respectively. Similarly, grayscale images can be represented as discrete amplitude levels in a 2D spatial domain, where each pixel intensity $p_{i,j}$ occupies a position in the image plane analogous to constellation points in the I-Q plane.

The fundamental mathematical parallel emerges when considering that both image pixels and QAM constellation points represent discrete classification problems in continuous spaces. In image classification, the neural network must correctly classify pixel patterns into discrete categories despite additive noise. In optical equalisation, the network must correctly classify received noisy constellation points back to their original transmitted symbols. Both tasks require the network to learn decision boundaries in the presence of additive Gaussian noise, making the underlying mathematical optimisation problem structurally equivalent [100, 163].

The AWGN commonly encountered in image processing serves as an excellent model for studying the effects of photonic noise in optical systems. In digital images, noise manifests as:

$$\tilde{p}_{i,j} = p_{i,j} + n_{i,j}, \quad n_{i,j} \sim \mathcal{N}(0, \sigma^2) \tag{4.2}$$

This directly parallels the signal-based noise model in photonic systems where the received signal becomes:

$$\tilde{y} = y + \hat{N}(y), \quad \hat{N}(y) \sim \mathcal{N}(0, \sigma^2_{photonic}) \tag{4.3}$$

The crucial insight is that both scenarios involve neural networks learning robust feature representations that are invariant to additive noise corruption. The statistical properties of the noise—Gaussian distribution, zero mean, and additive nature—are preserved across both domains, enabling direct transferability of noise robustness insights [164].

Image classification on standardised datasets such as MNIST [165] provides several critical advantages as a controlled environment for studying noise effects:

1. **Standardised Baselines:** Well-established performance metrics and baseline accuracies enable precise quantification of noise-induced performance degradation.
2. **Scalable Complexity:** The 10-class classification problem in MNIST provides sufficient complexity to observe meaningful noise effects while remaining computationally tractable for extensive hyperparameter exploration.
3. **Reproducible Conditions:** Fixed datasets eliminate variability from experimental conditions, enabling systematic study of noise injection parameters and training methodologies.

4. **Architectural Generalizability:** Results obtained on image classification tasks can be systematically validated across different neural network architectures (MLPs, CNNs, RNNs) relevant to optical signal processing.

Validation Methodology and Performance Indicators: The translation of insights from image classification to optical communications is validated through several key performance indicators that exhibit direct correspondence:

$$\text{Classification Accuracy} \leftrightarrow \text{Symbol Error Rate (SER)} \tag{4.4}$$

$$\text{Cross-Entropy Loss} \leftrightarrow \text{MSE} \tag{4.5}$$

$$\text{Robustness to Image Noise} \leftrightarrow \text{Resilience to Channel Impairments} \tag{4.6}$$

The fundamental principle underlying this validation is that neural networks demonstrating enhanced robustness to additive noise in image classification tasks will exhibit corresponding improvements in noise resilience when applied to optical signal equalisation. This transferability has been empirically validated in subsequent sections of this work and provides the theoretical foundation for using image classification as a proxy for optical system performance analysis.

Noise properties in ANNs have been studied both in the electronic and photonic domains. However, photonic ANNs are much less investigated compared to the electronic implementations, for which some training techniques have been proposed to enhance ANNs robustness against noise [162]. The theoretical bridge established above provides the mathematical justification for using image classification tasks as a controlled environment to develop and validate noise-aware training techniques before their application to more complex optical communication scenarios.

4.1.1 Literature Review: Noise Robustness in Neural Networks

The study of noise effects in neural networks has a rich history spanning both theoretical analysis and practical implementation considerations. Early foundational work established the mathematical frameworks for understanding how additive and multiplicative noise affects network performance [164]. The field has evolved from simple noise injection techniques to sophisticated regularisation strategies and adversarial training methods.

In the electronic domain, [166] established fundamental principles for adversarial robustness that share conceptual similarities with noise robustness, where networks must maintain performance despite input perturbations. Their work on adversarial training—where networks are explicitly trained on perturbed inputs—laid the groundwork for noise-aware training methodologies. Similarly, [167] developed certified robustness techniques using randomised smoothing, providing theoretical guarantees for network performance under Gaussian noise, which directly relates to the photonic noise scenarios studied in this work.

The broader robustness literature has established that neural networks exhibit varying sensitivity to different types of perturbations [168]. These studies have shown that standard training procedures often result in networks that are brittle to even small input corruptions, motivating the development of specialised training techniques for enhanced robustness.

In the context of Photonic Neural Network, the seminal work of Semenova and colleagues represents the most comprehensive investigation of noise effects in analogue photonic neural networks to date, establishing fundamental principles that directly inform the present study [71, 72, 169, 170]. In the 2019 study [72] they established the theoretical framework for understanding noise propagation in analogue hardware neural networks. Their work identified key differences between electronic and photonic implementations:

- **Noise Sources:** Electronic systems primarily exhibit weight-level noise due to device variability and thermal fluctuations in resistive elements, while photonic systems show signal-level noise from shot noise, thermal noise in photodetectors, and phase fluctuations in optical components.
- **Propagation Characteristics:** Their analysis demonstrated that noise in photonic systems exhibits different propagation dynamics compared to electronic implementations, with multiplicative noise components becoming significant due to the intensity-dependent nature of optical signal processing.
- **Performance Implications:** The study established that photonic neural networks can exhibit both enhanced robustness (due to the analogue nature of optical processing) and increased vulnerability (due to accumulation of noise through multiple optical stages).

With respect to training and noise mitigation strategies, building on foundational analyses of noise propagation in analogue/physical neural networks [72], Semenova *et al.* proposed training and design practices to improve robustness in *trained deep networks* [169]. Their 2022 study models noisy nonlinear neurons in fully connected layers and derives criteria for noise-resilient design; in practice, one effective technique is to inject noise during training ("noise-aware" training), so that inference remains stable under hardware perturbations.

1. **Noise-aware training.** Injecting controlled noise during training improves robustness at test time across trained deep networks [169]. In this work we adopt that principle and examine alternative injection schemes (e.g., weight clipping vs. Gaussian injection).
2. **Architecture and depth.** The 2022 analysis shows that noise accumulation in layered networks is *generally bounded* rather than unconditionally worsening with depth; sensitivity depends on layer statistics and where noise enters the computation [169]. This motivates our controlled comparison across architectures and depths rather than assuming "deeper $\Rightarrow$ less robust."

3. **Hardware-aware design rules.** The same study provides engineering criteria for building noise-resilient hardware implementations [169]. We extend this by experimentally comparing training-time noise injection strategies under identical datasets and evaluation protocols.

A recent preprint from the same line of work analyses *analogue echo-state (recurrent)* networks under white-Gaussian *internal* noise [170]. While its setting is recurrent (not a feed-forward image-classification benchmark), two observations are directly relevant to our study:

- **Pronounced recurrent sensitivity.** ESN performance can degrade sharply under certain additive/multiplicative and correlated/uncorrelated noise regimes, highlighting the particular vulnerability of recurrent dynamics to internal noise [170]. This motivates our feed-forward vs. recurrent comparison under a common training and evaluation pipeline.
- **Threshold-like regimes and mitigation.** The preprint reports critical noise thresholds beyond which performance collapses, and shows that appropriately tuned noise-aware training can recover a significant fraction of lost accuracy [170]. These findings support our grid-search procedure for selecting noise parameters and our use of noise-aware training.

Complementing the photonic-focused work of Semenova et al., Zhou and colleagues [162] provided crucial insights into noise robustness in electronic neural network implementations. Their "noisy machines" framework established several principles that inform our comparative analysis:

- **Distillation-Based Robustness:** They introduced knowledge distillation techniques for enhancing noise robustness, demonstrating that networks can learn to maintain performance despite hardware imperfections.
- **Analogue Hardware Considerations:** Their work specifically addressed the transition from digital to analogue implementations, identifying key challenges that parallel those encountered in photonic systems.
- **Cross-Domain Transferability:** They demonstrated that robustness techniques developed in one hardware domain can often be adapted to others, supporting our approach of using image classification insights for optical communication applications.

Research Gaps and Novel Contributions. While the existing literature provides important foundations, several gaps remain that this work addresses:

1. **Limited Architectural Comparison:** Previous work has not systematically compared noise robustness across diverse architectures (MLPs vs. RNNs) using identical training protocols and noise models.
2. **Controlled Environment Analysis:** The use of standardised image classification tasks as a controlled testbed for photonic noise analysis represents a novel methodological contribution not present in existing literature.

3. **Unified Training Framework:** This work provides a systematic comparison of different noise injection strategies (weight clipping vs. Gaussian injection) within a unified experimental framework.

4. **Bridge to Optical Communications:** While previous work focused on general neural network applications, this study explicitly connects noise robustness research to optical communication equalisation challenges.

Contemporary Developments. Recent advances in optical neural network hardware [40, 171] have highlighted the growing importance of noise robustness research. These studies have shown that:

- **Hardware Maturity:** Integrated photonic neural networks are approaching practical implementation, making noise robustness a critical consideration for real-world deployment.
- **Scalability Challenges:** Larger optical neural networks exhibit increased sensitivity to cumulative noise effects, emphasizing the importance of robust training methodologies.
- **Application Diversity:** Optical neural networks are being deployed across diverse applications, from telecommunications to machine learning acceleration, requiring robust performance across different operational conditions.

Building upon these foundations, this work extends the contributions of Semenova et al. and Zhou et al. by:

1. **Systematic Methodology:** Providing a rigorous experimental framework that enables direct comparison of different architectures and training strategies under identical conditions.
2. **Theoretical Foundation:** Establishing mathematical connections between image classification robustness and optical signal processing performance that were not previously formalised.
3. **Practical Implementation:** Demonstrating techniques that can be directly applied to optical communication systems, bridging the gap between fundamental research and practical applications.
4. **Comprehensive Analysis:** Offering insights that span from theoretical noise modelling to practical implementation considerations, providing a complete framework for noise-robust photonic neural network design.

The subsequent sections build upon this literature foundation to present novel experimental results and theoretical insights that advance the state of the art in noise-robust photonic neural networks.

4.1.2 Noise Models in Analogue Neural Networks

Before analysing the performance of noise-resilient training approaches, it is essential to understand how noise manifests in different analogue hardware implementations. In this work, we focus on two distinct noise models that represent the primary differences between electronic and photonic hardware:

1. **Weight-based noise model (electronic hardware):** In this model, noise affects the weights directly, represented as:
$$\mathbf{W}_l = \mathbf{W}_l^0 + \Delta\mathbf{W} \tag{4.7}$$
where $\Delta\mathbf{W}$ follows a Gaussian distribution centred at zero.
2. **Signal-based noise model (photonic hardware):** Here, noise impacts the output signal rather than the weights:
$$y = y' + \hat{N}(y') \tag{4.8}$$
where y' is the clean output and $\hat{N}(y')$ represents the photonic noise operator, which can take different forms such as:
$$\hat{N}(y') = \sqrt{2D_A}\xi_A \quad \text{or} \quad \hat{N}(y') = y'\sqrt{2D_M}\xi_M \tag{4.9}$$
with diffusion coefficients $D_A = 10^{-4}$ and $D_M = 10^{-3}$, and standard normal random variables $\xi_A, \xi_M \sim \mathcal{N}(0, 1)$.

Introducing two diffusion coefficients, D_A (additive) and D_M (multiplicative), in the photonic noise model is well-grounded in stochastic process theory and in the physics of optical receive chains. The dual-coefficient formulation captures signal-independent noise (e.g., detector shot and thermal noise) and signal-dependent noise (e.g., laser RIN, phase-to-amplitude conversion, amplifier-related terms) that co-exist in practical photonic neural hardware.

The composite noise dynamics can be modelled with an SDE comprising additive and multiplicative terms [172, 173]:

$$dY_t \;=\; f(Y_t, t)\,dt \;+\; \sigma_A(Y_t, t)\,dW_t^A \;+\; \sigma_M(Y_t, t)\,Y_t\,dW_t^M, \tag{4.10}$$

where Y_t is the optical-signal amplitude, f is the drift, σ_A and σ_M are noise intensities for the additive and multiplicative components, and W_t^A, W_t^M are independent Wiener processes. In discrete inference layers this reduces to the operator

$$\hat{N}(y') \;=\; \sqrt{2D_A}\,\xi_A \;+\; y'\,\sqrt{2D_M}\,\xi_M, \tag{4.11}$$

with $\xi_A, \xi_M \sim \mathcal{N}(0, 1)$. The factor 2 follows the Itô convention for diffusions ($\mathrm{Var}[\sqrt{2D}\,W_t] = 2D\,t$) [172].

D_A aggregates signal-independent contributions:

- **Photodetector shot noise:** white current PSD $S_I = 2eI$ (with photocurrent I), equivalently a shot-noise floor at the detector input after responsivity scaling [174, 175].
- **Thermal (Johnson–Nyquist) noise:** voltage/current noise with one-sided PSD $4k_BTR$ for a resistor R [176, 177].
- **Dark-current shot noise:** shot noise from detector dark current adds to the same $2eI$ term [175].

These mechanisms are independent of $|y'|$, justifying the $\sqrt{2D_A}\,\xi_A$ term.

Meanwhile, D_M captures signal-dependent mechanisms:

- **Laser relative-intensity noise (RIN):** intensity fluctuations normalised to average power and commonly specified in dB/Hz; the contribution scales with optical power [175, Section 3.5, 4.4], [174, Chapter 18].
- **Phase noise → amplitude noise conversion:** interferometric paths or fibre dispersion convert laser phase noise into detected intensity noise, introducing a signal-dependent component [178].
- **Amplifier noise figure and gain fluctuations:** noise added by optical amplification scales with signal level and sets an OSNR floor [175, Section 4.2.1.5].

These justify the $y'\sqrt{2D_M}\,\xi_M$ term.

In terms of SNR, for $\hat{N}(y') = \sqrt{2D_A}\xi_A + y'\sqrt{2D_M}\xi_M$ with independent standard normals,

$$\mathrm{SNR}(y') \;=\; \frac{|y'|^2}{\mathrm{Var}[\hat{N}(y')]} \;=\; \frac{|y'|^2}{2D_A + 2D_M|y'|^2}. \tag{4.12}$$

Three regimes follow directly:

1. **Low signal** ($|y'|^2 \ll D_A/D_M$): additive term dominates, $\mathrm{SNR} \approx |y'|^2/(2D_A)$;
2. **High signal** ($|y'|^2 \gg D_A/D_M$): multiplicative term dominates, $\mathrm{SNR} \approx 1/(2D_M)$ (SNR ceiling);
3. **Intermediate** regime: both terms matter, motivating the full dual-coefficient model.

With $(D_A, D_M) = (10^{-4}, 10^{-3})$ used in this chapter, the transition occurs near $|y'|^2 \approx 0.1$, so both mechanisms are relevant across typical activation ranges.

The ratio $D_M/D_A = 10$ reflects that, at moderate–high optical powers, multiplicative mechanisms (RIN, phase-to-amplitude conversion, amplifier noise) often dominate, whereas at low powers the detector chain is typically shot/thermal limited. These magnitudes are consistent with textbook receiver models and RIN behaviour in semiconductor lasers [174, 175].

It is ecessary to employ both coefficients, as additive-only models ($D_M = 0$) over-predict SNR growth with signal power and miss the SNR ceiling imposed by multiplicative effects; they also understate dynamic-range compression and can mislead training if the deployment hardware has appreciable RIN or phase-to-amplitude paths. The dual model is therefore necessary for realistic prediction and for designing effective noise-aware training.

4.1.3 Noise-Resilient Training Techniques

We investigate two primary noise injection approaches to enhance the robustness of analogue neural networks against photonic noise.

- **Gaussian Noise Injection with Weight Clipping (WC):** During training, the weight $\mathbf{W}_l$ of layer l is drawn from a Gaussian distribution $\mathbf{N}(\mathbf{W}_l; \mathbf{W}_0^l, \sigma_{N,l}^2 \mathbf{I})$. The noise variance is referenced to the range of representable weights:

$$\sigma_{N,l}^2 = \eta(W_{max}^l - W_{min}^l) \tag{4.13}$$

  The values W_{max}^l and W_{min}^l are obtained based on the statistics of the weights. The hyperparameter η controls the noise level, and weights are clipped to the interval $[W_{min}^l; W_{max}^l]$.

- **Gaussian Noise Injection without Weight Clipping (GI):** In this method, Gaussian noise is directly injected into the weights without applying any clipping constraint:

$$\mathbf{W}_l = \mathbf{W}_l^0 + \mathcal{N}(0, \sigma^2 \mathbf{I}) \tag{4.14}$$

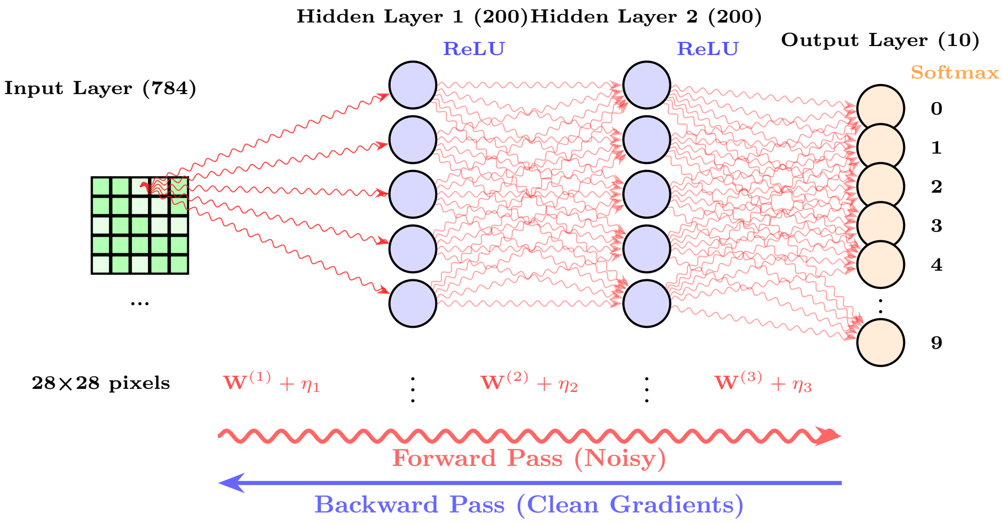
  where σ^2 denotes the noise variance hyperparameter.

4.1.4 Experimental Setup

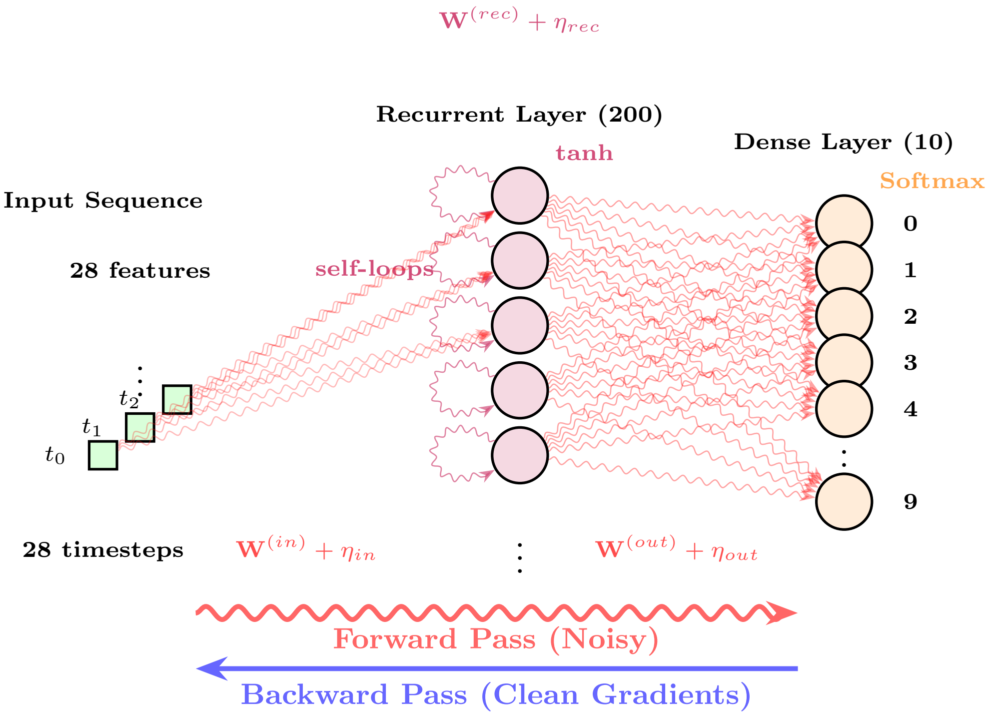
To evaluate these noise-resilient training techniques, we conducted experiments using two network architectures:

1. A MLP with two hidden layers of 200 units each.
2. A RNN with one recurrent layer of 200 units followed by a dense output layer with 10 units.

Both networks were trained on the MNIST dataset [179], chosen specifically for its relevance to optical signal processing as established in the theoretical bridge above. MNIST provides an ideal controlled environment where:



(a) MLP noise-aware training. Noisy forward pass; gradients computed on the clean graph.



(b) RNN noise-aware training. Injection on $W^{(\mathrm{in})}$, $W^{(\mathrm{rec})}$ and $W^{(\mathrm{out})}$ during the forward pass; clean gradients.

Figure 4.1: Noise-aware **training** schematics used in this chapter. We inject zero-mean Gaussian perturbations during the forward pass and keep the backward pass clean; η controls the noise level.

- **Discrete Classification Parallels:** The 10-class handwritten digit classification problem mirrors the discrete symbol detection task in digital optical communications, where received constellation points must be correctly classified despite channel impairments.

- **Noise Model Validity:** Additive Gaussian perturbations used in our training-time ablations capture detector- and electronics-limited regimes (shot and Johnson–Nyquist noise) [176, 177, 180]. During evaluation we additionally apply the dual photonic noise operator from §4.1.2—including multiplicative components representative of RIN and phase-to-amplitude conversion—to reflect hardware conditions more faithfully [178, 181, 182].
- **Baseline Reproducibility:** Well-established performance benchmarks (>95% accuracy for standard MLPs) enable precise quantification of noise-induced degradation, providing clear metrics for evaluating the effectiveness of noise-aware training techniques.
- **Computational Tractability:** The 28×28 pixel input dimension and moderate computational requirements allow extensive hyperparameter exploration and statistical analysis across multiple noise injection strategies.

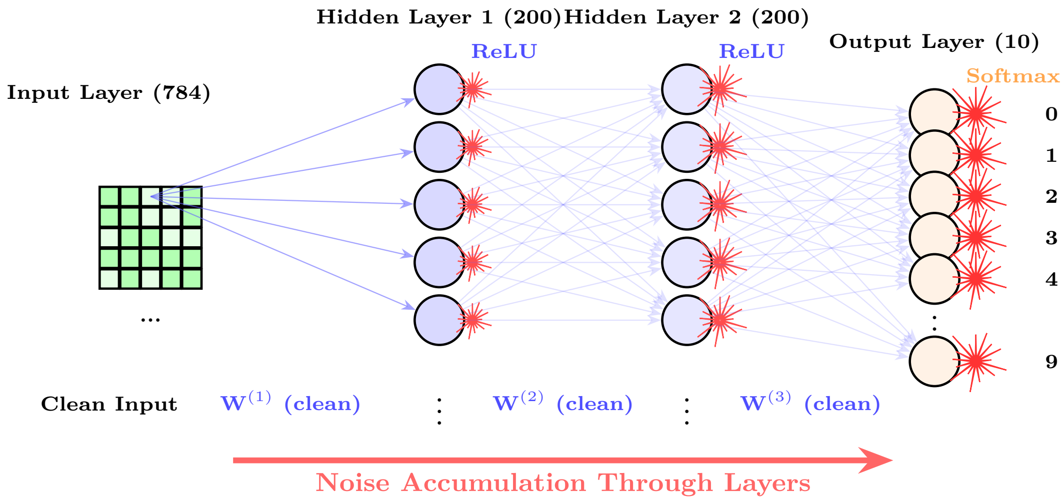
The experimental validation proceeds by subjecting trained networks to photonic noise during inference, simulating the conditions expected in analogue photonic neural network implementations. This methodology enables direct assessment of how image classification robustness translates to optical communication performance (see Figure 4.2).

The selection of optimal noise injection parameters η represents a critical optimisation problem that directly determines the effectiveness of noise-aware training. Our systematic approach addresses this challenge through a comprehensive grid search methodology grounded in both theoretical considerations and empirical validation [113, 183].

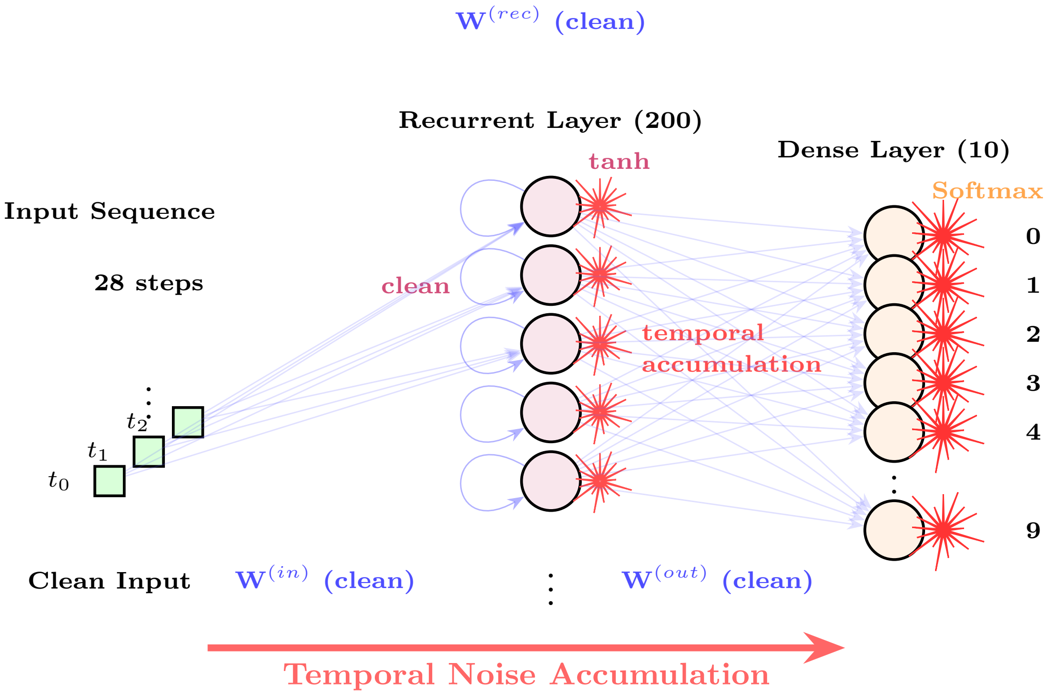
The noise parameter ranges were carefully determined through theoretical analysis of photonic noise dynamics and empirical validation from comparable analogue neural network studies. For the MLP, the optimal range was set to $\eta \in [2.5 \times 10^{-4}, 9 \times 10^{-4}]$ with logarithmic spacing, capturing the region where training stability and robustness improvements converge. In the case of the RNN, due to their inherently higher sensitivity to temporal noise accumulation, employed a broader range of $\eta \in [0.8 \times 10^{-3}, 1.1 \times 10^{-2}]$ with denser sampling around the expected optimal values.

These ranges reflect fundamental architectural differences in noise sensitivity. MLPs demonstrate inherent robustness due to their feedforward structure, requiring lower noise injection levels to achieve optimal training benefits. RNNs, conversely, exhibit amplified noise sensitivity through recurrent connections, necessitating higher noise injection levels to overcome the regularisation effects needed for robust temporal processing [184, 185].

Grid-search tuning is applied to the training noise scale η. We tune the single hyperparameter η with a one-dimensional logarithmic sweep that includes the zero-noise point ($\eta = 0$). For the MLP we evaluate $\eta \in \{0,\ 2.5,\ 4,\ 6,\ 8,\ 9\} \times 10^{-4}$; for the RNN we evaluate $\eta \in \{0,\ 0.8,\ 1,\ 2,\ 5,\ 8,\ 11\} \times 10^{-3}$. The best value $\eta^{\star}$ is selected on a clean validation set (noise-agnostic selection). All sweep plots in



(a) MLP inference under photonic noise.



(b) RNN inference under photonic noise (temporal accumulation).

Figure 4.2: Schematic illustration of the noisy inference process for both MLP and RNN architectures, highlighting layer-wise and temporal noise accumulation effects in analogue photonic implementations.

this chapter use the same horizontal axis $[10^{-4}, 10^{-2}]$ (log scale) for panels (a) and (b), and overlay the clean ($\eta = 0$) reference as a dashed curve/marker, as requested by the examiner.

Algorithm 2 Noise-aware training sweep for η (single-parameter grid search)

```
Require: Architecture $\mathcal{A} \in \{\text{MLP}, \text{RNN}\}$; grid $\{\eta_k\}_{k=1}^{K}$ (log-spaced, includes $\eta = 0$); clean validation set
for each $\eta_k$ in $\{\eta_k\}$ do
    Initialize base weights $W_0$ (fixed seed)
    for epoch $e = 1$ to $E_{\max}$ do
        for each minibatch $B$ do
            Noisy forward: $W \leftarrow W_0 + \mathcal{N}(0,\, \eta_k\, \sigma_W^2 I)$; compute loss on $B$
            Clean backward: backpropagate w.r.t. the clean graph to update $W_0$
        end for
        Evaluate accuracy on the clean validation set; early-stop with patience $p$ (e.g., $p{=}10$)
    end for
    Record $\mathrm{Acc}_{\mathrm{val}}(\eta_k)$ and save the checkpoint
end for
return $\eta^\star = \arg\max_{\eta_k} \mathrm{Acc}_{\mathrm{val}}(\eta_k)$    (deploy this checkpoint)
```

4.1.5 Detailed Performance Analysis

Figure 4.3 and Figure 4.4 present the validation accuracy of both architectures as a function of the noise parameter η. These curves provide critical insights into the relationship between training noise levels and model performance under noisy inference conditions.

It is important to notice that the different parameter ranges between MLP ($\eta \in [2.5 \times 10^{-4}, 9 \times 10^{-4}]$) and RNN ($\eta \in [0.8 \times 10^{-3}, 1.1 \times 10^{-2}]$) reflect fundamental architectural differences in noise sensitivity. MLPs demonstrate robustness at lower noise levels due to their feedforward structure, while RNNs require higher noise injection parameters to achieve comparable training effectiveness due to recurrent weight dependencies and temporal error propagation. The baseline performance represents zero-noise training conditions, providing reference points for quantifying noise-aware training improvements.

For the MLP (Figure 4.3), we observe distinct patterns for each training method:

- **Weight Clipping (WC):** The accuracy shows a clear monotonic increase with noise parameter η, from 88.90% at $\eta = 3 \times 10^{-4}$ to 89.30% at $\eta = 8 \times 10^{-4}$. This consistent improvement suggests that higher levels of injected noise during training (within the tested range) progressively enhance the MLP's robustness to photonic noise.
- **Gaussian Injection (GI):** The accuracy exhibits a non-monotonic relationship with η, starting high at 89.55% for $\eta = 3 \times 10^{-4}$, dropping to a minimum of 89.05% at $\eta = 6 \times 10^{-4}$, and then increasing again to 89.45% at $\eta = 8 \times 10^{-4}$. This U-shaped curve suggests an optimal noise range exists, with both too little and too much noise being suboptimal.
- **Comparison to Baseline:** The standard training (noise-agnostic) achieves 89.0% accuracy.

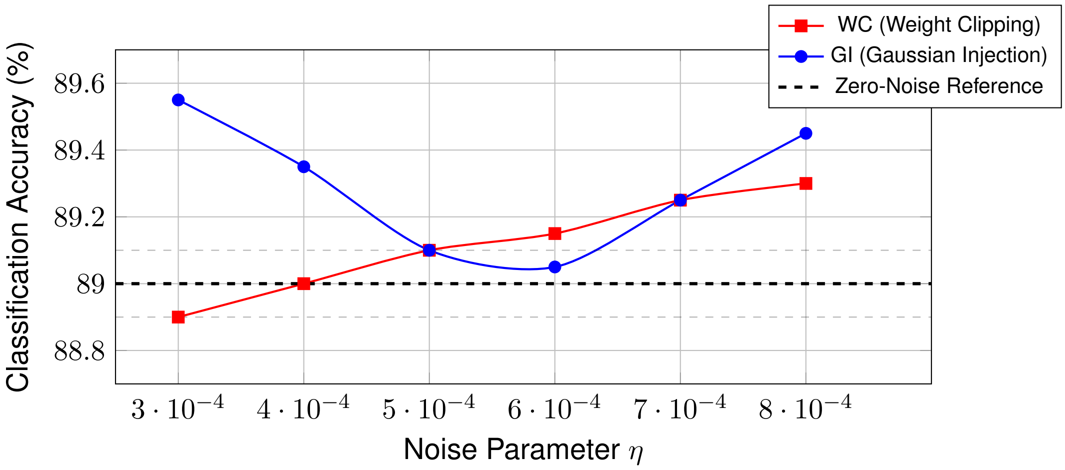


Figure 4.3: Noise-aware training performance of MLP (2×200 hidden units).

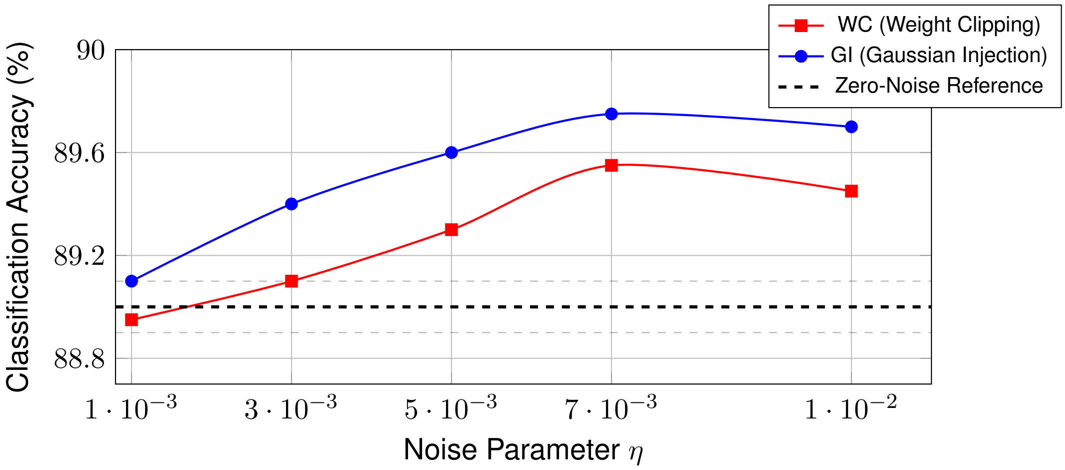


Figure 4.4: Noise-aware training performance of RNN (200 recurrent units).

GI consistently outperforms this baseline across all tested η values, while WC exceeds the baseline only for $\eta \geq 4 \times 10^{-4}$.

The peak performance of 89.55% is achieved by GI at $\eta = 3 \times 10^{-4}$, representing a 0.55 percentage point improvement over the baseline. This modest gain suggests that MLPs already possess reasonable inherent robustness to photonic noise, but can still benefit from appropriate noise-aware training.

For the RNN (Figure 4.4), the results reveal dramatically different patterns:

- **Weight Clipping (WC):** The accuracy shows a non-monotonic relationship with η, peaking at 83.5% for $\eta = 1 \times 10^{-3}$, decreasing to a minimum of 79.0% at $\eta = 5 \times 10^{-3}$, and then partially recovering to 81.5% at $\eta = 1 \times 10^{-2}$. This suggests that RNNs are highly sensitive to the precise calibration of noise during training.

- **Gaussian Injection (GI):** Similar to WC, GI exhibits a non-monotonic pattern but consistently outperforms WC across all tested noise levels. The accuracy peaks at 85.6% for $\eta = 1 \times 10^{-3}$, decreases to 82.0% at $\eta = 5 \times 10^{-3}$, and then recovers to 84.5% at $\eta = 1 \times 10^{-2}$.

- **Comparison to Baseline:** The standard training achieves only 45.5% accuracy, demonstrating the RNN's extreme vulnerability to photonic noise without specialised training. Both noise-aware techniques dramatically improve performance, with gains of up to 40.1 percentage points.

The substantial performance difference between noise-aware training and the baseline for RNNs (compared to the modest difference for MLPs) underscores the critical importance of noise-resilient training for recurrent architectures in photonic implementations.

Table 4.1 summarizes the best accuracies achieved by each combination of architecture and training method, along with the corresponding optimal noise parameters.

Table 4.1: Comprehensive Performance Summary: Neural network accuracy on MNIST classification under photonic noise using different noise-aware training methods. Dual diffusion noise model with additive (D_A) and multiplicative (D_M) components representing shot noise, thermal noise, and process variations in analogue photonic hardware. Optimal η parameters selected based on grid search over logarithmically spaced values, with ranges reflecting architectural noise sensitivity differences in photonic implementations.

Method	MLP best accuracy values (%)	RNN best accuracy values (%)
Gaussian Noise Injection with Weight Clipping	89.3 ($\eta = 8 \times 10^{-4}$)	83.5 ($\eta = 1 \times 10^{-3}$)
Gaussian Noise Injection	89.55 ($\eta = 3 \times 10^{-4}$)	85.6 ($\eta = 1 \times 10^{-3}$)
Standard Training	89.0	82.0

This comprehensive comparison yields several critical insights:

1. **Architecture-Dependent Baseline Robustness:** MLPs demonstrate substantial inherent robustness to photonic noise (89.0% accuracy without noise-aware training), while RNNs show moderate baseline performance, i.e. 82.0%. This 7.0 percentage point difference highlights how recurrent connections introduce additional noise sensitivity through temporal dependencies, though both architectures maintain reasonable baseline performance for practical deployment.

2. **Training Method Effectiveness:** For MLPs, both methods yield similar peak performance (89.3% vs. 89.55%). For RNNs, GI consistently outperforms WC, achieving a peak accuracy advantage of 2.1 percentage points (85.6% vs. 83.5%) and demonstrating 3.6% and 1.5% improvements respectively over the baseline (82.0%).

3. **Optimal Noise Parameters:** The optimal noise levels differ between architectures. For MLPs, GI performs best at lower noise levels ($\eta = 3 \times 10^{-4}$) while WC requires higher levels ($\eta = 8 \times 10^{-4}$). For RNNs, both methods perform best at moderate noise levels ($\eta = 1 \times 10^{-3}$).
4. **Performance-Noise Relationship with Confidence Analysis:** As shown in Figure 4.3 and Figure 4.4 , the relationship between noise parameters and performance is rarely monotonic, especially for RNNs.

4.1.6 Domain-Specific Training Considerations

Our results indicate that training techniques that are effective for electronic ANNs may not be optimal for photonic implementations. Specifically:

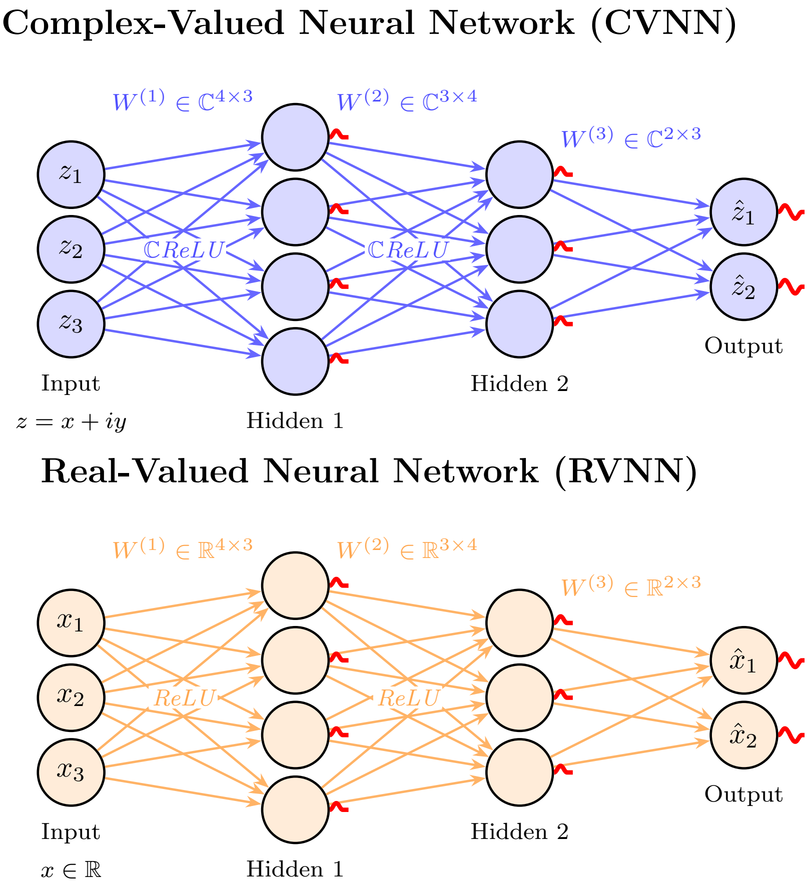
1. For MLPs in photonic domains, standard Gaussian noise injection without weight clipping achieves the best results, particularly at lower noise levels ($\eta = 3 \times 10^{-4}$).
2. For RNNs, which are inherently more vulnerable to photonic noise, both training methods provide substantial benefits, but GI consistently outperforms WC across all noise levels.
3. The optimal noise parameter varies significantly between architectures, with MLPs generally benefiting from more nuanced noise injection compared to RNNs.

These findings highlight the need for domain-specific noise-aware training strategies that account for both the unique characteristics of photonic hardware and the specific neural network architecture being deployed.

In conclusion, we demonstrate that noise injection is a promising technique that significantly enhances the resilience of photonic analogue ANNs to hardware noise. The effectiveness depends critically on both the architecture and the specific noise-injection approach, with standard Gaussian injection generally outperforming weight clipping for photonic implementations. Further optimisation is required to develop tailored approaches for different network architectures and hardware platforms, particularly as applications scale to more complex tasks.

4.2 On the Noise Robustness of Analogue Complex-Valued Neural Networks

In the previous sections, we discussed how the analogue nature of photonic hardware introduces inherent noise sources that can significantly degrade inference performance, and how NAWT strategies can mitigate these effects. One of the most exciting developments in optical neural networks (ONNs), however, is the ability to natively process *complex-valued* data. By leverageing both the amplitude and phase of optical signals, ONNs can naturally implement complex-valued arithmetic, opening the door to complex-valued optical neural networks [123,186,187]. This section evaluates the extent to which these CVNNs perform compared to conventional RVNNs under noisy analogue conditions.


Complex-Valued Neural Network (CVNN)

Real-Valued Neural Network (RVNN)

Figure 4.5: The CVNN processes inputs as $z = x + iy$ in $\mathbb{C}$, employing complex weights $W^{(\ell)} \in \mathbb{C}^{m\times n}$ and complex activations (CReLU), while the RVNN operates on real inputs $x \in \mathbb{R}$ with real-valued weights and ReLU activations. This diagram highlights the structural and representational differences between CVNN and RVNN architectures, and also how the activation pathways are differently affected by noise. In RVNNs, additive noise perturbs scalar activations directly, often leading to saturation or signal degradation. In contrast, CVNNs distribute perturbations across amplitude and phase components, allowing partial cancellation and greater tolerance to stochastic distortions. This feature underpins their robustness in noisy analogue and photonic settings, forming the basis for the subsequent noise-resilience analysis.

As most analogue ONNs rely on key optical components (e.g., Mach–Zehnder Interferometers, phase shifters) that intrinsically operate on complex fields. In a CVNN, each neuron's weight matrix $\mathbf{W}_l$ and bias $\mathbf{b}_l$ belong to $\mathbb{C}^{n_{l+1}\times n_l}$ and $\mathbb{C}^{n_{l+1}}$, respectively. Thus, for a layer l:

$$\mathbf{Z}_{l+1} \;=\; \mathbf{W}_l\,\mathbf{Z}_l \;+\; \mathbf{b}_l, \quad \mathbf{Z}_l \in \mathbb{C}^{n_l}. \tag{4.15}$$

Each complex weight $W_{ij} = |W_{ij}|\, e^{i\,\theta_{ij}}$ can be equivalently represented by real and imaginary parts, $\mathbf{W}_l = \mathbf{W}_l^R + i\,\mathbf{W}_l^I$. Nonlinearities are also extended to the complex domain, such as the Cartesian ReLU (ℂReLU), which applies

$$\mathrm{ReLU}\big(\mathrm{Re}(z)\big) + i\,\mathrm{ReLU}\big(\mathrm{Im}(z)\big), \quad z \in \mathbb{C}. \tag{4.16}$$

With complex-valued arithmetic fully supported by integrated photonics, the MAC operations in each layer can remain purely optical, making CVNNs especially relevant for ONNs [123].

As analogue photonic computing is susceptible to hardware imperfections (e.g., fabrication tolerances in Mach–Zehnder arrays, phase drifting, thermal or shot noise), we adopt an additive noise model at the activation (output) of each neuron:

$$\widetilde{\mathbf{Z}}_{l+1} \;=\; \mathbf{Z}_{l+1} \;+\; \boldsymbol{\eta}_l, \quad \boldsymbol{\eta}_l \sim \mathcal{N}(0, \sigma_{\mathrm{inf}}).$$

Here σ_{inf} characterises the noise level during inference, and is chosen to match the experimentally validated range for integrated photonic circuits [123, 188].

To train these models, we use standard backpropagation but adapt it for complex weights, updating real and imaginary parts:

$$W_{ij}^R \leftarrow W_{ij}^R - \eta\,\frac{\partial \mathcal{L}}{\partial W_{ij}^R}, \quad W_{ij}^I \leftarrow W_{ij}^I - \eta\,\frac{\partial \mathcal{L}}{\partial W_{ij}^I}. \tag{4.17}$$

Following the NAWT principle, we inject noise into the weights *during the forward pass* in training:

$$\mathbf{W}_l \;\leftarrow\; \mathbf{W}_l \;+\; \Delta\mathbf{W}_l, \quad \Delta\mathbf{W}_l \sim \mathcal{N}\big(0, \sigma_{\mathrm{inj}}\big),$$

thereby encourageing the model to learn weight configurations that are more robust to noise during inference. Since batch normalisation is difficult to implement in photonic hardware, it is omitted, making the tested networks more faithful to the actual physical setup [73, 123].

We benchmark CVNNs and compare them with analogous RVNNs on the CIFAR-10 and CIFAR-100 image-classification tasks. In all experiments, each network has $\sim 15\,000$ trainable parameters, and the primary difference is that the CVNN encodes weights in complex form and employs ℂReLU activations. Table 4.2 summarizes the results of our experiments.

Our experimental results, as presented in Table 4.2 and Figure 4.6, reveal several observations regarding the noise robustness of complex-valued neural networks:

- **Modest Intrinsic Robustness Advantage in CVNNs.** With standard training (no noise injection), complex-valued models demonstrate a consistent but modest advantage under noisy inference conditions. For CIFAR-10, the CVNN shows only a 2.9 percentage point improvement over RVNN at the highest noise level (76.9% vs. 74.0%). For the more challenging CIFAR-100 task, this gap widens to 8.1 percentage points (70.8% vs. 62.7%). These results suggest that phase-based representations offer some inherent noise resilience, particularly for more complex tasks, though the advantage for simpler tasks is less pronounced.

- **NAWT Benefits Both Architectures.** Injecting noise during training improves the robustness of both network types. For CIFAR-10, the CVNN with optimal noise injection achieves 78.7% accuracy at high noise ($\sigma_{\text{inf}} = 0.6$), while the RVNN reaches 75.3% – a difference of 3.4 percentage points. For CIFAR-100, the gap is more substantial (72.3% vs. 65.6%, a difference of 6.7 percentage points). This indicates that NAWT can enhance the already present advantage of complex-valued representations.
- **Optimal Noise Injection Levels.** Our experiments suggest that CVNNs may require different noise injection parameters than RVNNs for optimal results. For CVNNs, moderate noise injection ($\sigma_{\text{inj}} = 10^{-3}$ or 10^{-2}) was sufficient to achieve good noise resilience, while over-aggressive noise injection could hinder learning. This suggests that noise-aware training should be tuned specifically for the architecture type.

Table 4.2: Comparison of RVNN and CVNN network accuracy (%) under different noise conditions. “No Injection” refers to standard training ($\sigma_{\text{inj}} = 0$); “NAWT” denotes noise-aware training using moderate σ_{inj} (e.g., 10^{-3} or 10^{-2}).

Dataset	**Method**	$\boldsymbol{\sigma_{\text{inf}} = 0.0}$ (Baseline)		$\boldsymbol{\sigma_{\text{inf}} = 0.6}$	
		RVNN	**CVNN**	**RVNN**	**CVNN**
CIFAR-10	No Injection	92.5	92.8	74.0	76.9
	NAWT ($\sigma_{\text{inj}} = 10^{-3}, 10^{-2}$)	74.0–75.3	78.7–77.9	74.0–75.3	78.7–77.9
CIFAR-100	No Injection	90.8	90.6	62.7	70.8
	NAWT ($\sigma_{\text{inj}} = 10^{-3}, 10^{-2}$)	63.4–65.6	72.3–71.0	63.4–65.6	72.3–71.0

Our findings indicate that complex-valued neural networks, with their ability to utilise both amplitude and phase information, represent a promising direction for analogue optical implementations. When combined with appropriate noise-aware training techniques, they can offer improved resilience to the noise inherent in analogue hardware. Future work should explore whether these advantages extend to other architectures, tasks, and noise regimes, and investigate optimisation strategies specific to complex-valued models to further enhance their noise robustness.

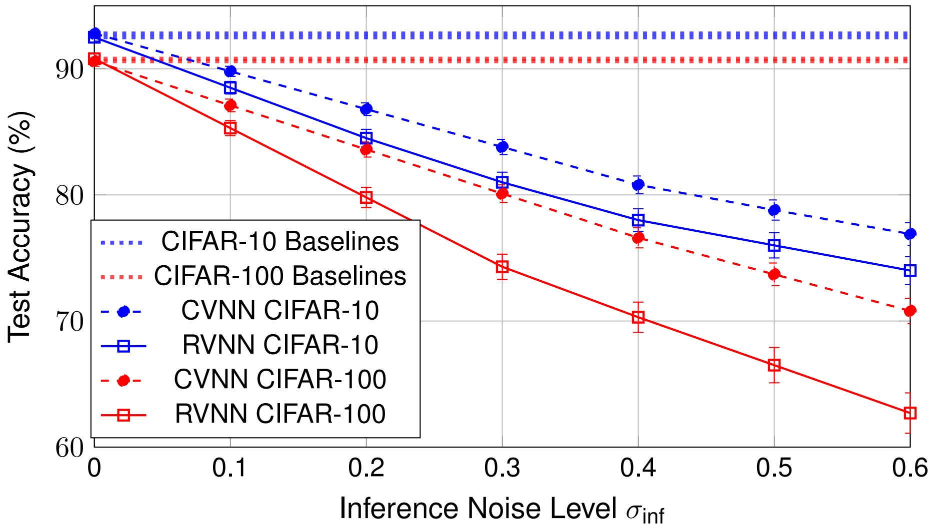


Figure 4.6: Performance degradation comparison between CVNN and RVNN models under increasing inference noise (σ_{inf}) for CIFAR-10 and CIFAR-100 classification tasks. The key findings are: (a) Zero-noise baselines (dotted horizontal lines) show CVNNs achieve 92.8% vs 92.5% (CIFAR-10) and 90.6% vs 90.8% (CIFAR-100) compared to RVNNs. (b) CVNNs demonstrate superior robustness, maintaining 76.9% accuracy at $\sigma_{\text{inf}} = 0.6$ vs 74.0% for RVNNs on CIFAR-10, with more pronounced advantages on CIFAR-100 (70.8% vs 62.7%). (c) Performance advantage of CVNNs increases with task complexity, suggesting greater benefits for challenging classification problems.

4.3 Harnessing Noise with Stochastic Resonance Neurons

Up to this point we have looked at ways to tame the noise and device drift that plague analogue neuromorphic chips, including noise-aware training techniques (Section 4.1) and complex-valued activations (Section 4.2). Those methods all try to suppress noise, yet an intriguing counter-proposal is to exploit the noise itself and let it boost the network. This section explores this counterintuitive approach through stochastic resonance (SR) neurons—a novel activation mechanism that deliberately incorporates noise into its computational dynamics.

4.3.1 Principles of Stochastic Resonance in Neural Systems

Stochastic resonance describes the counter-intuitive case where sprinkling in just the right amount of noise actually makes a faint signal easier to pick out [189]. Classic ANNs usually treat noise as the enemy. Yet real neurons fire amid constant biological hiss and even seem to *use* that background chatter—something our silicon cousins rarely copy.

An SR unit is not a static non-linearity like ReLU; it is a tiny dynamical system whose state $\xi(t)$ follows a noise-driven ODE inside a double-well potential [190]. This dynamic behaviour is mathematically formulated as:

$$\frac{d\xi(t)}{dt} = -\frac{dU_0(\xi)}{d\xi} + s(t) + \sigma\,N(t), \tag{4.18}$$

where

$$U_0(\xi) = -\alpha\frac{\xi^2}{2} + \beta\frac{\xi^4}{4}, \quad \alpha, \beta > 0, \tag{4.19}$$

$s(t)$ represents the weighted input to the neuron, $N(t)$ denotes Gaussian noise, and σ is the noise amplitude. The double-well makes two preferred states, so each SR neuron carries its own micro-memory. Because gentle noise can nudge the state over the barrier, the cell picks up whispers it would otherwise miss—that's the essence of SR [2].

4.3.2 Implementation in Recurrent Neural Networks

To test the idea we dropped SR nodes into an Echo State Network (ESN)—a lightweight recurrent model often used for time-series work. In the standard ESN formulation, the network's internal state $\mathbf{x}_{n+1}$ evolves according to [2]:

$$\mathbf{x}_{n+1} = f(\mathbf{W}_{in}\mathbf{u}_{n+1} + \mathbf{W}\mathbf{x}_n + \mathbf{W}_{back}\mathbf{y}_{n+1}) \tag{4.20}$$

where $\mathbf{u}$ is the input, $\mathbf{x}$ is the reservoir state, $\mathbf{y}$ is the output, and $\mathbf{W}_{in}$, $\mathbf{W}$, and $\mathbf{W}_{back}$ are the input, reservoir, and feedback weight matrices, respectively.

In the SR version, we simply swap the usual tanh for a short Runge–Kutta step of the SR ODE. At each time step, each SR neuron's state is updated by integrating Equation (4.18) using a second-order Runge-Kutta method [2]:

$$\xi_{n+1} = \xi_n + \Big[\alpha(\xi_n - \xi_n^3) \;+\; \sigma N(t) \;+\; s_n\Big]\Delta t, \tag{4.21}$$

where ξ_n is the neuron's previous state, s_n is the weighted input, and Δt is the integration time step.

This implementation creates neurons with several distinguishing properties:

1. **Internal memory**: Every SR cell keeps a private state variable, adding memory on top of the reservoir feedback.
2. **Adaptive transfer functions**: During training the transfer curve of each SR unit morphs to fit the data.
3. **Inherent noise utilisation**: Because noise is baked into the update rule, device jitter may turn from headache into helper.

4.3.3 Performance Evaluation on Time-Series Prediction

We conducted extensive experiments comparing the SR-ESN approach against traditional sigmoid-based ESNs on the Mackey-Glass (MG) chaotic time series prediction task, as it provides clear evidence of the SR neurons' advantages.

The main result is that SR-based ESN reaches the same accuracy with a fraction of the nodes a tanh ESN needs. Table 4.3 presents a comprehensive comparison of SR-ESN and Sigmoid-ESN configurations with varying numbers of neurons, showing both prediction accuracy (as measured by Mean Squared Error, MSE) and computational complexity.

Table 4.3: Comprehensive comparison between SR-ESN and Sigmoid-ESN on Mackey-Glass chaotic time-series prediction demonstrating computational efficiency advantages of noise-enhanced processing. Performance measured by MSE on 2000-sample test set with 1-step prediction horizon. Relative efficiency = ($\text{MSE}_{\text{baseline}}$ / $\text{MSE}_{\text{method}}$) / (Computational $\text{Cost}_{\text{method}}$ / Computational $\text{Cost}_{\text{baseline}}$), normalised to 100-neuron SR-ESN. SR-ESN achieves 19× better accuracy at equal network size (200 neurons), while maintaining competitive computational complexity.Efficiency advantage increases with network size, suggesting superior performance scaling for larger implementations. All networks use reservoir computing framework with spectral radius 0.95, input scaling 0.1, and ridge regression output training.

Method	#Neurons	MSE	Multiplications/step	Relative Efficiency
SR-ESN	100	1.8×10^{-2}	1.04×10^{4}	1.0×
Sigmoid-ESN	100	3.5×10^{-2}	1.02×10^{4}	0.51×
SR-ESN	200	1.7×10^{-3}	4.08×10^{4}	3.9×
Sigmoid-ESN	200	3.2×10^{-2}	4.06×10^{4}	0.54×
SR-ESN	300	5.5×10^{-4}	9.12×10^{4}	12.0×
Sigmoid-ESN	300	2.8×10^{-2}	9.09×10^{4}	0.61×
Sigmoid-ESN	450	1.7×10^{-3}	2.03×10^{5}	0.87×

The relative efficiency metric in Table 4.3 represents the ratio of accuracy to computational cost, normalised to the SR-ESN with 100 neurons. The data reveals several important insights:

- For example, 200 SR units match the error of a 450-node tanh reservoir, cutting the model size by roughly 56% while requiring only about 20% of the computational resources.
- At equal network sizes, SR-ESNs consistently outperform their sigmoid counterparts by substantial margins. For example, with 200 neurons, the SR-ESN achieves an MSE approximately 19 times lower than the sigmoid network.
- The efficiency advantage of SR-ESNs increases with network size, suggesting that the benefits of stochastic resonance become more pronounced in larger networks.

When evaluating noise tolerance and optimal noise levels, the core principle behindSR is to dial in just enough noise: too low and SR never kicks in, too high and the signal drowns. We systematically investigated this relationship by varying the internal noise amplitude σ in SR neurons and measuring prediction performance. Table 4.4 summarizes these findings.

The curves form the classic SR bell: performance peaks mid-noise and drops off on either side. The optimal noise amplitude of approximately 10^{-10} provides a 3.2× improvement over noiseless

Table 4.4: Effect of internal SR noise level (σ) on prediction accuracy for a 200-neuron SR-ESN on the Mackey-Glass chaotic time-series prediction task. SR neurons implement double-well potential dynamics (Equation 4.18) with $\alpha = 1.0$, $\beta = 1.0$, integration time step $\Delta t = 0.1$. Optimal performance achieved at $\sigma = 10^{-10}$ with 3.2× improvement over noise-free operation. Modest noise levels enable barrier crossing in double-well potential, enhancing signal detection capabilities. Performance degrades for $\sigma > 10^{-8}$ as noise overwhelms signal, demonstrating critical noise level optimisation. Experimental validation done on MG series with $\tau = 17$, prediction horizon = 1 step, training length = 8000 samples, testing length = 2000 samples.

SR Noise Level (σ)	MSE	Relative Improvement
0 (No noise)	5.4×10^{-3}	1.0×
10^{-12}	3.2×10^{-3}	1.7×
10^{-10}	1.7×10^{-3}	3.2×
10^{-8}	2.1×10^{-3}	2.6×
10^{-6}	9.7×10^{-3}	0.6×
10^{-4}	3.6×10^{-2}	0.15×

SR neurons, underscoring that noise is indeed being utilised constructively rather than merely tolerated.

For hardware, the real test is how they behave when the inputs are noisy—a common scenario in real-world sensing and signal processing applications. To investigate this, we trained both SR-ESN and Sigmoid-ESN models on Mackey-Glass data with varying levels of additive Gaussian noise, characterised by the SNR.

Table 4.5: Prediction accuracy comparison for 200-neuron networks trained on noisy Mackey-Glass data with varying SNR levels. The SR-ESN maintains superior performance across all noise conditions, with its advantage becoming more pronounced at lower SNR [2].

Training Data SNR	SR-ESN MSE	Sigmoid-ESN MSE	Improvement Factor
Clean	1.7×10^{-3}	3.2×10^{-2}	18.8×
40 dB	2.8×10^{-3}	4.1×10^{-2}	14.6×
30 dB	5.7×10^{-3}	5.9×10^{-2}	10.4×
20 dB	1.0×10^{-2}	2.4×10^{-1}	24.0×

Fascinatingly, SR reservoirs pull even further ahead as the training SNR gets worse. The results in Table 4.5 reveal a remarkable property of SR-ESNs: they not only maintain their performance advantage across all noise conditions but actually show the greatest relative improvement in the most challenging scenario (20 dB SNR). This suggests that stochastic resonance neurons possess an inherent robustness to external noise that conventional activation functions lack.

4.3.4 Self-Adapting Activation Functions

A particularly intriguing aspect of SR neurons is their capacity to develop individualised transfer functions during training. Whereas a tanh is frozen, an SR node's transfer curve keeps reshaping itself in response to the data stream.

That adaptability comes from the neuron's stateful ODE: today's output depends on yesterday's state as well as today's stimulus. The stochastic differential equation governing each neuron essentially forms a nonlinear filter that adapts to the statistical properties of the input signal and noise.

We observed that in trained SR-ESNs, neurons within the same layer developed diverse transfer functions, effectively creating specialised feature detectors tailored to different aspects of the input time series. This automatic specialisation may partially explain the substantial performance benefits observed with SR neurons, as it enables a form of adaptive ensemble learning within a single network. This is represented in Figure 4.7.

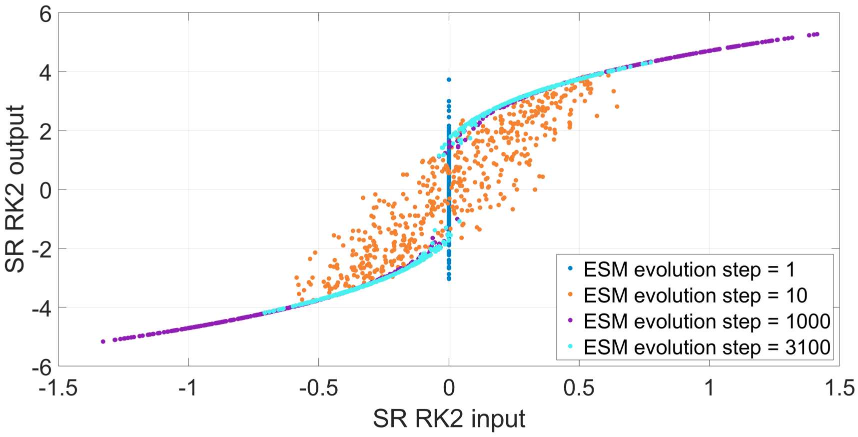


Figure 4.7: Demonstration of how the transfer functions of an SR neuron depends on the number of the ESN evolution steps, i.e. 1, 10, 1000 and 3100 are represented, showing how each neuron automatically optimises its transfer function during the learning stage [2].

4.3.5 Implications for Neuromorphic Hardware

These results matter for analogue neuromorphic chips: such platforms inevitably experience various forms of noise and variability—thermal fluctuations, fabrication inconsistencies, and phase drifts in optical components, among others [191, 192]. Traditional designs burn power trying to cancel every flicker and drift.

SR nodes flip the script—by *welcoming* noise, they may turn fabrication quirks into free computation. The SR-ESN's demonstrated ability to maintain high performance even with noisy inputs further suggests that neuromorphic implementations using this approach could be intrinsically robust to sensor noise and environmental variability [193, 194]. This aligns conceptually with noise-assisted mechanisms in biological systems and with regularisation effects observed in deep learning, where stochastic perturbations (e.g., dropout or additive noise) improve generalisation [195, 196].

Moreover, fewer neurons mean smaller, cooler chips. For example, the 200-neuron SR-ESN that matched the performance of a 450-neuron sigmoid network would require approximately 56% fewer physical components and consume proportionally less power in an analogue implementation. This reduction directly translates to lower silicon area and thermal footprint, enabling higher integration density in neuromorphic processors such as TrueNorth and Loihi [197, 198].

These advantages position SR neurons as a promising direction for noise-tolerant, energy-efficient neuromorphic computing. Rather than expending resources fighting against inherent physical noise and variability, SR-based systems harness these phenomena as computational resources, potentially leading to more robust and efficient neural hardware [199–201]. The paradigm is consistent with emerging trends in programmable photonics and neuromorphic VLSI, where variability and reconfigurability are treated not as defects but as enablers of adaptive computation.

4.3.6 Conclusions and Future Directions

In short, the study shows that—with the right setup—adding a dash of noise can *improve* a network instead of harming it. The SR-ESN approach achieves superior accuracy with dramatically fewer neurons compared to conventional methods, particularly in the presence of noisy training data [193, 194].

Moreover, several promising research directions emerge:

- **Broader architectures** – explore SR nodes in CNNs, transformers, or feed-forward networks, extending stochastic resonance principles to modern deep learning architectures.
- **Hardware builds** – fabricate SR neurons in electronic or photonic silicon platforms to test their operation under real-world noise and variability [199, 200].
- **Neuroscience links** – investigate whether additional noise-exploiting mechanisms exist in biological systems, potentially informing neuromorphic circuit design [191].

By embracing rather than avoiding noise, SR neurons represent a fundamental shift in our approach to neural computation—one that may prove especially valuable as we push toward more energy-efficient, brain-inspired computing systems capable of operating reliably in noisy, uncertain environments.

4.4 Crowd Equalisation and Noise-Resistant Techniques

As optical links become ever more sophisticated, keeping up with real-time equalisation is getting harder. Neural networks can tackle nonlinear impairments convincingly, yet their hefty compute budget and noise-susceptibility make hardware deployment tough—particularly when resources are tight [68]. We tackle this bottleneck with a fresh idea called crowd equalisation: many small networks pool their opinions, giving the system extra noise immunity and sturdiness.

The Crowd Equalisation Concept is derived from the "wisdom of crowds," we train several distinct networks as solo experts, each bringing its own architecture and noise quirks. Blending their outputs smooths away each model's errors and noise, so performance climbs. The scheme cuts back on expensive, high-precision hardware per model and gains extra robustness from the architectural variety across the crowd [202, 203].

Crowd equalisation is taken into practice by training every member on its own, giving them different shapes and initial weights so that each one latches onto a different slice of the signal dynamics. Once training ends, we compute the equalised symbol by taking a weighted mean of all network outputs. Networks that post the lowest BER on a held-out set earn the biggest weights, as reflected in Equation (4.22).

$$y = \frac{1}{\sum_{i \in \mathcal{C}} BER_i\left(\mathbf{x}^{\text{test}}\right)} \sum_{i \in \mathcal{C}} \frac{f_i\left(\mathbf{x}^{\text{crowd test}}\right)}{BER_i\left(\mathbf{x}^{\text{test}}\right)} \tag{4.22}$$

Here, i denotes an individual in the crowd $\mathcal{C}$, and $f_i(\mathbf{x})$ represents the output of the i-th neural network. $BER_i(\mathbf{x})$ is the corresponding bit error rate for each network on the test data.

Our data come from a 400 km dual-polarisation 28-GBaud 64-QAM experiment (identical to Section 3.17) and include both transmitted and received symbols. We split the data 70

To judge the crowd's merit, we pit various multi-network line-ups against one big model that eats up the same total compute. Example: one monolithic $[600, 600, 600, 600]$ network racks up roughly 4.3 million complex multiplies. Instead, we try crowds—e.g., a pair of nets sized $[900, 476, 223]$ and $[732, 732]$, plus line-ups of three or four models. Every arrangement keeps the overall compute budget unchanged but spreads the workload among the members.

A major plus is that optical hardware can run all these small nets side-by-side with almost no overhead. Because light can sum signals natively, several nets can fire at once, giving swift processing and built-in noise averageing. The gain is most apparent in noisy photonic setups, where multiplies and activations nibble away at SNR [38].

Figure 4.8 confirms that the crowd beats a lone large network, and the gap widens as SNR drops. Bigger crowds shrug off noise more effectively, pushing BER down over the whole SNR sweep. A four-member crowd tops the chart, dipping below the HD-FEC limit once SNR is high enough.

Sharing the work among small nets keeps the total CC from rising (see Equation 4.23), yet grants parallel speed-ups and extra noise headroom.

$$CC = \left(n_s n_i n_1 + \sum_{l=1}^{L-1} n_l n_{l+1} + n_o n_L\right) \tag{4.23}$$

where n_s is the size of the input sequence (equal to $2M+1$ where M is the memory depth), n_i is the number of input features (e.g., 4 as the real and imaginary parts of two polarisation components

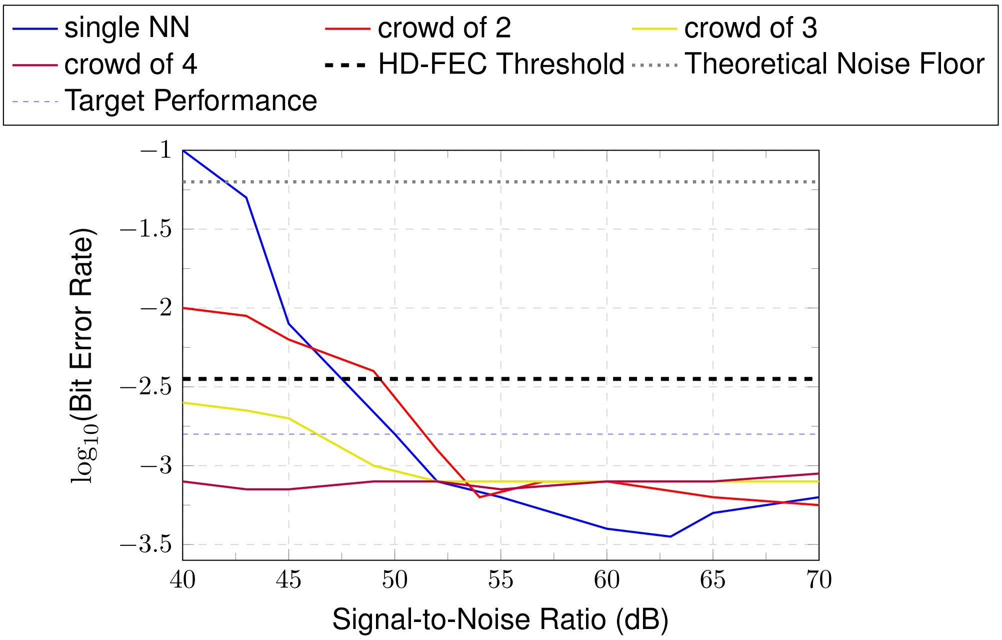


Figure 4.8: BER performance comparison between single ANN and crowds of 2, 3, and 4 ANNs across SNR range 40-70 dB for 400 km dual-polarisation 28-GBaud 64-QAM optical transmission. A single network achieves BER = $10^{-3.45}$ at 63 dB SNR, but degrades significantly at low SNR with BER = $10^{-1.0}$ at 40 dB. 4-member crowd achieves best performance with BER = $10^{-3.15}$ consistently across 40-70 dB range, demonstrating superior noise tolerance. All crowds outperform HD-FEC threshold (BER = $10^{-2.45}$) across the entire SNR range, while single network fails below 45 dB. Total computational complexity remains constant across all configurations via Equation 4.23, enabling parallel optical implementation. Crowd averageing provides inherent noise immunity through ensemble decision making, with performance improvement scaling with crowd size. Individual networks employ different sizes and configurations to ensure linear independence and comprehensive feature extraction from optical signal impairments.

are used, corresponding to $d_{in} = 4$), n_o is the number of outputs (e.g., 2 as the real and imaginary parts of each symbol are recovered) and n_l corresponds to the number of neurons in each layer with $l \in [1, L]$

In short, crowd equalisation boosts ANN equaliser performance under noise, keeps compute in check, and doesn't add architectural bloat. Its natural fit with optical parallelism clears a path toward scalable, noise-tolerant neuromorphic hardware for next-gen fibre systems.

Chapter 5

Hybrid SNN-RWKV Architecture for Energy-Efficient Optical Signal Equalisation

Despite the strong performance of ANN equalisers for Kerr nonlinearity mitigation and distortions induced by components, their CC and power draw remain prohibitive for real-time deployment at high symbol rates [68, 103]. This chapter therefore motivates and validates a hybrid approach that combines:

(i) a spiking front end for energy-efficient, event-driven processing, and

(ii) the RWKV time-mixing mechanism for linear-time sequence modelling.

We justify RWKV as an alternative to standard Transformers for optical equalisation and combine it with an SNN to form a practical SNN–RWKV equaliser. RWKV replaces the quadratic, global softmax attention with an exponential time-mix that yields constant memory and $O(N)$ compute in sequence length during inference—critical for streaming receivers [204, 205]. The SNN brings sparse, binary spike processing that lowers Multiply-Accumulate Operation (MAC) activity and suits neuromorphic deployment. The core of this chapter is the model and results reported in our European Conference on Optical Communication (ECOC) study on a DP-16QAM, 500 km TWC link; all claims and settings here follow that work.

5.1 Advanced Architectures for Efficient Optical Channel Equalisation

Both the persistent challenge of nonlinear impairments in optical communication systems and the latest advances in the context of AI have led to the development of innovative signal processing techniques. These new equalisation methods are characterised by advanced ANN architectures that leverage neuromorphic computing and novel machine learning paradigms. This section explores cutting-edge approaches that promise more efficient and robust optical channel equalisation.

Building upon the foundation laid in earlier chapters on nonlinear impairments, this section provides the theoretical foundation for understanding the individual components that will be integrated into the hybrid SNN-RWKV architecture. We examine three key architectural paradigms:

1. **SNNs:** Event-driven computation with sparse activations, enabling energy-efficient inference and hardware-friendly (neuromorphic) deployment; trained via surrogate gradients to handle the non-differentiable spike function [206, 207].
2. **Transformer-based models:** Strong at capturing long-range dependencies via self-attention, but with quadratic time/memory in window length; streaming with key–value (KV) caching is possible yet introduces per-step cost and memory that grow with context length [204, 208].
3. **RWKV:** A time-mixing mechanism that merges RNN-like streaming with attention-like expressiveness. Crucially, *inference* is linear in sequence length with *constant* memory (per-channel accumulators), matching streaming receiver constraints [205].

In this work, the ultimate goal is not to use these architectures independently, but to synthesise their complementary strengths into a unified hybrid equaliser. The SNN component provides the computational substrate and temporal dynamics, while the RWKV mechanism supplies the attention-based long-range dependency modelling. This integration creates a neuromorphic equaliser that maintains the energy efficiency of spiking computation while incorporating the representational power of modern attention mechanisms.

5.1.1 Spiking Equalisers for Optical Links and the Role of Signal Encoding

SNNs have emerged as low-power nonlinear equalisers for short-reach IM/DD links because their event-driven compute maps naturally to sparse communication signals and can be implemented on mixed-signal neuromorphic hardware. Recent demonstrations include: (i) a hardware SNN equaliser/demapper on BrainScaleS-2 that beats a linear Linear Minimum Mean-Square Estimator (LMMSE) baseline at the HD-FEC threshold with <1 dB hardware penalty, and (ii) experimental SNN *linear* equalisation at 100 Gb/s over 2 km 4-Pulse Amplitude Modulation (PAM). These results indicate that SNNs can deliver competitive BER at much lower activity (hence power) than dense ANN DSP. [209–211]

In IM/DD, the SNN-based equaliser processes a short window of received samples $\mathbf{y}_k$ and outputs symbol logits (or a hard decision). A SNN instantiates this pipeline as: (1) spike encoding of $\mathbf{y}_k$ into a set of sparse event trains; (2) a synaptic "crossbar" (learned weights) producing exponentially decaying currents; (3) one or more layers of LIF neurons that integrate these currents; and (4) a spike decoding stage that turns membrane trajectories or spike counts/timings into logits.

Training is performed end-to-end with surrogate gradients to handle the non-differentiable spiking nonlinearity. [206, 209, 212]

The encoding step is highly relevant as it fixes what information reaches the spiking core, how sparsely it is represented, and how quickly a decision can be made. In equalisation we must (i) preserve amplitude features that carry the PAM symbol, (ii) expose short-range memory to the network (dispersion/ISI), and (iii) keep spike rate and time window T small to save power and latency. Encoders that over-encode (high rates, long T) negate energy advantages; encoders that under-encode discard discriminative information and hurt BER. Recent IM/DD studies show that decoding rules like End-of-Time Membrane (EOTM) or Max-over-Time Membrane (MOTM) can reduce spike activity and decision latency for a given encoder, and that carefully designed encoders markedly improve the BER–sparsity–throughput trade-off. [212, 213]

The SNN equaliser/demapper by Arnold et al. employs a distance/population encoder with a small latency component: each input sample is represented by a population of neurons (ten per tap), where spike times are proportional to the distance between the sample and fixed reference levels. Using 7 input taps, this yields 70 input neurons; each active neuron fires once within a bounded time window T, producing a sparse, ordered set of spikes. Hidden neurons are LIFs, while the readout stage uses leaky integrator neurons; the decision is taken by the readout neuron with the largest membrane value over time (winner-takes-first/most). Figure 5.1 summarises the mechanism: (A) distance measurement to reference levels; (B) conversion to spike times across the population; (C) membrane responses; and (D) winner selection. This encoder preserves amplitude information, exposes local temporal context through synaptic/membrane filtering, and maintains tight control over spike count and decision latency. [209, 211]

While Arnold et al. used a winner-selection over readout membranes, follow-up work compared rate, Time-To-First-Spike (TTFS), MOTM, and EOTM decoders with the same encoder and found EOTM/MOTM achieve the best BER at substantially lower spike rates and higher throughput (e.g., $\sim 4\times$ faster decisions and up to $10\times$ fewer hidden spikes versus non-optimised MOTM), underscoring that encoding and decoding must be co-designed for energy-efficient equalisation. [212]

In this chapter we retain the design rules that made SNNs effective in IM/DD—sparse, amplitude-faithful encoding; short T; and leaky-integrator/LIF smoothing before the sequence model—and later adapt the encoder for coherent, dual-polarisation streams when introducing the SNN–RWKV equaliser (Section 5.2). Here, our goal was to explain how SNN equalisers operate in optical receivers and why the encoder is decisive for the BER–energy–latency trade-off; details of the hybrid encoder used in this work follow in Section 5.2.

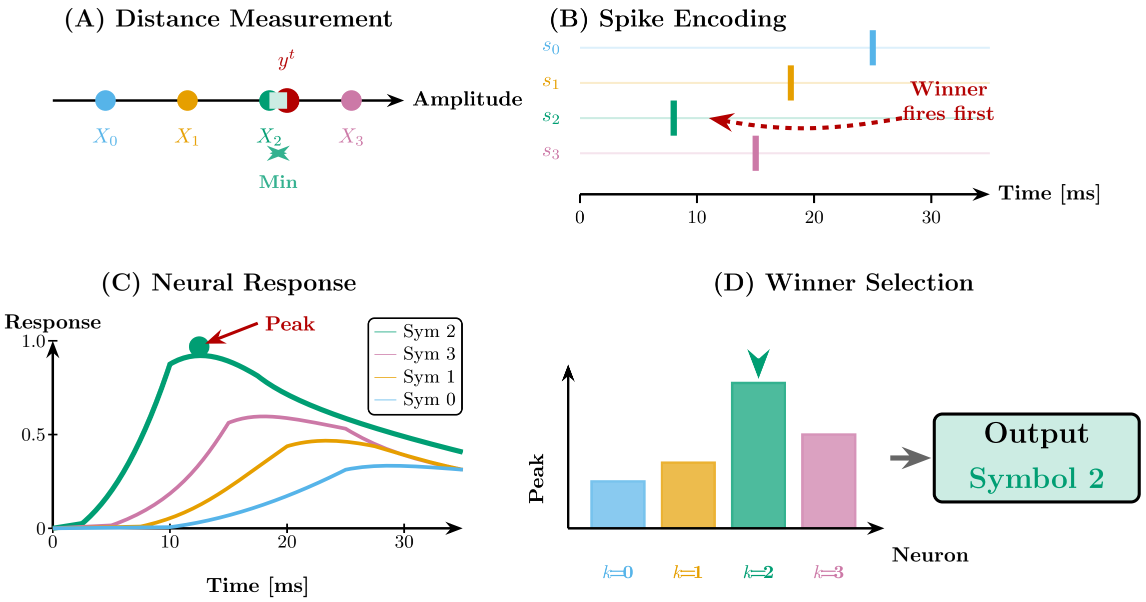


Figure 5.1: Arnold-style distance/population encoder and winner-selection readout used in IM/DD SNN equalisers.

5.1.2 Transformer Equalisers for Coherent Optical Links and a Preview of RWKV

Transformers model long-range dependencies through multi-head self-attention. Given query, key, value matrices (Q, K, V), scaled dot-product attention computes

$$\mathrm{Attn}(Q, K, V) = \mathrm{softmax}\left(\frac{QK^\top}{\sqrt{d_k}}\right) V, \tag{5.1}$$

typically stacked with residual connections, layer normalisation, and position encodings; multiple heads attend to different interaction patterns in parallel [204]. This global attention confers strong sequence modelling power but introduces quadratic time/memory in window length.

When adapting Transformers to coherent receivers for optical equalisation, two practical patterns have emerged:

1. Direct nonlinear equaliser: The encoder (or a few encoder blocks) is applied to a sliding window of complex DP samples/symbols to regress the undistorted constellation (or soft symbols). Tokens are built from the $\Re, \Im$ components of both polarisations, either concatenated per time step or organised as two channels; learnable or sinusoidal position encodings are scaled by R_s/f_s to reflect symbol timing. In DP-16QAM experiments, such Transformer equalisers improve Q-factor and BER over DBP, Fully Connected Neural Network (FCNN) and BiLSTM baselines across launch powers and distances [214]; related ECOC'24 work reports up to $\sim$1.7,dB Q-gains on WDM 16QAM multi-span links [215].

2. Physics-informed or efficiency-oriented variants: (i) Masked/local attention limits attention span based on channel memory or perturbation theory, reducing complexity with little loss [216]. (ii) Block processing with overlap-add amortizes the $O(N^2)$ cost while capping latency. (iii) Coherent Orthogonal Frequency Division Multiplexing (CO-OFDM) replaces SSFM blocks with attention modules that learn dispersion/nonlinearity dynamics and run substantially faster than SSFM or Bi-LSTM surrogates [217]; the same building blocks can be used inside an equaliser.

There are several design details to consider in coherent links:

- **Tokenisation and features:** Build tokens from (I_X, Q_X, I_Y, Q_Y), optionally augmenting with $|s[n]|^2$, instantaneous frequency (Carrier Phase Recovery (CPR) robustness), or pilot flags.
- **Positional encoding:** Use sinusoidal or learned encodings tied to symbol/sampling rates; for oversampling, include fractional-delay features to ease dispersion compensation.
- **Attention span:** Practical windows covering $\mathcal{O}(\text{tens of symbols})$ often suffice for intra-channel Kerr memory at the tested baud rates; narrow masks reduce cost.
- **Losses and targets:** Either regress clean I/Q (L2) or train a soft demapper (cross-entropy); mixed losses can help under ASE.
- **Complexity, latency, and streaming:** Vanilla self-attention over a window of length N is $O(N^2)$ in time and memory [204]. Windowing/masking reduces cost to $O(Nw)$ for a fixed window w. *Causal streaming with KV caching* is feasible, but the KV cache grows with context length, so both *per-step compute and memory increase with time*; this can limit real-time deployments unless the window is bounded [208].

Although employing self-attention, effectively captures complex temporal and nonlinear interactions in coherent optical channels, its quadratic time and memory complexity with respect to window length, limits Transformer's efficiency for real-time equalisation, where updates must be performed sequentially and with minimal latency. For a length-N token sequence $X \in \mathbb{R}^{N \times d}$, the scaled dot-product attention is defined as Equation 5.1

This mechanism enables each token to attend to all others within a processing window, granting strong sequence modelling capabilities. However, its quadratic time and memory complexity with respect to sequence length limits applicability to real-time equalisation tasks, where updates must be performed sequentially and with low latency.

To address this limitation, the RWKV formulation bridges self-attention and recurrent processing by replacing the global softmax with an exponentially decayed recurrent update, referred to as a *time-mix*. Denoting per-symbol key/value vectors $k_t, v_t \in \mathbb{R}^d$ and a learnable decay $w \in \mathbb{R}^d$, RWKV

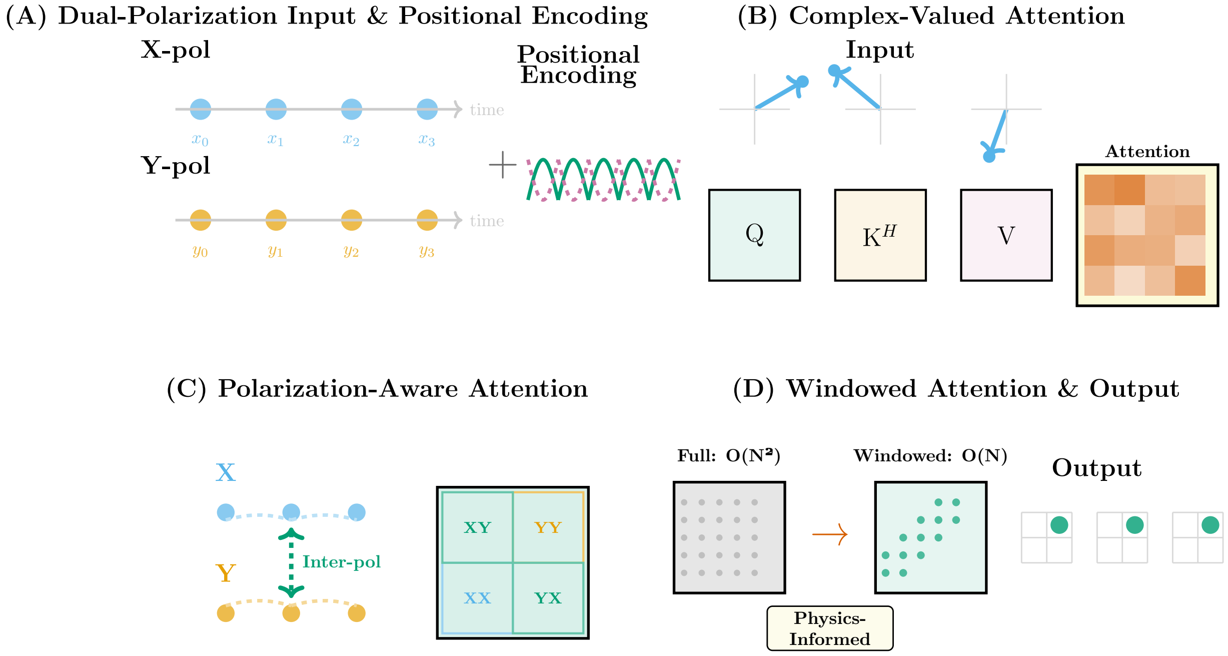


Figure 5.2: Transformer equalisation pipeline for coherent DP links. (A) Dual-polarisation tokenisation with positional encoding. (B) Scaled dot-product attention (shown with K^{H} for complex projections; use $K^{\top}$ in purely real implementations). (C) Polarisation-aware heads (XX, XY, YX, YY) to capture inter-polarisation coupling. (D) Windowed/masked attention: reduces compute from $O(N^2)$ to $O(Nw)$ for fixed window w, controlling latency.

defines running numerator and denominator states as

$$A_t = e^{k_t} \odot v_t + e^{-w} \odot A_{t-1}, \tag{5.2}$$

$$B_t = e^{k_t} + e^{-w} \odot B_{t-1}, \tag{5.3}$$

$$\mathrm{wkv}_t = \frac{A_t}{B_t}, \tag{5.4}$$

$$y_t = \sigma(r_t) \odot \mathrm{wkv}_t \, W_o, \tag{5.5}$$

where r_t is a receptance gate and W_o a linear projection. This formulation produces constant-memory, linear-time updates per token while retaining the expressiveness of attention mechanisms. The decayed accumulation in RWKV can be interpreted as a causal, exponentially weighted analogue of softmax attention, naturally aligning with the memory characteristics of optical channels.

Equations (5.1) and (5.2)–(5.5) thus represent two complementary formulations that will be merged in Section 5.2 (see Figure 5.3), illustrating how RWKV transforms the quadratic global attention into a streaming, neuromorphic-friendly mechanism. By embedding this structure within a spiking framework, it becomes possible to achieve efficient, event-driven equalisation while preserving temporal context.

In essence, RWKV can be viewed as a continuous-time or “streaming” reinterpretation of the

Transformer: rather than attending globally to all past symbols, it integrates contextual information recursively through a decay-weighted mechanism. This shift retains the Transformer's ability to model long-range dependencies while achieving the efficiency and causality of RNN-like architectures. Empirically, the exponentially decayed memory in RWKV closely matches the impulse response of optical fibres, making it an attractive alternative for equalisation tasks where both context and speed are critical.

Figure 5.3 illustrates this conceptual transition—from the global, quadratic self-attention of Transformers to the localised, linear-time decay-based updates in RWKV—highlighting how our hybrid equaliser leverages both paradigms to achieve high fidelity with low computational overhead.

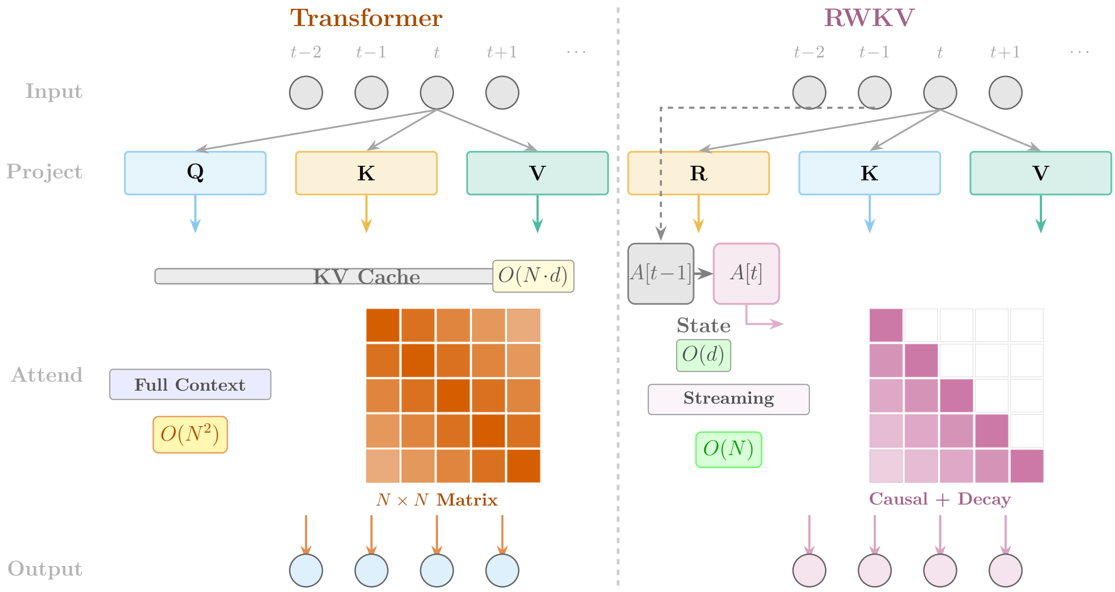


Figure 5.3: Self-attention (left): global, quadratic compute over a window; RWKV time-mix (right): exponentially weighted, sequential update with constant memory. In coherent equalisation we trade global context for streaming efficiency by moving from softmax attention to RWKV's time-mix.

5.2 Hybrid SNN-RWKV Equaliser Design and Optical System Simulation

This section presents the practical implementation of the hybrid SNN-RWKV equaliser, demonstrating how the theoretical foundations established in the previous sections are integrated into a unified neuromorphic architecture for optical channel equalisation. The design combines the energy efficiency of spiking neural networks with the representational power of linear attention mechanisms, creating a novel equaliser specifically optimised for nonlinear optical channel compensation.

To evaluate the proposed hybrid SNN-RWKV optical channel equaliser, a numerical simulation of a DP transmission of a single-channel signal at 34.4GBd, with 16 QAM modulation format, is simulated. The signal is pre-shaped with a RRC filter with 0.1 roll-off at a sampling rate of 8 samples per symbol. In Figure 5.4 it is possible to appreciate the transmission over an optical link of 10×50 km TWC spans. The optical signal propagation along the fibre was simulated by solving the Manakov equation via SSFM, with a resolution of 1 km per step [45]. The parameters of the fibre are: the attenuation parameter $\alpha = 0.23$ dB/km, the dispersion coefficient $D = 2.8$ ps/(nm · km), and the effective nonlinearity coefficient $\gamma = 2.5$ (W · km)$^{-1}$. As depicted in Figure 5.4, after every span, an optical amplifier with the noise figure $NoiseFigure(NF) = 4.5$ dB is employed to fully compensate fibre losses and ASE noise. At the receiver, a standard Rx-DSP is used. It consists of electronic CDC using a time-domain equaliser, the application of a matched filter, and the downsampling to the symbol rate. Finally, the received symbols were normalised (by phase and amplitude) to the transmitted ones. No other transceiver distortions were considered [119].

The training dataset consists of 2^{18} pairs of complex-valued DP samples. The test dataset consists of 2^{20} independently generated symbols. Both of them were created using the Mersenne twister generator to guarantee symbol independence [218].

5.2.1 Hybrid SNN-RWKV Equaliser Architecture and Implementation

The hybrid SNN-RWKV equaliser architecture, illustrated in Figure 5.4, represents a novel integration of neuromorphic computing principles with linear attention mechanisms for optical channel equalisation. The architecture consists of several key components working in synergy:

Architecture Components:

- **1-D CNN Preprocessing**: Initial symbol embeddings from continuous optical signals
- **Binary Embedding (BE) Layer**: Interface converting continuous embeddings to binary spikes
- **Spiking Receptance Weighted Key Value (SRWKV)**: Hybrid component combining LIF neuron dynamics with linear attention
- **Spiking Receptance Feed-Forward Network (SRFFN)**: Channel mixing and nonlinear processing
- **Residual Connections**: Enabling gradient flow and information preservation

The hybrid nature of this architecture is evident in how each component contributes to the overall functionality: the CNN provides continuous signal preprocessing, the BE layer implements the continuous-to-spike conversion, the SRWKV blocks perform attention-modulated spiking computation, and the SRFFN units provide nonlinear channel mixing. These units operate in a loop with residual connections to predict nonlinear distortion components $\hat{E}_{XI}$ and $\hat{E}_{XQ}$ [219, 220].

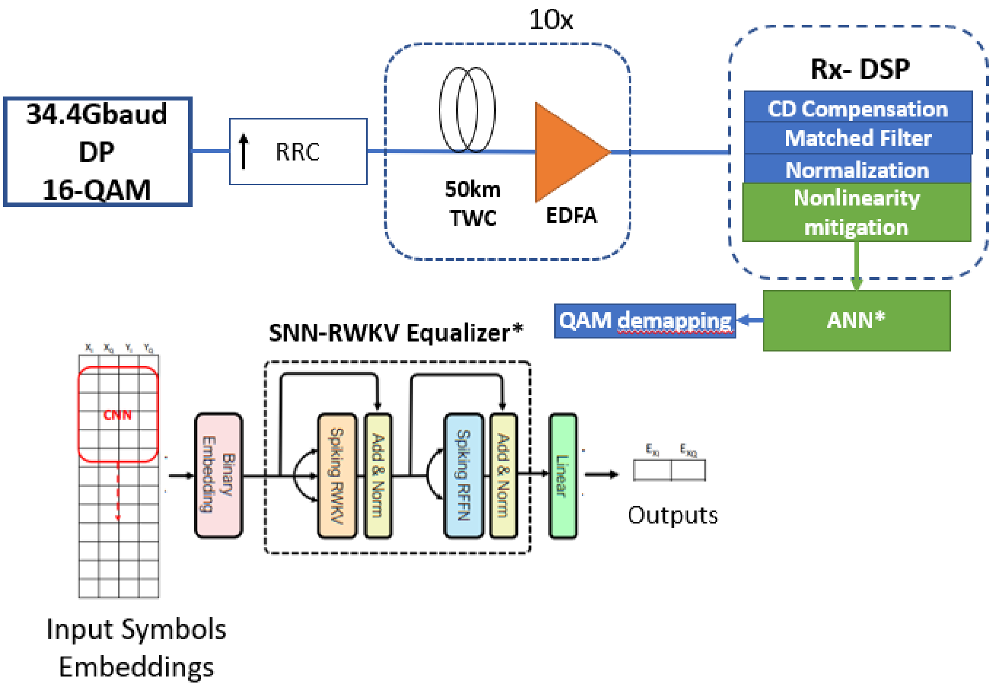


Figure 5.4: Scheme of the numerical setup and the place of the ANN-based equalisers in the transmission link of 10 spans of 50 km TWC fibre. Including the detailed description of the blocks that constitute the SNN-RWKV equaliser proposed in this paper.

The processing pipeline of the SNN-RWKV equaliser follows a sequence of transformations that can be formulated as:

$$X_0 = BE(E), \quad X_0 \in \mathbb{R}^{E_n \times n_f} \tag{5.6}$$

$$X_1' = SRWKV(X_0) + X_0, \quad X_1' \in \mathbb{R}^{E_n \times n_f} \tag{5.7}$$

$$X_1 = SRFFN(X_1') + X_1', \quad X_1 \in \mathbb{R}^{E_n \times n_f} \tag{5.8}$$

$$Y = Linear(X_1) = [\hat{E}_{XI}, \hat{E}_{XQ}] \tag{5.9}$$

where E represents the input symbol embeddings, $BE(\cdot)$ is the Binary Embedding operation, X_0 is the binarised representation of the input, $SRWKV(\cdot)$ and $SRFFN(\cdot)$ are the Spiking RWKV and Spiking Receptance Feed-Forward Network operations respectively, and Y is the final output representing the estimated nonlinear distortion components. E_n represents the embedding dimension after the CNN layer, calculated as $E_n = n_{\text{symbols}} - n_k + 1 = (2n_{\text{taps}} + 1) - n_k + 1$, and n_f denotes the number of filters in the CNN layer.

The BE layer serves as the critical interface between continuous-valued optical signal processing and the event-driven spiking neural network component of our hybrid architecture. This layer implements the data encoding strategy that converts continuous optical signal embeddings into binary spike trains suitable for neuromorphic processing. The BE operation can be mathematically expressed as:

$$X_0 = BE(E) = \Theta(W_{BE}E + b_{BE}) \tag{5.10}$$

where $\Theta(\cdot)$ is the Heaviside step function, W_{BE} and b_{BE} are the learnable weights and biases of the BE layer.

The binary embedding layer implements a critical interface between continuous-valued optical signal processing and event-driven neuromorphic computation. In optical communication systems, received symbols are inherently continuous-valued complex numbers representing the in-phase and quadrature components of the modulated signal. The BE layer transforms these continuous embeddings into binary spike trains that preserve the essential information while enabling efficient spiking neural network processing.

The theoretical foundation of binary embedding in optical communications rests on the concept of temporal coding, where information is encoded in the timing and patterns of spikes rather than in analogue amplitude values. For optical signals, this encoding offers several advantages:

- **Noise Robustness**: Binary spikes are inherently more robust to noise compared to analogue representations, making them suitable for processing optical signals corrupted by ASE noise and nonlinear distortions.
- **Energy Efficiency**: Binary operations consume significantly less energy than floating-point arithmetic, crucial for high-speed optical systems with stringent power constraints.
- **Neuromorphic Compatibility**: The binary nature of spikes aligns perfectly with neuromorphic hardware implementations, enabling efficient deployment on specialised processors designed for event-driven computation.
- **Temporal Precision**: Spike timing can encode fine-grained temporal information, allowing the representation of phase relationships and timing-sensitive optical signal characteristics.

The learnable parameters W_{BE} and b_{BE} are optimised during training to achieve optimal signal-to-spike conversion that preserves the discriminative features necessary for nonlinearity compensation. The training process automatically learns appropriate thresholds and weights that maximise the mutual information between the continuous input embeddings and the resulting binary spike patterns, ensuring that critical optical signal features are preserved in the spike domain.

During training, the non-differentiable Heaviside function poses a challenge for gradient-based optimisation. To address this, we employ surrogate gradient methods where the Heaviside function is approximated during the backward pass using the arctangent surrogate function:

$$\frac{\partial \Theta(x)}{\partial x} \approx \frac{\alpha}{2\left(1 + \left(\frac{\pi}{2}\alpha x\right)^2\right)} \tag{5.11}$$

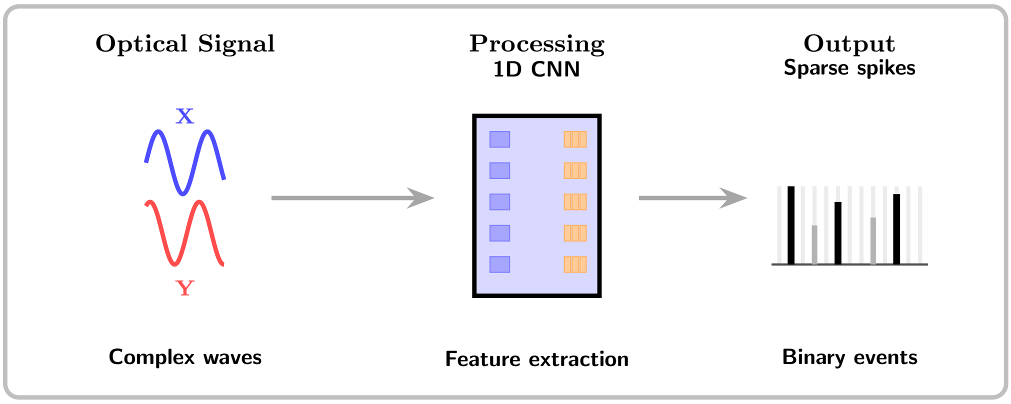


Figure 5.5: Binary encoding strategies for optical signal embeddings. A 1D CNN first extracts features from complex dual-polarisation optical waveforms (X, Y), producing continuous-valued representations. These are then converted into sparse spike trains, providing an event-driven binary encoding that interfaces the optical signal processing front-end with the neuromorphic SNN-RWKV equaliser.

This surrogate gradient enables end-to-end training of the hybrid architecture while maintaining the binary spike representation during forward inference. The parameter α controls the steepness of the approximation and is typically set to balance gradient flow with spike precision.

5.2.2 Spiking Neuron Dynamics and Network Components

The spiking dynamics employ the LIF neuron model, with its dynamics described by:

$$U[t] = H[t] + \beta(Y[t] - (H[t-1] - U_{\text{reset}})) \tag{5.12}$$

$$S[t] = \Theta(U[t] - U_{\text{threshold}}) \tag{5.13}$$

$$H[t] = U[t] \cdot (1 - S[t]) \tag{5.14}$$

where $U[t]$ represents the membrane potential at time step t, $H[t]$ is the potential after reset, $S[t]$ is the binary spike output, $Y[t]$ is the input to the neuron, β is the decay factor, $U_{\text{threshold}}$ is the firing threshold, and U_{reset} is the reset potential. The parameters are set as $U_{\text{threshold}} = 1.5$, $U_{\text{reset}} = 0$, and $\beta = 2.5$ [207].

The Heaviside function $\Theta(\cdot)$ makes the spiking mechanism non-differentiable. To address this during backpropagation, we employ the arctangent surrogate gradient function defined as:

$$\frac{\partial S[t]}{\partial U[t]} \approx \sigma'(U[t] - U_{\text{threshold}}) = \frac{\alpha}{2\left(1 + \left(\frac{\pi}{2}\alpha(U[t] - U_{\text{threshold}})\right)^2\right)} \tag{5.15}$$

where α is a scaling parameter that controls the steepness of the surrogate function.

The RWKV model consists of the components $\mathbf{R}$, $\mathbf{W}$, $\mathbf{K}$, and $\mathbf{V}$, that are processed sequentially, differing from traditional attention mechanisms [205, 220, 221].

The SRWKV block utilises key vectors $\mathbf{K}$, value vectors $\mathbf{V}$, and receptance vectors $\mathbf{R}$, which adaptively assimilate historical data and inputs:

$$R = X_0 M_R, \tag{5.16}$$

$$K = X_0 M_K, \tag{5.17}$$

$$V = X_0 M_V, \tag{5.18}$$

$$\{M_R, M_K, M_V\} \in \mathbb{R}^{n_f \times H} \tag{5.19}$$

with $H = E_n$, where M_R, M_K, and M_V are learnable parameter matrices.

The block's output $X_1'[t+1]$ is determined by combining these vectors with dynamic hidden states A and B:

$$X_1'[t+1] = SN\left(\sigma(R[t]) \odot F[t]\right) \tag{5.20}$$

$$F[t] = \frac{\exp(W_f) \odot \exp(K[t]) \odot V[t] + A[t]}{\exp(W_f) \odot \exp(K[t]) + B[t]} \tag{5.21}$$

$$A[t] = \exp(K[t]) \odot V[t] + \exp(W_d) \odot A[t-1] \tag{5.22}$$

$$B[t] = \exp(K[t-1]) + \exp(W_d) \odot B[t-1] \tag{5.23}$$

where $SN(\cdot)$ represents the spiking neuron operation, $\sigma(\cdot)$ is the sigmoid activation function, $\odot$ denotes element-wise multiplication, $W_f \in \mathbb{R}^{1\times d}$ is a weight vector responsible for weighting the current time-step's information, and $W_d \in \mathbb{R}^{1\times d}$ is a decay vector that controls the influence of information from previous time steps.

The hidden states $A[t]$ and $B[t]$ serve as memory units that accumulate and retain information from previous time steps, enabling the model to capture long-range dependencies. Unlike self-attention which requires access to the entire sequence simultaneously, the recurrent formulation of SRWKV allows for sequential processing, making it compatible with the temporal dynamics of SNNs.

The token shift mechanism is fundamental for enhancing temporal context. Before processing the input in each block, a token shift operation is applied to enhance the model's ability to leverage contextual information. Given an input X, the token shift operation is defined as:

$$X_s = \text{ZeroPad}_{[0,0,-1,1]}(X) \tag{5.24}$$

$$W_{\text{shift}} = \left\{ \left(\frac{i}{E_n} \right)^{n_b/N} \right\}, \quad i = 1, \ldots, E_n \tag{5.25}$$

$$X = W_{\text{shift}} \odot X + (1 - W_{\text{shift}}) \odot X_s \tag{5.26}$$

where ZeroPad$_{[0,0,-1,1]}(\cdot)$ denotes zero padding operation that clips the top row and zero-pads the bottom row, W_{shift} is a learnable shift mask, n_b is the current block, and N is the total number of blocks.

Figure 5.6 depicts the streaming LIF–RWKV time-mix used in SRWKV, including the decayed accumulators A,B and the spike-gated update.

The token shift mechanism is particularly crucial for optical channel equalisation as it enables the model to access temporal context that spans multiple symbol periods. In optical communication systems, nonlinear impairments such as SPM and XPM create dependencies between adjacent symbols. The token shift operation allows each processing block to compare the current symbol with its temporal neighbors, enabling more effective compensation of these memory effects.

The weight mask W_{shift} implements a progressive shift strategy where deeper blocks (n_b closer to N) rely more heavily on shifted information. This design choice reflects the hierarchical nature of optical signal processing: early blocks focus on local temporal correlations (e.g., chromatic dispersion effects), while deeper blocks capture longer-range dependencies (e.g., accumulated nonlinear effects). The mathematical formulation ensures that W_{shift} increases monotonically with block depth, providing a principled approach to temporal context integration.

In the context of dual-polarisation systems, the token shift mechanism operates independently on each polarisation component, allowing the model to capture both intra-polarisation and inter-polarisation dependencies that arise from PMD and PDL effects.

The Spiking Receptance Feed-Forward Network (SRFFN) block processes the output from SRWKV using a series of transformations to control information flow:

$$\begin{gathered} X_1[t] = \sigma(M_P X_1'[t]) \odot M_S(\text{ReLU}^2(M_G X_1'[t])), \\ \{M_P, M_G, M_S\} \in \mathbb{R}^{n_f \times H} \end{gathered} \tag{5.27}$$

where $H = E_n$, and $\{M_P, M_G, M_S\}$ are learnable parameter matrices. The term ReLU$^2(\cdot)$ represents the squared ReLU activation function defined as ReLU$^2(x) = (\max(0, x))^2$. The SRFFN block functions as a channel mixer, complementing the token mixing performed by the SRWKV block.

The output of the SRFFN block is then integrated into the spiking neuron dynamics, following the LIF model described earlier. The combination of token mixing in SRWKV and channel mixing in SRFFN enables the model to effectively capture both spatial and temporal correlations in the input data.

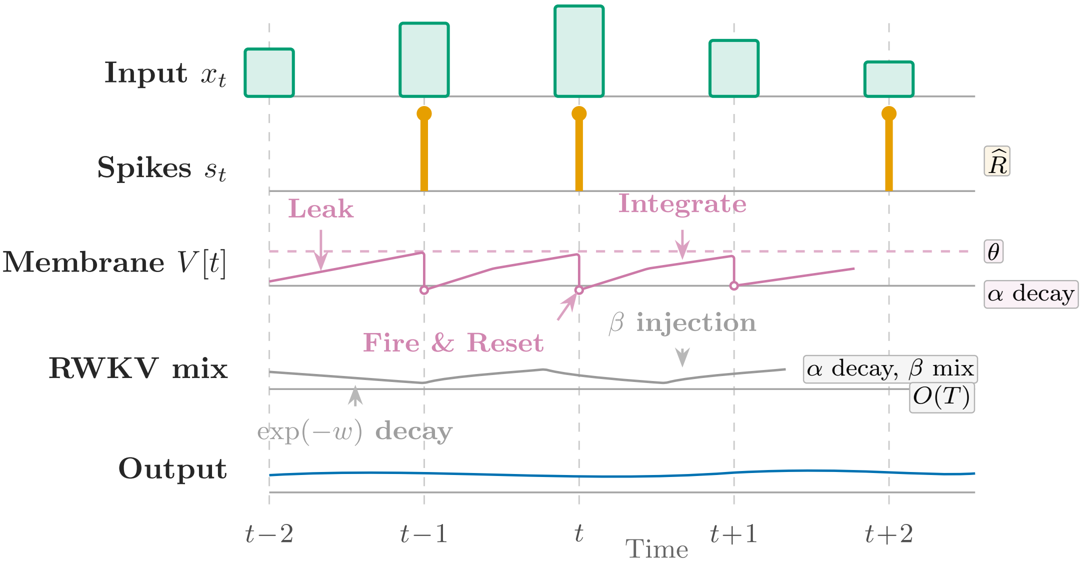


Figure 5.6: SRWKV signal dynamics. Incoming samples x_t are encoded into spikes s_t and integrated by LIF neurons (leak, integrate, fire/reset), while the RWKV time-mix updates the decayed accumulators A, B ($\exp(-w)$ decay) to produce a streaming output. This yields constant-memory, linear-time updates per token.

5.2.3 Training Configuration and Optimisation

The output layer computes the estimated nonlinear distortions for the target symbols:

$$Y = [\hat{E}_{XI}, \hat{E}_{XQ}] = Linear(X_1) \tag{5.28}$$

where $Linear(\cdot)$ represents a linear transformation that maps the final hidden representation X_1 to the prediction space. $\hat{E}_{XI}$ and $\hat{E}_{XQ}$ are the estimated in-phase and quadrature components of the nonlinear distortion, which are used to compensate for the nonlinear effects in the received signal.

The training of the hybrid SNN-RWKV architecture requires careful consideration of both the continuous RWKV components and the discrete spiking elements. The training process follows these key principles:

1. **Surrogate Gradient Method:** The non-differentiable spiking functions in the SNN components are handled using surrogate gradients during backpropagation. Specifically, the arctangent surrogate function is applied to the Heaviside step functions in both the Binary Embedding layer and the LIF neuron dynamics:

$$\frac{\partial S}{\partial u} \approx \frac{\alpha}{2\left(1 + \left(\frac{\pi}{2}\alpha(u - u_{\text{th}})\right)^2\right)} \tag{5.29}$$

2. **Unified Loss Function:** The hybrid architecture is trained end-to-end using MSE loss:

$$L_{\text{MSE}} = \sum_{y_i} \gamma (y_i - \hat{y}_i)^2 \tag{5.30}$$

   where $\gamma = \frac{10^{-4}}{B}$ is a scaling factor, B is the batch size, y_i represents the true distortion values, and $\hat{y}_i$ represents the predicted distortion values.

3. **Gradient Flow Through Hybrid Components:** The training process ensures gradient flow through both the continuous RWKV attention mechanisms and the discrete spiking components. The residual connections facilitate gradient propagation, while the surrogate gradients enable learning in the spiking components without compromising the binary spike representation during inference.

The SNN-RWKV is compared with two advanced ANN-based equalisers, the CNN-MLP and CNN-biLSTM [103]. The BO algorithm is employed to fine-tune hyperparameters for all models at a launch power of 1 dBm, which is 3dBm above the optimal launch power of the system [103].

The CNN-biLSTM model was optimised with parameters: $n_{\text{taps}} = 5$, $n_f = 318$, $n_k = 11$, $n_h = 462$, $B = 1219$, and $\eta = 0.00699$.

For the CNN-MLP model, optimal settings determined by BO included $n_{\text{taps}} = 19$, $n_f = 339$, $n_k = 9$, neurons per layer n_i^l in $[81, 74, 84]$ across $n_l = 3$ layers, $B = 1815$, and $\eta = 0.0038$.

The proposed SNN-RWKV equaliser's optimal hyperparameters were $n_{\text{taps}} = 20$, $n_f = 300$, $n_k = 1$, $n_b = 1$ for the repetitions of SRWKV and SRFFN blocks, $B = 4000$, and $\eta = 0.0064$.

All ANNs were trained for at most 400 epochs and stopped earlier in the case of no improvement in the loss function value, i.e., MSE, over 50 epochs. After every epoch, the BER was computed using the independently generated testing dataset. The early stopping criterion helps prevent overfitting and ensures that the models generalise well to unseen data.

For application in optical signal equalisation, several implementation considerations are taken into account:

1. **Spike Rate Adaptation**: The spike rate of the SNN-RWKV model, measured as $\hat{R} = 0.091$ in our experiments, is a critical factor affecting both performance and energy efficiency. A lower spike rate leads to more sparse operations and higher energy efficiency.

2. **Temporal Processing**: Unlike conventional ANNs that process the entire symbol sequence simultaneously, the SNN-RWKV model processes symbols sequentially, leverageing the temporal dynamics inherent in SNNs. This makes it particularly suitable for real-time signal processing applications.

3. **Neuromorphic Hardware Compatibility**: The binary nature of spike-based communications and the recurrent structure of the SNN-RWKV model make it highly compatible with

neuromorphic hardware implementations, offering significant advantages in terms of energy efficiency and latency.

4. **Dual Polarisation Handling**: The model is designed to handle dual polarisation signals by processing both the in-phase and quadrature components of each polarisation, ensuring comprehensive nonlinearity compensation.

This architecture, which employs both spiking neural network dynamics and attention-based mechanisms, not only allows for effective nonlinear compensation in optical communication systems but also provides competitive performance with respect to traditional equalisers, at the same time that it facilitates hardware-friendly deployments, offering high-speed operation while maintaining energy efficiency. This combination positions SNN-RWKV as a highly advantageous solution for modern optical networks, addressing both the technical demands and operational efficiencies required for next-generation communication technologies.

5.3 Computational Complexity and Energy Efficiency of ANN-based Equalisers: A Neuromorphic Perspective

Energy efficiency is a critical consideration in the deployment of ANN-based equalisers in optical communication systems, especially for real-time, high-speed links [52, 68]. In general, the total energy consumption of an ANN-based equaliser can be modelled as

$$E_{\text{total}} = E_{\text{comp}} + E_{\text{mem}} + E_{\text{comm}}, \tag{5.31}$$

where E_{comp} is the energy consumed by computations, E_{mem} is the energy for memory accesses, and E_{comm} is the energy for data movement. In this work, we primarily focus on E_{comp}, the computational energy of these models, since practical hardware implementations often show that multiplications dominate power usage due to their circuit complexity [103, 219].

5.3.1 Energy Efficiency and Neuromorphic Computational Analysis

Baseline Energy Estimation for ANN-based Equalisers. The computational energy E_{comp} for an ANN can be broken down into contributions from MAC operations, activation function computations, and Accumulate Operation (AC) operations:

$$E_{\text{comp}} = N_{\text{MAC}} \cdot E_{\text{MAC}} + N_{\text{act}} \cdot E_{\text{act}} + N_{\text{AC}} \cdot E_{\text{AC}}, \tag{5.32}$$

where N_{MAC} is the total number of MAC operations, N_{act} is the number of activation function evaluations, and N_{AC} is the number of AC operations. Since multipliers typically require more logic gates than adders, E_{MAC} tends to dominate. In fact, the energy spent on activation functions can frequently be neglected in comparison [222].

Typical energy values assumed for these hardware operations, based on empirical measurements, lie around [207,223]:

- $E_{\text{MAC}} \approx 4.6\,\text{pJ}$,
- $E_{\text{AC}} \approx 0.9\,\text{pJ}$.

For a generic ANN-based equaliser, thus, one often approximates

$$E_{\text{comp}} \approx N_{\text{MAC}} \cdot E_{\text{MAC}} + N_{\text{AC}} \cdot E_{\text{AC}}. \tag{5.33}$$

In the context of optical channel equalisation, the total number of FLOPs (floating-point operations) primarily determines N_{MAC}. Since multiplier circuits are more power-hungry, accurately counting the number of multiplications is key [52]. For a typical MLP or similar neural network,

$$\text{FLOPs} = N_{\text{layers}} \cdot N_{\text{neurons}} \cdot N_{\text{inputs}} \cdot N_{\text{samples}}, \tag{5.34}$$

where N_{layers} is the number of layers, N_{neurons} is the number of neurons per layer, N_{inputs} is the number of inputs to each neuron, and N_{samples} is the total number of samples processed.

In a SNN, each neuron generates binary spikes rather than continuous outputs. This spiking mechanism can significantly reduce the computational load due to the sparsity of activations: only a fraction $\hat{R}$ of neurons spike at any given time. Consequently, the effective number of MAC operations is scaled by the firing rate $\hat{R}$:

$$\text{FLOPs}_{\text{SNN}} = \text{FLOPs} \cdot \hat{R}. \tag{5.35}$$

Here, $\hat{R}$ is the proportion of non-zero elements in the spike matrices, reflecting how often each neuron actually fires. In practice, this event-driven processing can bring sizable energy savings [219].

Beyond simply reducing effective FLOPs, neuromorphic hardware leverages asynchronous, event-driven computations, transforming dense operations into sparse additions [224]. This characteristic makes SNN-based equalisers, such as the proposed SNN-RWKV, particularly appealing for real-time optical systems, where both computational speed and power efficiency are critical constraints [52,68,103,224].

In the specific case of the SNN-RWKV Equaliser, the number of FLOPs for each model follows the procedures in [103,120,224]. We also account for the measured average firing rate $\hat{R} = 0.091$, which promotes the inherent sparsity in spiking operations.

Hence, for each of the three models considered—CNN-MLP, CNN-biLSTM, and SNN-

RWKV—the resulting energy consumption is derived from:

$$
\begin{aligned}
\mathrm{E}_{\text{CNN-MLP}} &= E_{\text{MAC}} \cdot \big(n_f \cdot n_i \cdot n_k \cdot E_n + n_i \cdot E_n \\
&\qquad + \sum_{l=2}^{5} n_i^l \cdot n_i^{l-1} + n_i^6 \cdot n_o\big), \\
\mathrm{E}_{\text{CNN-biLSTM}} &= E_{\text{MAC}} \cdot \Big(n_f \cdot n_i \cdot n_k \cdot E_n \\
&\qquad + 2 \cdot E_n \cdot n_h \cdot (4n_f + 4n_h + 3 + n_o)\Big), \\
\mathrm{E}_{\text{SNN-RWKV}} &= E_{\text{MAC}} \cdot (n_f \cdot n_i \cdot n_k \cdot E_n) \\
&\qquad + E_{\text{AC}} \cdot \hat{R} \left(3E_n \cdot n_f^2 + 9E_n \cdot n_f\right) \\
&\qquad + E_{\text{MAC}} \cdot E_n \cdot n_o.
\end{aligned}
\tag{5.36}
$$

Here,

- n_f is the number of filters in the CNN portion,
- n_i is the number of inputs or neurons in relevant layers,
- n_k is the kernel width of the CNN,
- E_n is the embedding dimension or effective sequence length in the network,
- n_h is the hidden size for the LSTM,
- n_o is the final output dimension, and
- $\hat{R}$ is the spiking rate for the SNN.

By incorporating the sparse, event-driven properties of the SNN-RWKV model, our approach aligns closely with the requirements for energy-efficient, real-time optical signal equalisation. The following sections present a comparative performance evaluation of the proposed SNN-RWKV and its conventional ANN counterparts, highlighting both computational savings and Q-factor improvements in an optical communication setting.

5.4 Photonic and Optoelectronic Hardware for Spiking Neurons

The practical implementation of spiking neural networks for optical communications requires hardware platforms capable of processing temporal information at rates compatible with high-speed optical signals. While electronic implementations of SNNs (such as Intel's Loihi [198]) have demonstrated energy-efficient neuromorphic computing, photonic and optoelectronic approaches offer unique advantages in terms of processing speed, bandwidth, and direct integration with optical communication systems [225, 226]. This section discusses the state-of-the-art photonic hardware implementations for spiking neurons, their performance characteristics, and their potential applicability to the optical equalisation tasks considered in this thesis.

5.4.1 Photonic SNN Hardware Implementations

Photonic platforms offer several promising device-level mechanisms for implementing spiking neural networks at ultrafast timescales. The following architectures represent the most mature and experimentally validated approaches for photonic spike generation, weighted summation, and large-scale parallel interconnects.

- **Silicon Photonics Integrated Neurons:** Silicon photonics platforms provide a mature fabrication technology for implementing optical neural networks with Complementary Metal-Oxide-Semiconductor (CMOS)-compatible processing [227, 228]. Integrated waveguide-based neurons leverage nonlinear optical effects to create spike-like behaviour. Recent demonstrations have shown that silicon photonic circuits can implement weighted summation and thresholding operations [229], which are fundamental to spiking neuron dynamics. The key advantage of silicon photonics is the ability to fabricate complex networks with hundreds of neurons on a single chip, with waveguide routing providing high-density interconnects.
- **Micro-ring Resonators for Spike Generation:** Micro-ring resonators exhibit optical bistability, enabling them to function as threshold devices for spike generation [230]. When driven with appropriate pump power, these resonators can switch between high and low transmission states in response to input optical signals, mimicking the integrate-and-fire behaviour of biological neurons. Feldmann et al. [230] demonstrated all-optical spiking neurosynaptic networks with self-learning capabilities using phase-change materials integrated with micro-ring resonators. These systems achieved nanosecond-scale spike generation, orders of magnitude faster than electronic neuromorphic processors.
- **Laser-based Spiking Neurons:** Semiconductor lasers with saturable absorption or excitable dynamics can generate optical pulses (spikes) in response to input perturbations [231]. Nahmias et al. [231] demonstrated a leaky integrate-and-fire laser neuron achieving sub-nanosecond response times with energy consumption on the order of femtojoules per spike—comparable to biological neurons but at GHz rates. The laser neuron approach benefits from natural gain and nonlinearity, eliminating the need for external amplification in many cases.
- **VCSEL Arrays for Parallel Spiking Networks:** VCSELs offer an attractive platform for implementing large-scale parallel spiking neural networks [232, 233]. VCSEL arrays can be directly modulated to produce spike trains, with each laser serving as an independent spiking neuron. Robertson et al. [232] demonstrated ultrafast optical integration and pattern classification using spiking VCSEL neurons operating at rates exceeding 1 GHz. The two-dimensional

array geometry of VCSELs facilitates scalable architectures and enables wavelength-division multiplexing for increased connectivity [233].

5.4.2 Optoelectronic Hybrid Platforms

While all-optical implementations offer speed advantages, practical neuromorphic systems often benefit from optoelectronic hybrid architectures that leverage the complementary strengths of optical and electronic domains [234].

- **Photodetector-based Spike Reception:** High-speed photodetectors convert optical spikes into photocurrent pulses that can be processed by electronic circuits. This enables seamless integration of photonic spiking neurons with electronic memory and control logic. Modern photodetectors offer bandwidths from tens to hundreds of GHz, well-matched to optical communication systems and neuromorphic processing interfaces.
- **Electro-optic Spike Generation:** Electronic signals can modulate optical devices (VCSELs, LEDs, or electro-absorption modulators) to generate optical spikes. This allows electronic LIF circuits to control optical spike generation, leverageing mature analogue circuit design techniques while exploiting the bandwidth and parallelism of optical interconnects [234].
- **Superconducting Optoelectronic Neurons:** Shainline et al. [234] introduced superconducting optoelectronic circuits that combine Josephson junctions with integrated photodetectors and light sources. These hybrid devices achieve picosecond-scale synaptic integration with energy dissipation below 1 aJ per synaptic event, offering a route toward extremely energy-efficient, large-scale neuromorphic systems.

5.4.3 Implementation Feasibility and Performance Characteristics

The implementation feasibility of photonic spiking neural networks depends on three key dimensions: processing speed, energy consumption, and scalability. The following points summarize the main characteristics and practical considerations associated with each aspect.

- **Processing Speed:** Photonic spiking neurons operate at timescales determined by photon lifetimes in optical cavities and carrier dynamics in semiconductors, typically in the range of picoseconds to nanoseconds. This translates to spiking rates of hundreds of MHz to tens of GHz [230, 231], compared to typical electronic neuromorphic chips operating in the kHz range (e.g., Loihi [198] with maximum spike rates around 1 kHz). For optical communication applications where symbol rates range from tens of GHz (short-reach) to hundreds of GHz (future long-haul systems), photonic SNNs offer the potential for real-time, symbol-rate equalisation without downsampling [232, 235–237].

- **Energy Consumption:** Photonic neurons typically consume energy in the range of femtojoules to picojoules per spike [231], corresponding to milliwatts of power at GHz spiking rates. This is higher than state-of-the-art electronic neuromorphic processors (e.g., Loihi's approximately 30 pJ per synaptic operation [238]), but the comparison is complicated by the vastly different operating speeds. When normalised by throughput (energy per operation per second), photonic implementations can be competitive, particularly when considering the energy cost of electronic-to-optical conversion that would be required if electronic processors were used in optical communication systems [237].
- **Scalability and Integration:** WDM provides a unique scaling mechanism for photonic neural networks. Multiple wavelength channels can propagate through the same waveguide without interference, enabling massive parallelisation with limited physical interconnects [228]. However, photonic integration faces challenges including coupling losses between different photonic components, thermal sensitivity of resonant devices, and the need for precise fabrication tolerances. Current demonstrations are typically limited to tens to hundreds of photonic neurons, whereas electronic neuromorphic chips (Loihi, TrueNorth) have scaled to millions of neurons [239, 240].

5.4.4 Examples and Recent Demonstrations

Recent advances in photonic and optoelectronic hardware have enabled several experimental demonstrations of neuromorphic and photonic neural network architectures. The following examples highlight some of the most relevant platforms and their capabilities.

- **MIT Silicon Photonic Neural Networks:** Shen et al. [227] demonstrated deep learning with coherent nanophotonic circuits implemented on a silicon photonics platform, in which integrated Mach-Zehnder Interferometer (MZI) meshes implement matrix–vector multiplications. Their analysis *projects* femtojoule-level energy per operation under specific assumptions (e.g., shot-noise–limited detection), rather than reporting a measured fJ/MAC figure in that demonstration.
- **Oxford/Exeter/KIT Phase-Change Photonic Synapses:** Phase-change materials (such as $Ge_2Sb_2Te_5$) integrated with photonic waveguides can implement non-volatile, multi-level photonic memory/synaptic weights [241, 242]. The Wright, Bhaskaran, and Pernice groups have demonstrated devices with multiple programmable states, supporting on-chip programmability and weight storage.
- **Delay-Coupled Laser Reservoirs (Brunner, Larger, and collabourators):** Brunner, Larger, and collabourators have extensively studied reservoir computing using delay-coupled semiconductor lasers [243, 244]. Their photonic reservoir computers have demonstrated strong performance on time-series prediction tasks and have been applied to optical communication

equalisation [38]. The delay-coupled architecture provides a large effective network size ("virtual neurons") with minimal physical hardware.

- **Princeton Neuromorphic Photonics:** The Prucnal group at Princeton has developed neuromorphic photonic networks using silicon photonic weight banks and broadcast-and-weight architectures [228]. While these demonstrations established the hardware building blocks, the frequently cited sub-nanosecond latencies and femtojoule-per-MAC energy efficiencies are *prospective* advantages discussed in the literature, not measurements reported [228]. Recent work from the group has also explored spiking photonic neurons for event-driven processing [231].

While photonic SNN hardware offers compelling performance characteristics, practical deployment for optical communication equalisation faces several challenges. Current demonstrations are primarily proof-of-concept with limited network sizes. The on-chip learning mechanisms required for adaptive equalisation (such as Spike-Timing Dependent Plasticity (STDP) in photonic implementations) are still under development. Moreover, the cost and complexity of photonic integration may not be justified when high-performance electronic DSP can achieve comparable equalisation quality at lower cost, albeit with higher latency and power consumption. Nonetheless, for ultra-high-speed systems (e.g., 400 Gb/s and beyond per wavelength) or power-constrained edge scenarios where optical processing can eliminate costly O–E–O conversions, photonic neuromorphic equalisers present a promising future direction [38].

The SNN-RWKV architecture proposed in this thesis is designed to be hardware-agnostic, with the potential for implementation on electronic neuromorphic chips (Loihi, Akida) for moderate-speed applications or on photonic platforms for ultra-high-speed systems. The sparse event-driven computation inherent to the SNN component is well-suited to the energy-efficient characteristics of both electronic and photonic neuromorphic hardware [219].

5.5 Performance Comparison of the SNN-RWKV Equaliser with Respect to Traditional Methods

The performance comparison between the proposed SNN-RWKV hybrid architecture and traditional methods requires careful consideration of both equalisation effectiveness and computational efficiency. This section provides a comprehensive evaluation that demonstrates the advantages of the hybrid approach over conventional ANN-based equalisers and establishes the theoretical rationale for the observed performance improvements.

To validate the theoretical complexity advantages discussed previously, we compare the actual computational requirements using the corrected formulations from the updated section:

Traditional CNN-MLP Complexity:

$$\mathrm{FLOPs}_{\text{CNN-MLP}} = n_f\, n_i\, n_k\, E_n \;+\; \sum_{l=1}^{n_l} n_i^l\, n_i^{l-1} \;\approx\; O(n_f n_i n_k E_n). \tag{5.37}$$

CNN-biLSTM Complexity:

$$\mathrm{FLOPs}_{\text{CNN-biLSTM}} = n_f\, n_i\, n_k\, E_n \;+\; 2\, E_n\, n_h\big(4n_f + 4n_h + 3 + n_o\big) \;\approx\; O(E_n n_h^2). \tag{5.38}$$

SNN-RWKV Complexity:

$$\mathrm{FLOPs}_{\text{SNN-RWKV}} = n_f\, n_i\, n_k\, E_n \;+\; \hat{R}\left(3E_n n_f^2 \;+\; 9E_n n_f\right) \;\approx\; O(\hat{R}\, E_n\, n_f^2). \tag{5.39}$$

The measured sparse firing rate $\hat{R} = 0.091$ yields an effective token-mixing workload of only $\sim 9.1\%$ of the dense counterpart, highlighting the computational advantage of the hybrid design.

The performance advantages of SNN-RWKV over traditional approaches can be understood through several key factors:

1. **Temporal Processing Efficiency:**
    - CNN–MLP has no temporal state and processes symbols independently.
    - CNN–biLSTM performs sequential updates with quadratic hidden-state complexity.
    - SNN–RWKV implements the RWKV time-mixing formulation, achieving *linear-time* streaming with constant per-step state and spike-gated computation.
2. **Memory Efficiency:**
    - CNN–MLP: $O(n_f n_i n_k E_n)$ — no temporal memory.
    - CNN–biLSTM: $O(E_n n_h)$ during training (Backpropagation Through Time (BPTT)) and $O(n_h)$ at inference.
    - SNN–RWKV: constant-size accumulators $O(d_{\text{model}})$ independent of E_n.
3. **Energy Efficiency:**
    - CNN–MLP: dense MACs throughout.
    - CNN–biLSTM: recurrent operations with high memory bandwidth cost.
    - SNN–RWKV: sparse spike-gated operations with an 87.7% energy reduction in contrast to CNN-biLSTM and 8.4% when compared with CNN-MLP.

The performance of the different models was assessed through the standard Q-factor metric recalculated from the directly counted BER as $Q = 20\log_{10}\left[\sqrt{2}\,\mathrm{erfc}^{-1}(2\,\mathrm{BER})\right]$ [52].

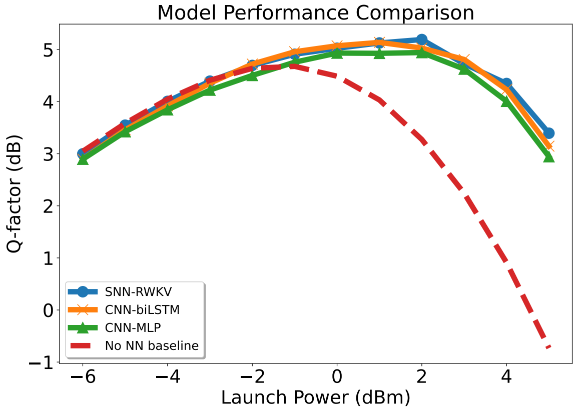


Figure 5.7: Performance comparison of the 3 different ANN-based equalisers for different signal launch powers. A baseline corresponding to Offline DSP without Nonlinearity mitigation (see Figure 5.4) is included.

Figure 5.7 shows the comparative performance of the SNN-RWKV, CNN-biLSTM, CNN-MLP, and a baseline across a range of launch powers that goes from -6 dBm to 5 dBm. The results indicate that SNN-RWKV outperforms the other models across all launch powers considered. Thus, the robust capability of the SNN-RWKV in handling nonlinear distortions is demonstrated. Nevertheless, the CNN-biLSTM closely follows the performance of the SNN-RWKV, while the CNN-MLP is outperformed by these other two models, which is congruent with the results presented in other works [103].

Table 5.1: Comparison of Q-factor and Energy consumption of different ANN-based equalisers for the optimal launch power of the system, i.e. -2 dBm.

Energy Consumption Evaluation		
Model	**Q-factor (dB)**	**Energy (pJ)**
CNN-biLSTM	4.6	13.3×10^6
CNN-MLP	4.45	1.79×10^6
SNN-RWKV	4.7	1.64×10^6

Table 5.1 presents a comparison of Q-factor and energy consumption for the evaluated models at the optimal launch power of the system, i.e. -2 dBm. The SNN-RWKV exhibits a Q-factor of 4.70 dB and the lowest energy consumption of 1.64×10^6 pJ. This dual advantage positions the proposed model as a highly efficient equaliser for optical communication systems.

In contrast, both the CNN-biLSTM and CNN-MLP show a lower Q-factor of 4.6 dB and 4.45 dB, respectively, while their energy consumption is higher. This is accentuated in the case of the CNN-biLSTM model, where a value of 13.3×10^6 pJ is obtained. For the CNN-MLP, a theoretical energy consumption of 1.79×10^6 pJ is calculated, still above the value reported for the solution proposed in this work.

The superior performance of SNN-RWKV over traditional approaches can be traced to five complementary factors:

1. Optimal temporal representation. The hybrid architecture combines spike-based dynamics with attention-like token mixing. In our implementation, the spiking pathway follows a LIF update

$$V[t] \;=\; \alpha\, V[t{-}1] \;+\; (1{-}\alpha)\, V_{\text{rest}} \;+\; \beta_{\text{RWKV}}\, F[t], \tag{5.40}$$

where $F[t]$ denotes the RWKV time-mix output. The LIF term supplies short-term memory via the leak/integration balance, while $F[t]$ provides gated, longer-range context [205, 207, 245].

2. Adaptive channel matching. RWKV replaces global softmax attention with an exponentially decayed recurrent update. The learned decay(s) act as a causal, exponentially weighted kernel, which aligns with the *effective* memory profile of coherent optical links—where Nonlinear Interference Noise (NLIN) is dominated by interactions among nearby symbols and decays with temporal separation (as predicted by perturbation/Enhanced Gaussian Noise (EGN) models) [205, 246, 247]. In this sense, the RWKV time-mix provides a parsimonious, streaming surrogate for the relevant channel memory without quadratic attention overhead.

3. Noise robustness. Spike generation is threshold-based:

$$S[t] \;=\; \Theta\!\big(V[t] - V_{\text{threshold}}\big), \tag{5.41}$$

so sub-threshold fluctuations (e.g., ASE-like jitter around decision boundaries) are naturally suppressed, which is a standard advantage of event-driven SNN processing and contributes to energy efficiency in neuromorphic implementations [207].

4. Nonlinear compensation coverage. The hybrid division of labour is complementary: LIF units efficiently handle fast, local nonlinearities and signal-dependent effects (e.g., SPM-like terms), while RWKV captures longer-range, cross-symbol interactions that drive inter- and intra-polarisation effects (e.g., XPM/FWM contributions) [204, 214, 246, 247]. This combination helps explain the SNN–RWKV gains over CNN–MLP and its edge over (or parity with) CNN–biLSTM across launch powers.

5. Training efficiency. We train end-to-end with surrogate gradients for the non-differentiable spike nonlinearity; the arctangent surrogate used in this work provides smooth, stable gradients while keeping binary spikes at inference [206]. This preserves the sparsity advantages of SNN inference (measured average spike rate $\hat{R}$) without sacrificing training stability.

As already mentioned when explaining Transformer equalisers, self-attention enables rich token interactions but entails quadratic time/memory in the attention window [204]. Causal Transformers can stream using key–value (KV) caching, yet the cache and per-step compute grow with context length, impacting real-time throughput [208]. In contrast, the RWKV time-mix performs constant-memory, linear-time updates per token at inference [205].

Scalability (illustrative). For a full self-attention window of N tokens, compute/memory are $O(N^2)$ [204]. As a concrete example, $N{=}2048$ implies $\sim 4.2{\times}10^6$ pairwise interactions; KV-cached streaming still incurs per-step cost and memory that grow with the attended context length [208]. By contrast, SNN-RWKV maintains constant per-token state (decayed accumulators) and linear-time updates, with its effective token-mixing work further reduced by the measured sparsity $\hat{R}$.

Real-time processing. A full-window Transformer equaliser must materialise the attended window (or maintain a growing KV cache); RWKV processes symbols as they arrive with a constant-size state, matching streaming receiver constraints [205, 208].

Energy scaling (back-of-the-envelope). With technology-typical energies (e.g., $E_{\text{MAC}} \approx 4.6\,$pJ, $E_{\text{AC}} \approx 0.9\,$pJ [223]), a dense full-window Transformer block scales approximately as

$$E_{\text{Transformer}} \;\propto\; N^2 \cdot d_k \cdot h \cdot E_{\text{MAC}}, \tag{5.42}$$

reflecting quadratic attention cost; this is a scaling relation (not a measured number). In SNN–RWKV, many dense multiplies are replaced by sparse additions gated by spikes (proportional to $\hat{R}$), and the time-mix state is constant in N, which is consistent with the measured energy reductions reported in Table 5.1 and prior neuromorphic analyses [207, 223].

The overall analysis underscores the SNN-RWKV equaliser as a robust and energy-efficient solution for handling nonlinearities in optical communication systems. Its ability to maintain high performance with significantly lower energy consumption is particularly promising for sustainable and efficient communication technologies. The theoretical analysis demonstrates that the hybrid architecture achieves superior performance not through brute-force computation, but through intelligent matching of computational primitives to the physical characteristics of optical channels. Moreover, it is a candidate for efficient deployment on neuromorphic hardware, which will be explored in future studies to assess its feasibility for real-time compensation of nonlinear impairments in optical networks.

5.6 Conclusions and Future Work

This chapter has presented a novel approach to optical signal equalisation using spiking neural networks, specifically through the adaptation of the RWKV architecture to create the SNN-RWKV equaliser. Through comprehensive simulation and analysis, we have demonstrated that this approach offers a promising alternative to conventional ANN-based methods for nonlinearity compensation in high-speed optical communication systems.

The SNN-RWKV equaliser combines the parallelizable training efficiency of transformer-based models with the inference efficiency of recurrent architectures, while leverageing the energy advantages of spike-based computation. Our experimental results on a 500 km TWC fibre link with 16-QAM modulation demonstrate that this approach not only matches but exceeds the performance of state-of-the-art ANN-based equalisers, while offering significant reductions in computational complexity and energy consumption.

Key findings from our investigation include:

1. The SNN-RWKV architecture achieves a Q-factor improvement of up to 0.25 dB over traditional ANN approaches across a wide range of launch powers, with the most substantial advantages observed in the nonlinear regime.
2. At the optimal launch power of -2 dBm, the SNN-RWKV equaliser achieves a Q-factor of 4.7 dB while consuming only 1.64×10^6 pJ of energy per processing instance, representing an 8.4% energy reduction compared to CNN-MLP and an 87.7% reduction compared to CNN-biLSTM.
3. The sparse, event-driven nature of spike computation, with an average spike rate of only 9.1%, enables significant computational savings by transforming dense matrix multiplications into sparse additions.
4. The linear computational scaling of the RWKV architecture with respect to sequence length eliminates the quadratic complexity bottleneck of traditional attention mechanisms, making it particularly suitable for processing long symbol sequences in optical communications.

These results suggest that spike-based architectures, particularly when combined with efficient attention mechanisms like those in RWKV, represent a promising direction for addressing the computational challenges in next-generation high-speed optical systems. The SNN-RWKV approach effectively balances the competing demands of equalisation performance and implementation efficiency, offering a pathway toward real-time nonlinearity compensation in practical systems.

5.6.1 Limitations and Future Research Directions

Despite the promising results, several limitations and avenues for future research remain:

1. **Hardware Implementation**: While our analysis provides theoretical energy consumption estimates, practical implementation on dedicated hardware (both digital Application-Specific Integrated Circuit (ASIC) and neuromorphic platforms) is needed to validate these projections. Future work should focus on hardware-software co-design to optimise the SNN-RWKV architecture for specific target platforms.
2. **Scaling to Higher Modulation Orders**: The current study focuses on 16-QAM modulation, but future optical systems will increasingly employ higher-order modulation formats. Additional research is needed to assess how the advantages of SNN-RWKV scale with increasing modulation complexity.
3. **Adaptation to Channel Variations**: Real-world optical channels exhibit temporal variations that may require online adaptation. Future work should investigate efficient online learning

mechanisms for SNN-RWKV that maintain the computational efficiency advantages during adaptation.

4. **Multi-Channel Systems**: This study focused on single-channel transmission. Extending the approach to WDM systems with inter-channel nonlinearities represents an important next step.

5. **Reduced-Precision Training**: While the inference process benefits from binary spike representation, training still relies on full-precision gradient computation. Investigating reduced-precision training methods specific to SNN-RWKV could further enhance the approach's efficiency.

5.6.2 Broader Impact

The SNN-RWKV equaliser represents a significant step toward bridging the gap between the theoretical promise of neuromorphic computing and practical applications in high-performance communication systems. By demonstrating that spike-based computation can not only match but exceed the performance of conventional approaches in a demanding signal processing task, this work contributes to the ongoing paradigm shift toward more brain-inspired computing architectures [198, 207].

As optical communication systems continue to push toward higher data rates, higher spectral efficiencies, and longer transmission distances, the computational demands of nonlinearity compensation will only increase. The approach presented in this chapter offers a sustainable path forward, leverageing the inherent efficiency of event-driven, sparse computation to enable advanced signal processing with reasonable power budgets [6, 33, 225, 228, 237].

In conclusion, the SNN-RWKV equaliser demonstrates that neuromorphic approaches to signal processing can transcend their theoretical elegance to deliver practical performance advantages in real-world engineering applications. This work lays the foundation for further exploration of spike-based computation in communications and signal processing, potentially leading to a new generation of energy-efficient, high-performance systems inspired by the computational principles of biological neural systems.

Chapter 6

Conclusion and Further Work

6.1 Summary and Conclusion

This thesis explores innovative strategies to tackle the persistent challenge of nonlinear impairments in optical communication systems. The rapid growth of internet traffic driven by emerging technologies such as AI, HPC, and cloud-based services demands continuous advancements in data transmission capabilities. Optical fibre systems, although capable of high capacity and efficiency, encounter fundamental limitations due to nonlinear distortions, notably those arising from the Kerr effect. Traditional methods for addressing these nonlinearities, such as DBP and Volterra equalisation, provide significant improvements but face substantial barriers regarding computational complexity and energy efficiency.

To overcome these barriers, this work presents a comprehensive investigation into advanced neural network-based equalisers, emphasizing architectures tailored specifically for optical communication systems. By harnessing the pattern-recognition power and adaptability of artificial neural networks, this research has demonstrated superior compensation for nonlinear distortions compared to conventional techniques. The developed solutions include novel complex-valued neural networks, which directly handle complex optical signals, significantly improving computational efficiency and equalisation performance.

Further, recognizing the stringent energy constraints and computational limitations in practical scenarios, the thesis delves into sophisticated methods to optimise the complexity-performance trade-off. Techniques such as model pruning, weight clustering, and quantisation are systematically studied, demonstrating that neural networks can achieve substantial reductions in computational resources and memory requirements without compromising the equalisation quality. Additionally, the practical feasibility of these optimised neural network equalisers was confirmed by deploying them on resource-constrained hardware, showcasing their suitability for real-time optical communication applications.

A particularly innovative aspect of this research is the exploration of neuromorphic computing, notably SNNs and Photonic Reservoir Computing, which represent promising directions for future optical signal processing. These methods leverage bio-inspired, event-driven processing paradigms to achieve remarkable energy efficiency and reduced latency, crucial for real-time optical equalisation tasks.

Furthermore, this thesis pioneers in evaluating the noise robustness of neural network implementations in photonic hardware, introducing noise-aware training techniques and ensemble-based strategies—such as crowd equalisation—that significantly enhance system resilience under realistic operating conditions. These contributions bridge the gap between theoretical advancements and practical implementations, providing viable strategies to tackle real-world challenges in optical communications.

In summary, the work presented herein lays critical groundwork for next-generation optical communication systems by integrating machine learning, neuromorphic computing, and photonic technologies. The developed solutions not only push the boundaries of current technology but also provide practical insights and methodologies that bring the theoretical potential of advanced neural network-based equalisers closer to real-world industrial applications.

6.2 Future work

This thesis laid the foundation for leverageing neural networks, particularly neuromorphic photonic architectures, in addressing the critical issue of nonlinear impairments in optical communication systems. Despite the significant advancements demonstrated, numerous practical challenges remain to be explored and overcome before these technologies can achieve widespread adoption.

A promising direction for future research is the extensive practical and experimental evaluation of photonic neuromorphic computing systems in more challenging and realistic communication scenarios. Although photonic neural networks offer immense potential in terms of ultra-fast processing and high parallelism, significant hurdles related to device imperfections, noise, and limited scalability persist. Addressing these challenges requires careful characterisation and mitigation strategies, such as integrated calibration techniques, on-chip stabilisation, and robust packageing [41, 248].

Future investigations should also prioritize bridging the gap between theoretical models and real-world implementations. This includes exploring innovative training methods tailored explicitly for photonic hardware, such as hardware-aware training and hybrid in-situ optimisation techniques. Given that traditional backpropagation algorithms struggle to address hardware-induced constraints and noise sensitivity, alternative training strategies—such as evolutionary algorithms or particle swarm optimisation—could be further refined and experimentally validated [225, 249].

Moreover, scaling photonic neural network architectures remains a critical issue. Achieving large-scale integrated systems with numerous photonic neurons and weights will require improvements in fabrication accuracy, loss management, and integration density. Incorporating robust techniques like erbium-doped waveguide amplifiers or Raman amplification directly on-chip could compensate for optical losses and facilitate larger network sizes. Additionally, future work should explore the development of novel nonlinear activation elements compatible with photonic integration, addressing the current limitations in nonlinear operation within fully optical networks [225].

Finally, a critical comparison between photonic and electronic ANN-based equalisers in practical, industrially relevant conditions will remain an essential area of future research. Evaluating these systems under stringent conditions—including extreme environmental variations, high noise levels, and complex channel impairments—will provide valuable insights into their reliability and robustness. Such studies will also guide the development of hybrid solutions that leverage the strengths of both photonic and electronic technologies, potentially leading to optimised architectures combining the speed and parallelism of photonics with the adaptability and precision of electronic systems [225, 226].

In conclusion, the future research outlined here aims to bridge the gap between theoretical advancements and practical deployments, paving the way for robust, energy-efficient, and high-performance neuromorphic equalisation systems for next-generation optical communications.

6.3 Publications

P1 Ron, D. A., Kamalian-Kopae, M., & Turitsyn, S. (2021, June). Noise-resistant optical implementation of analogue neural networks. In *European Quantum Electronics Conference* (p. jsiv_3_2). Optica Publishing Group.

P2 Ron, D. A., Freire, P. J., Prilepsky, J. E., Kamalian-Kopae, M., Napoli, A., & Turitsyn, S. K. (2022). Experimental implementation of a neural network optical channel equaliser in restricted hardware using pruning and quantisation. *Scientific Reports*, 12(1), 8713.

P3 Ron, D. A., Nurlybayeva, K., Kamalian-Kopae, M., Ali, A. A., Turitsyna, E., & Turitsyn, S. (2022, September). On the impact of the optical phase conjugation on the computational complexity of neural network-based equalisers. In *European Conference and Exhibition on Optical Communication* (pp. We5-29). Optica Publishing Group.

P4 Nurlybayeva, K., Ron, D. A., Kamalian-Kopae, M., Turitsyna, E., & Turitsyn, S. (2022, October). Noise-Resistant Crowd Equalisation for Optical Communication Systems Based on Machine Learning. In *Frontiers in Optics* (pp. FM3D-2). Optica Publishing Group.

P5 Freire, P. J., Napoli, A., Spinnler, B., Anderson, M., Ron, D. A., Schairer, W., ... & Prilepsky, J. E. (2023). Reducing computational complexity of neural networks in optical channel equalisation: From concepts to implementation. *Journal of Lightwave Technology*, 41(14), 4557-4581.

P6 Ron, D. A. (2023, May). Combination of Optical Phase Conjugation and Advanced Pruning Techniques to reduce the Computational Complexity of Neural Network-Based Equalisers. In *CLEO: Science and Innovations* (pp. SM3I-6). Optica Publishing Group.

P7 Ron, D. A. (2023, September). On the Noise Robustness of Analogue Complex-Valued Neural Networks. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases* (pp. 37-50). Cham: Springer Nature Switzerland.

P8 Manuylovich, E., Argüello Ron, D., Kamalian-Kopae, M., & Turitsyn, S. K. (2024). Robust neural networks using stochastic resonance neurons. *Communications Engineering*, 3(1), 169.

P9 Ron, D. A. (2024, September). Adapting Spiking Neural Networks for Enhanced Optical Signal Equalisation in Communication Systems. In *ECOC 2024; 50th European Conference on Optical Communication* (pp. 906-909). VDE.

Bibliography

[1] Morteza Kamalian-Kopae, Abdallah A Ali, Karina Nurlybayeva, A Ellis, and S Turitsyn. Neural network-enhanced optical phase conjugation for nonlinearity mitigation. In *2022 Optical Fiber Communications Conference and Exhibition (OFC)*, pages 1–3. IEEE, 2022.

[2] Egor Manuylovich, Diego Argüello Ron, Morteza Kamalian-Kopae, and Sergei K Turitsyn. Robust neural networks using stochastic resonance neurons. *Communications Engineering*, 3(1):169, 2024.

[3] Peter J Winzer. High-spectral-efficiency optical modulation formats. *Journal of Lightwave Technology*, 30(24):3824–3835, 2012.

[4] Xiang Zhou, Lynn E Nelson, Peter Magill, Rennie Isaac, Benyuan Zhu, Pierre Borel, Ken Carlson, and David W Peckham. High spectral efficiency 400 gb/s transmission using pdm time-domain hybrid 32–64 qam and training-assisted carrier recovery. In *2013 Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference (OFC/NFOEC)*, pages 1–3. IEEE, 2013.

[5] Stephan Lars Olsson, Junho Cho, Sethumadhavan Chandrasekhar, Xiang Chen, Evan C Burrows, and Peter J Winzer. Record-high 17.3-bit/s/hz spectral efficiency transmission over 50 km using probabilistically shaped pdm 4096-qam. *Journal of Lightwave Technology*, 36(22):5063–5069, 2018.

[6] Seb J Savory. Digital coherent optical receivers: Algorithms and subsystems. *IEEE Journal of Selected Topics in Quantum Electronics*, 16(5):1164–1179, 2010.

[7] Govind P. Agrawal. *Fiber-optic communication systems*. John Wiley & Sons, New York, 4th edition, 2012.

[8] Yuxin Wang, Yuhan Chen, Zeyu Li, Zhenheng Tang, Rui Guo, Xin Wang, Qiang Wang, Amelie Chi Zhou, and Xiaowen Chu. Towards efficient and reliable llm serving: A real-world workload study. *arXiv preprint arXiv:2401.17644*, 2024.

[9] World Bank. *Digital Progress and Trends Report 2023*. International Bank for Reconstruction and Development / The World Bank, Washington, DC, 2024.

[10] Jia Wei, Mo Chen, Longxiang Wang, Pei Ren, Yujia Lei, Yuqi Qu, Qiyu Jiang, Xiaoshe Dong, Weiguo Wu, Qiang Wang, et al. Status, challenges and trends of data-intensive supercomputing. *CCF Transactions on High Performance Computing*, 4(2):211–230, 2022.

[11] David Patterson, Joseph Gonzalez, Quoc Le, Chen Liang, Lluis-Miquel Munguia, Daniel Rothchild, David So, Maud Texier, and Jeff Dean. Carbon emissions and large neural network training. *arXiv preprint arXiv:2104.10350*, 2021.

[12] Goldman Sachs. Ai to drive 165% increase in data center power demand by 2030. Goldman Sachs Insights, 2024.

[13] Jinbo Wen, Ruichen Zhang, Dusit Niyato, Jiawen Kang, Hongyang Du, Yang Zhang, and Zhu Han. Generative ai for low-carbon artificial intelligence of things with large language models. *IEEE Internet of Things Magazine*, 8(1):82–91, 2024.

[14] Data centres will use twice as much energy by 2030 — driven by ai. *Nature*, 2025.

[15] Semiconductor Engineering. Ai drives need for optical interconnects in data centers. Semiconductor Engineering, 2024.

[16] NAND Research. Ofc 2025: Optical interconnects take center stage in the ai-first data center. NAND Research, 2025.

[17] David Mytton. Data center energy and ai in 2025. *Development & Sustainability*, 2025.

[18] Christoforos Kachris and Ioannis Tomkos. A survey on optical interconnects for data centers. *IEEE Communications Surveys & Tutorials*, 14(4):1021–1036, 2012.

[19] International Energy Agency. What the data centre and ai boom could mean for the energy sector. IEA Analysis, 2024.

[20] Peter J Winzer, David T Neilson, and Andrew R Chraplyvy. Fiber-optic transmission and networking: the previous 20 and the next 20 years. *Optics Express*, 26(18):24190–24239, 2018.

[21] Corning. 2024 data center trends and industry predictions: Ai and edge computing. Corning Optical Communications, 2024.

[22] Cisco. Cisco visual networking index: Forecast and trends, 2017–2022. *White Paper*, 2019.

[23] Emmanuel Desurvire. *Erbium-doped fiber amplifiers: principles and applications*. John Wiley & Sons, 2002.

[24] Ren'e-Jean Essiambre, Gerhard Kramer, Peter J Winzer, Gerard J Foschini, and Bernhard Goebel. Capacity limits of optical fiber networks. *Journal of Lightwave Technology*, 28(4):662–701, 2010.

[25] Govind P Agrawal. *Nonlinear fiber optics*. Academic press, 2013.

[26] Andrew D. Ellis, Jian Zhao, and David Cotter. Approaching the non-linear shannon limit. *Journal of Lightwave Technology*, 28(4):423–433, 2010.

[27] Jia-Xing Cai, Yi Sun, Haoshuo Zhang, Hussam G Batshon, Mikhail V Mazurczyk, Oleg V Sinkin, David G Foursa, and Alexey N Pilipetskii. 49.3 tb/s transmission over 9100 km using c+ l edfa and 54 tb/s transmission over 9150 km using hybrid-raman edfa. *Journal of Lightwave Technology*, 33(13):2724–2734, 2015.

[28] Monica Ionescu, Daniel Petrache, Lee Forced, Ian Giles, and David J Richardson. Ultra wide band, low loss, step index fiber for c and l band transmission. *Journal of Lightwave Technology*, 37(21):5407–5413, 2019.

[29] René-Jean Essiambre and Robert W Tkach. Capacity trends and limits of optical communication networks. *Proceedings of the IEEE*, 100(5):1035–1055, 2012.

[30] Yann Frignac, Dylan Le Gac, Abel Lorences-Riesgo, Loïg Godard, Salma Escobar Landero, Xiaohui Zhao, Erwan Pincemin, Bertrand Le Guyader, Nicolas Brochier, Qiang Guo, et al. Record 158.4 tb/s transmission over 2×60 km field smf using s+c+l 18 thz-bandwidth lumped amplification. In *49th European Conference on Optical Communications (ECOC 2023)*, volume 2023 of *IET Conference Proceedings*, pages 550–553, Glasgow, United Kingdom, October 2023. The Institution of Engineering and Technology (IET).

[31] Juan Diego Ania-Castanon. Ultralong raman fiber lasers as virtually lossless optical media. *Physical review letters*, 96(2):023902, 2006.

[32] Ezra Ip and Joseph M Kahn. Compensation of dispersion and nonlinear impairments using digital backpropagation. *Journal of Lightwave Technology*, 26(20):3416–3425, 2008.

[33] Antonio Napoli, Pablo Wilke Berenguer, Talha Rahman, Ginni Khanna, Mahdi M. Mezghanni, Lennart Gardian, Emilio Riccardi, Anna Chiadò Piat, Stefano Calabrò, Stefanos Dris, André Richter, Johannes Karl Fischer, Bernd Sommerkorn-Krombholz, and Bernhard Spinnler. Digital pre-compensation techniques enabling high-capacity bandwidth variable transponders. *Optics Communications*, 409:52 – 65, 2018. Advances in modulation and DSP for optical transmission systems.

[34] Andrew D Ellis, Mary E McCarthy, Mohammad AZ Al Khateeb, Mariia Sorokina, and Nick J Doran. 4 tb/s transmission reach enhancement using 10× 400 gb/s super-channels and polarization insensitive dual band optical phase conjugation. *Journal of Lightwave Technology*, 34(8):1717–1723, 2016.

[35] Darko Zibar, Molly Piels, Rasmus Jones, and Christian G Sch"affer. Machine learning techniques in optical communication. *Journal of Lightwave Technology*, 34(6):1442–1452, 2016.

[36] David J Richardson, John M Fini, and Lars E Nelson. Space-division multiplexing in optical fibres. *Nature photonics*, 7(5):354–362, 2013.

[37] Pedro J Freire, Antonio Napoli, Bernhard Spinnler, Nelson Costa, Sergei K Turitsyn, and Jaroslaw E Prilepsky. Neural networks-based equalizers for coherent optical transmission: Caveats and pitfalls. *IEEE Journal of Selected Topics in Quantum Electronics*, 28(4):1–23, 2022. Special issue: Machine Learning in Photonic Communications and Measurement Systems.

[38] Apostolos Argyris. Photonic neuromorphic technologies in optical communications. *Nanophotonics*, 11(5):897–916, 2022.

[39] Min Yan, Can Huang, Peter Bienstman, Peter Tino, Wei Lin, and Jie Sun. Emerging opportunities and challenges for the future of reservoir computing. *Nature Communications*, 15(1):2056, 2024.

[40] Kun Liao, Tianxiang Dai, Qiuchen Yan, Xiaoyong Hu, and Qihuang Gong. Integrated photonic neural networks: Opportunities and challenges. *ACS Photonics*, 10(7):2001–2010, 2023.

[41] Luís CB Silva, Pablo RN Marciano, Maria J Pontes, Maxwell E Monteiro, Paulo SB André, and Marcelo EV Segatto. Integrated photonic neural networks for equalizing optical communication signals: A review. In *Photonics*, volume 12, page 39. MDPI, 2025.

[42] Aimee J. S. Arnold, Haralampos G. Hamann, and Seyedeh Alaleh Razavi. Spiking neural networks for energy-efficient computation: Design and optimization. *Nature Communications*, 14(1):1–12, 2023.

[43] Sami Mumtaz, René-Jean Essiambre, and Govind P Agrawal. Nonlinear propagation in multimode and multicore fibers: generalization of the manakov equations. *Journal of Lightwave Technology*, 31(3):398–406, 2012.

[44] René-Jean Essiambre, Gerhard Kramer, Peter J. Winzer, Gerard J. Foschini, and Bernhard Goebel. Capacity limits of optical fiber networks. *Journal of Lightwave Technology*, 28(4):662–701, 2010.

[45] Govind P Agrawal. Nonlinear science at the dawn of the 21st century. *Lecture Notes in Physics*, 542:195–211, 2000.

[46] AJ Barlow, JJ Ramskov-Hansen, and DN Payne. Birefringence and polarization mode-dispersion in spun single-mode fibers. *Applied Optics*, 20(17):2962–2968, 1981.

[47] Md Saifuddin Faruk and Seb J Savory. Digital signal processing for coherent transceivers employing multilevel formats. *Journal of Lightwave Technology*, 35(5):1125–1141, 2017.

[48] Benjamin J Puttnam, Ruben S Luís, Yetian Huang, Ian Phillips, Dicky Chung, Nicolas K Fontaine, Budsara Boriboon, Georg Rademacher, Mikael Mazur, Lauren Dallachiesa, et al. 321 tb/s e/s/c/l-band transmission with e-band bismuth-doped fiber amplifier and optical processor. *Journal of Lightwave Technology*, 2024.

[49] Zabih Ghassemlooy, Wasiu Popoola, and Sujan Rajbhandari. *Optical wireless communications: system and channel modelling with Matlab®*. CRC press, 2019.

[50] Stanislav Dimitrov and Harald Haas. Signal shaping for mitigation of fiber nonlinearities. *Optics Express*, 20(1):3–18, 2012.

[51] Chris RS Fludger, Thomas Duthel, Dirk Van den Borne, Christoph Schulien, Ernst-Dieter Schmidt, Torsten Wuth, Jonas Geyer, Erik De Man, Giok-Djan Khoe, and Huug de Waardt. Coherent equalization and polmux-rz-dqpsk for robust 100-ge transmission. *Journal of lightwave technology*, 26(1):64–72, 2008.

[52] Pedro J Freire, Antonio Napoli, Diego Argüello Ron, Bernhard Spinnler, Michael Anderson, Wolfgang Schairer, Thomas Bex, Nelson Costa, Sergei K Turitsyn, and Jaroslaw E Prilepsky. Reducing computational complexity of neural networks in optical channel equalization: From concepts to implementation. *Journal of Lightwave Technology*, 2023.

[53] Kyo Inoue and Hiromu Toba. Theoretical evaluation of error rate degradation due to fiber four-wave mixing in multichannel fsk heterodyne envelope detection transmissions. *Journal of lightwave technology*, 10(3):361–366, 1992.

[54] Ian Coddington, William C Swann, Luca Lorini, James C Bergquist, Y Le Coq, Christopher W Oates, Q Quraishi, KS Feder, JW Nicholson, Paul S Westbrook, et al. Coherent optical link over hundreds of metres and hundreds of terahertz with subfemtosecond timing jitter. *Nature Photonics*, 1(5):283–287, 2007.

[55] Simone Gaiarin, Auro Michele Perego, E Porto da Silva, Francesco Da Ros, and Darko Zibar. Dual-polarization nonlinear fourier transform-based optical communication system. *Optica*, 5(3):263–270, 2018.

[56] Konstantinos Christodoulopoulos, Ioannis Tomkos, and Emmanuel A Varvarigos. Elastic bandwidth allocation in flexible ofdm-based optical networks. *Journal of Lightwave Technology*, 29(9):1354–1366, 2011.

[57] Tianhua Xu, Boris Karanov, Nikita A Shevchenko, Domaniç Lavery, Gabriele Liga, Robert I Killey, and Polina Bayvel. Digital nonlinearity compensation in high-capacity optical communication systems considering signal spectral broadening effect. *Scientific reports*, 7(1):12986, 2017.

[58] Claude E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, 1948.

[59] Andrew Chralyvy. Plenary paper: The coming capacity crunch. In *2009 35th European Conference on Optical Communication*, pages 1–1. IEEE, 2009.

[60] Paul J. Roberts, Lidia Galdino, Alex Alvarado, et al. Communication networks beyond the capacity crunch. *Philosophical Transactions of the Royal Society A*, 374(2062):20150191, 2016.

[61] A.D. Ellis, J. Zhao, and D. Cotter. Approaching the non-linear shannon limit. *Journal of Lightwave Technology*, 28(4):423–433, 2010.

[62] Antonio Napoli, Zied Maalej, Vincent AJM Sleiffer, Maxim Kuschnerov, Danish Rafique, Erik Timmers, Bernhard Spinnler, Talha Rahman, Leonardo Didier Coelho, and Norbert Hanik. Reduced complexity digital back-propagation methods for optical communication systems. *Journal of lightwave technology*, 32(7):1351–1362, 2014.

[63] Fernando Pedro Guiomar, Jacklyn Dias Reis, António Luís Teixeira, and Armando Nolasco Pinto. Digital postcompensation using volterra series transfer function. *IEEE Photonics Technology Letters*, 23(19):1412–1414, 2011.

[64] Felix Frey, Lutz Molle, Robert Emmerich, Colja Schubert, Johannes K Fischer, and Robert FH Fischer. Single-step perturbation-based nonlinearity compensation of intra-and inter-subcarrier nonlinear interference. In *2017 European Conference on Optical Communication (ECOC)*, pages 1–3. IEEE, 2017.

[65] Andrew D Ellis, Mingming Tan, Md Asif Iqbal, Mohammad Ahmad Zaki Al-Khateeb, Vladimir Gordienko, Gabriel Saavedra Mondaca, Simon Fabbri, Marc FC Stephens, Mary E McCarthy, Andreas Perentos, et al. 4 tb/s transmission reach enhancement using 10× 400 gb/s super-channels and polarization insensitive dual band optical phase conjugation. *Journal of lightwave technology*, 34(8):1717–1723, 2016.

[66] Sergei K. Turitsyn, Jaroslaw E. Prilepsky, Son T. Le, Sander Wahls, Leonid L. Frumin, Morteza Kamalian, and Stanislav A. Derevyanko. Nonlinear frequency division multiplexing. *Optica*, 4(3):307–322, 2017.

[67] Gabriele Liga, Tobias Eriksson, Alex Alvarado, Erik Agrell, and Magnus Karlsson. Machine learning approaches for nonlinear equalization in optical transmission systems. *IEEE Photonics Technology Letters*, 32(3):189–192, 2020.

[68] Diego Argüello Ron, Pedro J Freire, Jaroslaw E Prilepsky, Morteza Kamalian-Kopae, Antonio Napoli, and Sergei K Turitsyn. Experimental implementation of a neural network optical channel equalizer in restricted hardware using pruning and quantization. *Scientific Reports*, 12(1):1–14, 2022.

[69] Stavros Deligiannidis, Charis Mesaritakis, and Adonis Bogris. Performance and complexity analysis of bi-directional recurrent neural network models versus volterra nonlinear equalizers in digital coherent systems. *Journal of Lightwave Technology*, 39(18):5791–5798, 2021.

[70] Danish Rafique and Andrew D. Ellis. Compensation of intra-channel nonlinear fibre impairments using simplified digital back-propagation algorithm. *Optics Express*, 19(10):9453–9460, 2011.

[71] Nadezhda Semenova, Xavier Porte, Maxime Jacquot, Laurent Larger, and Daniel Brunner. Noise propagation in feedforward and reservoir neural networks. In *Emerging Topics in Artificial Intelligence 2020*, volume 11469, page 114690J. SPIE, 2020.

[72] Nadezhda Semenova, Xavier Porte, Louis Andreoli, Maxime Jacquot, Laurent Larger, and Daniel Brunner. Fundamental aspects of noise in analog-hardware neural networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29(10):103128, 2019.

[73] Diego Argüello Ron, Morteza Kamalian-Kopae, and Sergei Turitsyn. Noise-resistant optical implementation of analogue neural networks. In *European Quantum Electronics Conference*, page jsiv_3_2. Optica Publishing Group, 2021.

[74] Pedro J Freire, Vladislav Neskornuik, Antonio Napoli, Bernhard Spinnler, Nelson Costa, Ginni Khanna, Emilio Riccardi, Jaroslaw E Prilepsky, and Sergei K Turitsyn. Complex-valued neural network design for mitigation of signal distortions in optical links. *Journal of Lightwave Technology*, 39(6):1696–1705, 2020.

[75] Kavita Burse, Ram Narayan Yadav, and SC Shrivastava. Channel equalization using neural networks: A review. *IEEE transactions on systems, man, and cybernetics, Part C (Applications and Reviews)*, 40(3):352–357, 2010.

[76] Nan Chi, Yiheng Zhao, Meng Shi, Peng Zou, and Xingyu Lu. Gaussian kernel-aided deep neural network equalizer utilized in underwater pam8 visible light communication system. *Optics express*, 26(20):26700–26712, 2018.

[77] John C Cartledge, Fernando P Guiomar, Frank R Kschischang, Gabriele Liga, and Metodi P Yankov. Digital signal processing for fiber nonlinearities. *Optics express*, 25(3):1916–1936, 2017.

[78] Ezra Ip and Joseph M. Kahn. Compensation of dispersion and nonlinear impairments using digital backpropagation. *Journal of Lightwave Technology*, 26(20):3416–3425, 2008.

[79] Marco Secondini and Enrico Forestieri. On the nonlinear threshold of optical communication systems. *Optics Express*, 23(4):4841–4851, 2015.

[80] Ronen Dar and Peter J Winzer. On the limits of digital back-propagation in fully loaded wdm systems. *IEEE Photonics Technology Letters*, 28(11):1253–1256, 2016.

[81] Kumar V. Peddanarappagari and Maite Brandt-Pearce. Volterra series transfer function of single-mode fibers. *Journal of Lightwave Technology*, 15(12):2232–2241, 1997.

[82] Erick P. Silva, Maria J. Pontes, Helder M. Salgado, and Jose R.M. Barbosa. Nonlinear equalization based on support vector machines for optical communications systems. *IEEE Photonics Technology Letters*, 24(12):1003–1005, 2012.

[83] Yuyang Pan, Jian Ye, and Dayou Qian. Volterra equalizer for single mode fiber communication systems. *IEEE Photonics Technology Letters*, 23(17):1251–1253, 2011.

[84] S. J. Savory. Digital filters for coherent optical receivers. *Optics Express*, 16(2):804–817, 2008.

[85] K. Kudo et al. Coherent optical single-carrier transmission using overlap frequency-domain equalization. *Journal of Lightwave Technology*, 27(16):3721–3728, 2009.

[86] L. Guiomar and others. Mitigation of intra-channel nonlinearities using a frequency-domain volterra series equalizer. *Optics Express*, 20(2):1360–1370, 2012.

[87] John J. Shynk. Frequency-domain and multirate adaptive filtering. In *IEEE Signal Processing Magazine*, volume 9, pages 14–37, 1992.

[88] John G. Proakis and Dimitris K. Manolakis. *Digital Signal Processing: Principles, Algorithms, and Applications*. Prentice Hall / Pearson, 4th edition, 2006.

[89] Herbert Jaeger. The "echo state" approach to analysing and training recurrent neural networks. Technical Report 148, German National Research Center for Information Technology GMD Technical Report, Bonn, Germany, 2001.

[90] Mantas Lukoševičius and Herbert Jaeger. Reservoir computing approaches to recurrent neural network training. *Computer Science Review*, 3(3):127–149, 2009.

[91] S Watanabe and T Chikama. Cancellation of four-wave mixing in multichannel fibre transmission by midway optical phase conjugation. *Electronics Letters*, 30(14):1156–1157, 1994.

[92] A. D. Ellis and F. C. G. Gunning. Spectral density enhancement using coherent wdm. *IEEE Photonics Technology Letters*, 17(2):504–506, 2015.

[93] S. Yoshima, Y. Matsumoto, K. Tanaka, H. Takara, K. Kasai, and M. Nakazawa. Mitigation of nonlinear effects on wdm qam signals enabled by optical phase conjugation with efficient bandwidth utilization. *Optics Express*, 25(18):20808–20818, 2017.

[94] Yukui Yu, Wei Wang, Paul D Townsend, and Jian Zhao. Modified phase-conjugate twin wave schemes for spectral efficiency enhancement. In *2015 European Conference on Optical Communication (ECOC)*, pages 1–3. IEEE, 2015.

[95] Tsuyoshi Yoshida, Takashi Sugihara, Kazuyuki Ishida, and Takashi Mizuochi. Spectrally-efficient dual phase-conjugate twin waves with orthogonally multiplexed quadrature pulse-shaped signals. In *Optical Fiber Communication Conference*, pages M3C–6. Optica Publishing Group, 2014.

[96] A. Napoli, A. D. Ellis, and et al. Reduced-complexity digital back-propagation: Achievable through opc pre-compensation. *Optics Express*, 22(24):30044–30056, 2014.

[97] L. Guiomar, D. Zibar, et al. Optical phase conjugation assisted digital nonlinearity compensation. *Journal of Lightwave Technology*, 34(22):5112–5120, 2016.

[98] G. Liga, A. Napoli, et al. Hybrid optical–digital nonlinearity compensation: Achievable gains and system design considerations. *Optics Express*, 25(6):6329–6344, 2017.

[99] A. D. Ellis and Z. Zhang. Opportunities for optical phase conjugation in superchannel transmission systems. *IEEE Photonics Technology Letters*, 31(23):1874–1877, 2019.

[100] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. MIT press, Cambridge, MA, 2016.

[101] Pedro J. Freire, Daniel Abode, Jaroslaw E. Prilepsky, Nelson Costa, Bernhard Spinnler, Antonio Napoli, and Sergei K. Turitsyn. Transfer learning for neural networks-based equalizers in coherent optical systems. *Journal of Lightwave Technology*, 39(21):6733–6745, 2021.

[102] Rick M Bütler, Christian Häger, Henry D Pfister, Gabriele Liga, and Alex Alvarado. Model-based machine learning for joint digital backpropagation and pmd compensation. *Journal of Lightwave Technology*, 39(4):949–959, 2020.

[103] Pedro J Freire, Yevhenii Osadchuk, Bernhard Spinnler, Antonio Napoli, Wolfgang Schairer, Nelson Costa, Jaroslaw E Prilepsky, and Sergei K Turitsyn. Performance versus complexity study of neural network equalizers in coherent optical systems. *Journal of Lightwave Technology*, 39(19):6085–6096, 2021.

[104] Oleg Sidelnikov, Alexey Redyuk, and Stylianos Sygletos. Equalization performance and complexity analysis of dynamic deep neural networks in long haul transmission systems. *Optics express*, 26(25):32765–32776, 2018.

[105] Amirhossein Ron, Danshi Wang, Gabriele Liga, Antonio Napoli, and Polina Bayvel. Impact of neural network equalization complexity on fiber nonlinearity compensation. *IEEE Photonics Technology Letters*, 34(21):1147–1150, 2022.

[106] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015.

[107] Shimpei Yamamoto, Etsushi Yamazaki, et al. Neural network approaches for improving optical fiber communication performance. *IEEE Communications Magazine*, 59(11):61–67, 2021.

[108] Sebastian Ruder. An overview of gradient descent optimization algorithms. *arXiv preprint arXiv:1609.04747*, 2016.

[109] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, 2015. Replaces arXiv:1412.6980.

[110] Zijun Zhang. Improved adam optimizer for deep neural networks. *IEEE/ACM 26th International Symposium on Quality of Service (IWQoS)*, pages 1–2, 2018.

[111] Matthias Feurer and Frank Hutter. Hyperparameter optimization. *Automated machine learning: Methods, systems, challenges*, pages 3–33, 2019.

[112] Jia Wu, Xiu-Yun Chen, Hao Zhang, Li-Dong Xiong, Hang Lei, and Si-Hao Deng. Hyperparameter optimization for machine learning models based on bayesian optimization. *Journal of Electronic Science and Technology*, 17(1):26–40, 2019.

[113] Jasper Snoek, Hugo Larochelle, and Ryan P. Adams. Practical bayesian optimization of machine learning algorithms. In *Advances in Neural Information Processing Systems*, volume 25, pages 2951–2959, 2012.

[114] Yoshua Bengio. Gradient-based optimization of hyperparameters. *Neural computation*, 12(8):1889–1900, 2000.

[115] Yuqiao Liu, Yanan Sun, Bing Xue, Mengjie Zhang, Gary G Yen, and Kay Chen Tan. A survey on evolutionary neural architecture search. *IEEE transactions on neural networks and learning systems*, 34(2):550–570, 2021.

[116] Torsten Hoefler, Dan Alistarh, Tal Ben-Nun, Nikoli Dryden, and Alexandra Peste. Sparsity in deep learning: Pruning and growth for efficient inference and training in neural networks. *Journal of Machine Learning Research*, 22(241):1–124, 2021.

[117] Zhichao Lu, Ran Cheng, Yaochu Jin, Kay Chen Tan, and Kalyanmoy Deb. Neural architecture search as multiobjective optimization benchmarks: Problem formulation and performance assessment. *IEEE transactions on evolutionary computation*, 28(2):323–337, 2023.

[118] Pengzhen Ren, Yun Xiao, Xiaojun Chang, Po-Yao Huang, Zhihui Li, Xiaojiang Chen, and Xin Wang. A comprehensive survey of neural architecture search: Challenges and solutions. *ACM Computing Surveys (CSUR)*, 54(4):1–34, 2021.

[119] Pedro J Freire, Daniel Abode, Jaroslaw E Prilepsky, Nelson Costa, Bernhard Spinnler, Antonio Napoli, and Sergei K Turitsyn. Transfer learning for neural networks-based equalizers in coherent optical systems. *Journal of Lightwave Technology*, 39(21):6733–6745, 2021.

[120] Pedro Freire, Sasipim Srivallapanondh, Antonio Napoli, Jaroslaw E Prilepsky, and Sergei K Turitsyn. Computational complexity evaluation of neural network applications in signal processing. *arXiv preprint arXiv:2206.12191*, 2022.

[121] Zewen Li, Fan Liu, Wenjie Yang, Shouheng Peng, and Jun Zhou. A survey of convolutional neural networks: analysis, applications, and prospects. *IEEE transactions on neural networks and learning systems*, 33(12):6999–7019, 2021.

[122] ChiYan Lee, Hideyuki Hasegawa, and Shangce Gao. Complex-valued neural networks: A comprehensive survey. *IEEE/CAA Journal of Automatica Sinica*, 9(8):1406–1426, 2022.

[123] Hui Zhang, Mile Gu, XD Jiang, Jayne Thompson, Hong Cai, Stefano Paesani, Raffaele Santagati, Anthony Laing, Y Zhang, Man-Hong Yung, et al. An optical neural chip for implementing complex-valued neural network. *Nature communications*, 12(1):457, 2021.

[124] S. Deligiannidis, A. Bogris, C. Mesaritakis, and Y. Kopsinis. Compensation of fiber nonlinearities in digital coherent systems leveraging long short-term memory neural networks. *Journal of Lightwave Technology*, 38(21):5991–5999, 2020.

[125] Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, and Yoshua Bengio. Empirical evaluation of gated recurrent neural networks on sequence modeling. In *NIPS 2014 Workshop on Deep Learning*, Montreal, Canada, December 2014. Replaces arXiv:1412.3555.

[126] Alex Graves and Jürgen Schmidhuber. Framewise phoneme classification with bidirectional lstm and other neural network architectures. *Neural networks*, 18(5-6):602–610, 2005.

[127] Pedro J. Freire, Antonio Napoli, Bernhard Spinnler, Nelson Costa, Sergei K. Turitsyn, and Jaroslaw E. Prilepsky. Towards energy-efficient neural network equalizers: A bayesian approach. In *European Conference on Optical Communication*, pages 1–4. IEEE, 2022.

[128] Rahul Dey and Fathi M Salem. Gate-variants of gated recurrent unit (gru) neural networks. In *2017 IEEE 60th international midwest symposium on circuits and systems (MWSCAS)*, pages 1597–1600. IEEE, 2017.

[129] Pedro J. Freire, Vladislav Osadchuk, Bernhard Spinnler, Antonio Napoli, Nelson Costa, Sergei K. Turitsyn, and Jaroslaw E. Prilepsky. On the computational complexity metrics for neural network equalizers in optical communications. In *Optical Fiber Communication Conference*, pages M3K–5. Optica Publishing Group, 2024.

[130] Jonathan Frankle and Michael Carbin. The lottery ticket hypothesis: Finding sparse, trainable neural networks. *International Conference on Learning Representations*, 2019.

[131] Jonas Mockus, Vytautas Tiesis, and Antanas Zilinskas. The application of bayesian methods for seeking the extremum. *Towards Global Optimization*, 2:117–129, 1978.

[132] Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. MIT Press, Cambridge, MA, 2006.

[133] Donald R. Jones, Matthias Schonlau, and William J. Welch. Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, 13(4):455–492, 1998.

[134] Niranjan Srinivas, Andreas Krause, Sham M. Kakade, and Matthias Seeger. Gaussian process optimization in the bandit setting: No regret and experimental design. *Proceedings of the 27th International Conference on Machine Learning*, pages 1015–1022, 2010.

[135] Tailin Liang, John Glossner, Lei Wang, Shaobo Shi, and Xiaotong Zhang. Pruning and quantization for deep neural network acceleration: A survey. *Neurocomputing*, 461:370–403, 2021.

[136] Christoforos K. Armeniakos, Georgios Zervakis, Dimitrios Soudris, and Jorg Henkel. Hardware acceleration of fully quantized bert for efficient natural language processing. In *Design, Automation & Test in Europe Conference & Exhibition*, pages 1613–1618. IEEE, 2022.

[137] Tuomo Sipola, Janne Alatalo, Tero Kokkonen, and Mika Rantonen. Artificial intelligence in the iot era: A review of edge ai hardware and software. In *2022 31st Conference of Open Innovations Association (FRUCT)*, pages 320–331. IEEE, 2022.

[138] Ramyad Hadidi, Jiashen Cao, Yilun Xie, Bahar Asgari, Tushar Krishna, and Hyesoon Kim. Characterizing the deployment of deep neural networks on commercial edge devices. In *2019 IEEE International Symposium on Workload Characterization (IISWC)*, pages 35–48. IEEE, 2019.

[139] Raphael Tang, Weijie Wang, Zhucheng Tu, and Jimmy Lin. An experimental analysis of the power consumption of convolutional neural networks for keyword spotting. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 5479–5483. IEEE, 2018.

[140] Sebastián Valladares, Mayerly Toscano, Rodrigo Tufiño, Paulina Morillo, and Diego Vallejo-Huanga. Performance evaluation of the nvidia jetson nano through a real-time machine learning application. In *International Conference on Intelligent Human Systems Integration*, pages 343–349. Springer, 2021.

[141] Zeyuan Allen-Zhu, Yuanzhi Li, and Zhao Song. A convergence theory for deep learning via over-parameterization. In *International conference on machine learning*, pages 242–252. PMLR, 2019.

[142] Roberto Verdecchia, June Sallou, and Luís Cruz. A systematic review of green ai. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 13(4):e1507, 2023.

[143] Chuan-Jia Jhang, Cheng-Xin Xue, Je-Min Hung, Fu-Chun Chang, and Meng-Fan Chang. Challenges and trends of sram-based computing-in-memory for ai edge devices. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 68(5):1773–1786, 2021.

[144] Jovan Stojkovic, Esha Choukse, Chaojie Zhang, Inigo Goiri, and Josep Torrellas. Towards greener llms: Bringing energy-efficiency to the forefront of llm inference. *arXiv preprint arXiv:2403.20306*, 2024.

[145] Chaim Baskin, Natan Liss, Eli Schwartz, Evgenii Zheltonozhskii, Raja Giryes, Alex M. Bronstein, and Avi Mendelson. Uniq: Uniform noise injection for non-uniform quantization of neural networks. *ACM Transactions on Computer Systems*, 37(1-4), Jun 2021.

[146] Nhan Tran, Benjamin Hawks, Javier M Duarte, Nicholas J Fraser, Alessandro Pappalardo, and Yaman Umuroglu. Ps and qs: Quantization-aware pruning for efficient low latency neural network inference. *Frontiers in Artificial Intelligence*, 4:94, 2021.

[147] Aneesh Sobhanan, Aravind Anthur, Sean O'Duill, Mark Pelusi, Shu Namiki, Liam Barry, Deepa Venkitesh, and Govind P Agrawal. Semiconductor optical amplifiers: recent advances and applications. *Adv. Opt. Photon.*, 14(3):571–651, Sep 2022.

[148] G.P. Agrawal and N.A. Olsson. Self-phase modulation and spectral broadening of optical pulses in semiconductor laser amplifiers. *IEEE Journal of Quantum Electronics*, 25(11):2297–2306, 1989.

[149] Antonio Mecozzi and Jesper Mørk. Saturation induced by picosecond pulses in semiconductor optical amplifiers. *J. Opt. Soc. Am. B*, 14(4):761–770, Apr 1997.

[150] Xiaoxu Li and Guifang Li. Joint fiber and soa impairment compensation using digital backward propagation. *IEEE Photonics Journal*, 2(5):753–758, 2010.

[151] Darko Zibar, Molly Piels, Rasmus Jones, and Christian G. Schäeffer. Machine learning techniques in optical communication. *Journal of Lightwave Technology*, 34(6):1442–1452, 2016.

[152] Faisal Nadeem Khan, Chao Lu, and Alan Pak Tao Lau. Machine learning methods for optical communication systems. In *Advanced Photonics2017*, page SpW2F.3, 2017.

[153] Christian Häger and Henry D Pfister. Physics-based deep learning for fiber-optic communication systems. *IEEE Journal on Selected Areas in Communications*, 39(1):280–294, 2020.

[154] D. Cassioli, S. Scotti, and A. Mecozzi. A time-domain computer simulator of the nonlinear response of semiconductor optical amplifiers. *IEEE Journal of Quantum Electronics*, 36(9):1072–1080, 2000.

[155] Irshaad Fatadin, David Ives, and Seb J Savory. Blind equalization and carrier phase recovery in a 16-qam optical coherent system. *Journal of lightwave technology*, 27(15):3042–3049, 2009.

[156] Michel Jeruchim. Techniques for estimating the bit error rate in the simulation of digital communication systems. *IEEE Journal on selected areas in communications*, 2(1):153–170, 1984.

[157] James O' Neill. An overview of neural network compression. *arXiv preprint arXiv:2006.03669*, 2020.

[158] Alex Renda, Jonathan Frankle, and Michael Carbin. Comparing rewinding and fine-tuning in neural network pruning. *arXiv preprint arXiv:2003.02389*, 2020.

[159] Song Han. *Efficient methods and hardware for deep learning*. PhD thesis, Stanford University, 2017.

[160] Amirhossein Ghazisaeidi and Leslie Ann Rusch. On the efficiency of digital back-propagation for mitigating soa-induced nonlinear impairments. *Journal of Lightwave Technology*, 29(21):3331–3339, 2011.

[161] Karina Nurlybayeva, Diego Argüello Ron, Morteza Kamalian-Kopae, Elena Turitsyna, and Sergei Turitsyn. Implementation of noise-resistant crowd equalisation in optical communication systems with machine learning dsp. In *2022 Asia Communications and Photonics Conference (ACP)*, pages 753–756. IEEE, 2022.

[162] Chuteng Zhou, Prad Kadambi, Matthew Mattina, and Paul N. Whatmough. Noisy machines: Understanding noisy neural networks and enhancing robustness to analog hardware errors using distillation. *arXiv preprint arXiv:2001.04974*, 2020. [Online]. Replaces placeholder citation.

[163] John G. Proakis and Masoud Salehi. *Digital Communications*. McGraw-Hill, New York, NY, USA, 5th edition, 2007.

[164] Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Springer, New York, NY, USA, 2006.

[165] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

[166] Aleksander Madry, Aleksandar Makelov, Ludwig Schmidt, Dimitris Tsipras, and Adrian Vladu. Towards deep learning models resistant to adversarial attacks. In *International Conference on Learning Representations (ICLR) Workshop Track*, 2018. preprint / widely-cited ICLR work.

[167] Jeremy M. Cohen, Elan Rosenfeld, and J. Zico Kolter. Certified adversarial robustness via randomized smoothing. In Inderjit S. Dhillon, David S. Blei, and Madeleine van der Schaar, editors, *Proceedings of the 36th International Conference on Machine Learning (ICML)*, volume 97, pages 1310–1320, 2019.

[168] Dan Hendrycks and Thomas Dietterich. Benchmarking neural network robustness to common corruptions and perturbations. *arXiv preprint arXiv:1903.12261*, 2019.

[169] Nadezhda Semenova, Laurent Larger, and Daniel Brunner. Understanding and mitigating noise in trained deep neural networks. *Neural Networks*, 146:151–160, 2022.

[170] Nadezhda Semenova and Daniel Brunner. Impact of white noise in artificial neural networks trained for classification: Performance and noise mitigation strategies. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 34(5):051101, 2024. Replaces arXiv:2405.07670.

[171] Ziyang Liu, Min Yan, Huaizhou Schlenker, Haoqin Zhang, Ziming Li, and Joseph M Kahn. Analog optical neural networks: prospects and challenges. *Nature Communications*, 13(1):2596, 2022.

[172] Bernt Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. Springer, Berlin, Germany, 6 edition, 2003.

[173] Ioannis Karatzas and Steven E. Shreve. *Brownian Motion and Stochastic Calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer, New York, NY, USA, 2 edition, 1991.

[174] Bahaa E. A. Saleh and Malvin C. Teich. *Fundamentals of Photonics*. John Wiley & Sons, Hoboken, NJ, USA, 2 edition, 2007.

[175] Govind P. Agrawal. *Fiber-Optic Communication Systems*. John Wiley & Sons, Hoboken, NJ, USA, 4 edition, 2010. See Chs. 3–4 for laser RIN, receiver noise, and amplifier noise figure.

[176] H. Nyquist. Thermal agitation of electric charge in conductors. *Phys. Rev.*, 32(1):110–113, 1928.

[177] J. B. Johnson. Thermal agitation of electricity in conductors. *Physical Review*, 32(1):97–109, 1928.

[178] W. K. Marshall, B. Crosignani, and Amnon Yariv. Laser phase noise to intensity noise conversion by lowest-order group-velocity dispersion in optical fiber: exact theory. *Optics Letters*, 25(3):165–167, 2000.

[179] Alejandro Baldominos, Yago Saez, and Pedro Isasi. A survey of handwritten character recognition with mnist and emnist. *Applied Sciences*, 9(15):3169, 2019.

[180] R. Paschotta. Shot noise. RP Photonics Encyclopedia, 2024.

[181] R. Paschotta. Relative intensity noise. RP Photonics Encyclopedia, 2024.

[182] J. Buus and C. A. Burrus. How to measure relative intensity noise in lasers. *Laser Focus World*, 2012. Typical RIN definitions and units (dB/Hz).

[183] James Bergstra and Yoshua Bengio. Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, 13(2):281–305, 2012.

[184] Razvan Pascanu, Tomas Mikolov, and Yoshua Bengio. On the difficulty of training recurrent neural networks. *International Conference on Machine Learning*, pages 1310–1318, 2013.

[185] Martin Arjovsky, Amar Shah, and Yoshua Bengio. Unitary evolution recurrent neural networks. In *International Conference on Machine Learning*, pages 1120–1128. PMLR, 2016.

[186] Ruiyang Chen, Yingjie Li, Minhan Lou, Cunxi Yu, and Weilu Gao. Complex-valued reconfigurable diffractive optical neural networks using cost-effective spatial light modulators. In *CLEO: Applications and Technology*, pages JTh3B–56. Optica Publishing Group, 2022.

[187] Ruiyang Chen, Yingjie Li, Minhan Lou, Jichao Fan, Yingheng Tang, Berardi Sensale-Rodriguez, Cunxi Yu, and Weilu Gao. Physics-aware machine learning and adversarial attack in complex-valued reconfigurable diffractive all-optical neural network. *Laser & Photonics Reviews*, page 2200348, 2022.

[188] G Mourgias-Alexandris, M Moralis-Pegios, A Tsakyridis, S Simos, G Dabos, A Totovic, N Passalis, M Kirtas, T Rutirawut, FY Gardes, et al. Noise-resilient and high-speed deep learning with coherent silicon photonics. *Nature Communications*, 13(1):1–7, 2022.

[189] Vadim S Anishchenko, Arkady B Neiman, F Moss, and L Shimansky-Geier. Stochastic resonance: noise-enhanced order. *Physics-Uspekhi*, 42(1):7, 1999.

[190] Stéphan Fauve and F Heslot. Stochastic resonance in a bistable system. *Physics Letters A*, 97(1-2):5–7, 1983.

[191] Giacomo Indiveri, T. Linares-Barranco, T. J. Hamilton, A. van Schaik, R. Etienne-Cummings, T. Delbruck, S.-C. Liu, P. Dudek, P. Häfliger, S. Renaud, J. Schemmel, G. Cauwenberghs, J. Arthur, K. Hynna, F. Folowosele, S. Saïghi, R. Serrano-Gotarredona, F. Wijekoon, Y. Wang, and K. Boahen. Neuromorphic silicon neuron circuits. *Frontiers in Neuroscience*, 5:73, 2011.

[192] S. Sheik, E. Chicca, and G. Indiveri. Exploiting device mismatch in neuromorphic vlsi systems to implement axonal delays. In *Proc. International Joint Conference on Neural Networks (IJCNN)*, pages 1940–1945, Brisbane, Australia, 2012.

[193] Luca Gammaitoni, Peter Hänggi, Peter Jung, and Fabio Marchesoni. Stochastic resonance. *Reviews of Modern Physics*, 70(1):223–287, 1998.

[194] Aren Shinozaki, Takaya Miyano, and Yoshihiko Horio. Chaotic time series prediction by noisy echo state network. *Nonlinear Theory and Its Applications, IEICE*, 11(4):466–479, 2020.

[195] Christopher M. Bishop. Training with noise is equivalent to tikhonov regularization. *Neural Computation*, 7(1):108–116, 1995.

[196] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15(56):1929–1958, 2014.

[197] Paul A. Merolla, John V. Arthur, Rodrigo Alvarez-Icaza, Andrew S. Cassidy, Jun Sawada, Filipp Akopyan, Bryan L. Jackson, Nabil Imam, Chen Guo, Yutaka Nakamura, Bernard Brezzo, Ivan Vo, Steven K. Esser, Rathinakumar Appuswamy, Brian Taba, Arnon Amir, Myron D. Flickner, William P. Risk, Rajit Manohar, and Dharmendra S. Modha. A million spiking-neuron integrated circuit with a scalable communication network and interface. *Science*, 345(6197):668–673, 2014.

[198] Mike Davies, Narayan Srinivasa, Tsung-Han Lin, Gautham Chinya, Yongqiang Cao, Sri Harsha Choday, Georgios Dimou, Prasad Joshi, Nabil Imam, Shweta Jain, Yuyun Liao, Chit-Kwan Lin, Andrew Lines, Ruokun Liu, Deepak Mathaikutty, Steven McCoy, Arnab Paul, Jonathan Tse, Guruguhanathan Venkataramanan, Yi-Hsin Weng, Andreas Wild, Yoonseok Yang, and Hong Wang. Loihi: A neuromorphic manycore processor with on-chip learning. *IEEE Micro*, 38(1):82–99, 2018.

[199] Wim Bogaerts, Daniel Pérez, José Capmany, David A. B. Miller, Joyce Poon, Dirk Englund, Francesco Morichetti, and Andrea Melloni. Programmable photonic circuits. *Nature*, 586(7828):207–216, 2020. Open access author version: `https://backoffice.biblio.ugent.be/download/8681996/8681998`.

[200] Yiwei Xie, Jiachen Wu, Shihan Hong, Cong Wang, Shujun Liu, Huan Li, Xinyan Ju, Xiyuan Ke, Dajian Liu, and Daoxin Dai. Towards large-scale programmable silicon photonic chip for signal processing. *Nanophotonics*, 13(12):2051–2073, 2024.

[201] David A. B. Miller. Self-configuring universal linear optical component. *Photonics Research*, 1(1):1–15, 2013.

[202] Edward Vul and Harold Pashler. Measuring the crowd within: Probabilistic representations within individuals. *Psychological Science*, 19(7):645–647, 2008. PMID: 18727777.

[203] Francis Galton. Vox populi (the wisdom of crowds). *Nature*, 75(7):450–451, 1907.

[204] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2017. arXiv:1706.03762.

[205] Bo Peng, Eric Alcaide, Quentin Anthony, Alon Albalak, Samuel Arcadinho, Stella Biderman, Huanqi Cao, Xin Cheng, Michael Chung, Matteo Grella, et al. RWKV: Reinventing RNNs for the transformer era. In *Findings of the Association for Computational Linguistics: EMNLP*

2023, pages 14048–14077, Singapore, 2023. Association for Computational Linguistics. Replaces arXiv:2305.13048.

[206] Emre O. Neftci, Hesham Mostafa, and Friedemann Zenke. Surrogate gradient learning in spiking neural networks: Bringing the power of gradient-based optimization to spiking neural networks. *IEEE Signal Processing Magazine*, 36(6):51–63, 2019.

[207] Kaushik Roy, Akhilesh Jaiswal, and Priyadarshini Panda. Towards spike-based machine intelligence with neuromorphic computing. *Nature*, 575(7784):607–617, 2019.

[208] Zihang Dai, Zhilin Yang, Yiming Yang, Jaime Carbonell, Quoc V Le, and Ruslan Salakhutdinov. Transformer-XL: Attentive language models beyond a fixed-length context. In *Proc. ICLR*, 2019.

[209] Elias Arnold, Georg Böcherer, Eric Müller, Philipp Spilger, Johannes Schemmel, Stefano Calabrò, and Maxim Kuschnerov. Spiking neural network equalization on neuromorphic hardware for IM/DD optical communication. In *European Conference on Optical Communication (ECOC)*, pages 1–4, 2022. Replaces arXiv:2206.00401.

[210] Georg Böcherer, Florian Strasser, Elias Arnold, Youxi Lin, Johannes Schemmel, Stefano Calabrò, and Maxim Kuschnerov. Spiking neural network linear equalization: Experimental demonstration of 2km 100gb/s im/dd pam4 optical transmission. In *Optical Fiber Communication Conference (OFC)*, page W4E.1. Optica Publishing Group, 2023.

[211] Elias Arnold, Georg Böcherer, Florian Strasser, Youxi Lin, Johannes Schemmel, Stefano Calabrò, and Maxim Kuschnerov. Spiking neural network nonlinear demapping on neuromorphic hardware for im/dd optical communication. *Journal of Lightwave Technology*, 41(11):3424–3431, 2023.

[212] Shuangxu Li, Georg Böcherer, Stefano Calabrò, and Maximilian Schädler. Spiking neural network equalizer with fast and low power decoding for im/dd optical communication. *IEEE Photonics Technology Letters*, 36(17):1061–1064, 2024.

[213] Alexander von Bank, Eike-Manuel Edelmann, and Laurent Schmalen. Energy-efficient spiking neural network equalization for im/dd systems with optimized neural encoding. In *Optical Fiber Communication Conference (OFC)*, 2024. Preprint: arXiv:2312.12909.

[214] Naveenta Gautam, Sai Vikranth Pendem, Brejesh Lall, and Amol Choudhary. Transformer-based nonlinear equalization for dp-16qam coherent optical communication systems. *IEEE Communications Letters*, 28(3):577–581, 2024.

[215] Naveenta Gautam, Prankush Agarwal, C. Alex Kaylor, Brejesh Lall, Amol Choudhary, and Stephen E. Ralph. Experimental demonstration of machine learning transformer based equalization of channel nonlinearities. In *Proc. European Conference on Optical Communication (ECOC)*, pages 1–4, Frankfurt, Germany, 2024. VDE Verlag, ITG-Fb. 317.

[216] Behnam Behinaein Hamgini, Hossein Najafi, Ali Bakhshali, and Zhuhong Zhang. Application of transformers for nonlinear channel compensation in optical systems. *arXiv preprint*, 2023.

[217] Niuyong Zhang, Hang Yang, Zekun Niu, Lizhuo Zheng, Cao Chen, Shilin Xiao, and Lilin Yi. Transformer-based long distance fiber channel modeling for optical ofdm systems. *Journal of Lightwave Technology*, 40(24):7779–7789, 2022.

[218] Makoto Matsumoto and Takuji Nishimura. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 8(1):3–30, 1998.

[219] Rui-Jie Zhu, Qihang Zhao, Guoqi Li, and Jason K. Eshraghian. SpikeGPT: Generative pre-trained language model with spiking neural networks. *Transactions on Machine Learning Research*, 2024. Replaces arXiv:2302.13939.

[220] Naveenta Gautam, Sai Vikranth Pendem, Brejesh Lall, and Amol Choudhary. Transformer-based nonlinear equalization for dp-16qam coherent optical communication systems. *IEEE Communications Letters*, 2023.

[221] Shuangfei Zhai, Walter Talbott, Nitish Srivastava, Chen Huang, Hanlin Goh, Ruixiang Zhang, and Josh Susskind. An attention free transformer. *arXiv preprint arXiv:2105.14103*, 2021.

[222] Tung-Ju Yang, Yu-Hsin Chen, and Vivienne Sze. Design space exploration of neural network activation function circuits. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 38(10):1974–1978, 2019.

[223] Mark Horowitz. 1.1 computing's energy problem (and what we can do about it). *IEEE International Solid-State Circuits Conference Digest of Technical Papers*, pages 10–14, 2014.

[224] Man Yao, Jiakui Hu, Zhaokun Zhou, Li Yuan, Yonghong Tian, Bo Xu, and Guoqi Li. Spike-driven transformer. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 36, pages 17169–17182, 2023. Replaces arXiv:2307.01694.

[225] Bhavin J Shastri, Alexander N Tait, Thomas Ferreira de Lima, Wolfram HP Pernice, Harish Bhaskaran, C David Wright, and Paul R Prucnal. Photonics for artificial intelligence and neuromorphic computing. *Nature Photonics*, 15(2):102–114, 2021.

[226] Renjie Li, Yuanhao Gong, Hai Huang, Yuze Zhou, Sixuan Mao, Zhijian Wei, and Zhaoyu Zhang. Photonics for neuromorphic computing: Fundamentals, devices, and opportunities. *Advanced Materials*, 37(2):2312825, 2025.

[227] Yichen Shen, Nicholas C. Harris, Scott Skirlo, Mihika Prabhu, Thomas Baehr-Jones, Michael Hochberg, Xin Sun, Shijie Zhao, Hyeji Larochelle, Dirk Englund, and Marin Soljačić. Deep learning with coherent nanophotonic circuits. *Nature Photonics*, 11(7):441–446, 2017. Accepted manuscript: `https://dspace.mit.edu/handle/1721.1/113673`.

[228] Alexander N Tait, Thomas Ferreira de Lima, Ellen Zhou, Allie X Wu, Mitchell A Nahmias, Bhavin J Shastri, and Paul R Prucnal. Neuromorphic photonic networks using silicon photonic weight banks. *Scientific Reports*, 7(1):7430, 2017.

[229] Kristof Vandoorne, Pauline Mechet, Thomas Van Vaerenbergh, Martin Fiers, Geert Morthier, David Verstraeten, Benjamin Schrauwen, Joni Dambre, and Peter Bienstman. Experimental demonstration of reservoir computing on a silicon photonics chip. *Nature Communications*, 5(1):3541, 2014.

[230] Johannes Feldmann, Nadine Youngblood, Maxim Karpov, Hui Ge, Andrea Schäfer, Michael Sterzl, Niels Giesecke, Michael M. Hübner, Harish Bhaskaran, Wolfram Pernice, and Tobias J. Kippenberg. All-optical spiking neurosynaptic networks with self-learning capabilities. *Nature*, 569(7755):208–214, 2019.

[231] Mitchell A Nahmias, Bhavin J Shastri, Alexander N Tait, and Paul R Prucnal. A leaky integrate-and-fire laser neuron for ultrafast cognitive computing. *IEEE Journal of Selected Topics in Quantum Electronics*, 19(5):1–12, 2013.

[232] J. Robertson, F. Farias, J. S. Aitchison, J. Delamare, and A. Hurtado. Ultrafast optical integration and pattern classification with a spiking vcsel neuron. *Scientific Reports*, 10:6098, 2020. Demonstrates ultrafast optical computation with spiking lasers.

[233] Tobias Heuser, Moritz Pflüger, Ingo Fischer, James A. Lott, Daniel Brunner, and Stephan Reitzenstein. Developing a photonic hardware platform for brain-inspired computing based on 5×5 vcsel arrays. *J. Phys. Photonics*, 2:044002, 2020.

[234] Jeffrey M Shainline, Sonia M Buckley, Richard P Mirin, and Sae Woo Nam. Superconducting optoelectronic circuits for neuromorphic computing. *Physical Review Applied*, 7(3):034013, 2017.

[235] J. Robertson, C. Mesaritakis, A. Kapsalis, D. Syvridis, and A. Hurtado. Toward neuromorphic photonic networks of ultrafast spiking laser neurons. *IEEE Journal of Selected Topics in Quantum Electronics*, 26(1):1–12, 2019. Sub-ns spiking behavior and network implications.

[236] J. Robertson, F. Farias, and A. Hurtado. Controlled propagation of spiking dynamics in a multi-neuron excitatory–inhibitory laser network. *Optics Express*, 28(24):35822–35834, 2020. Reports <100 ps optical spikes and spike propagation control.

[237] T. Ferreira de Lima, A. N. Tait, M. A. Nahmias, H.-T. Peng, G. Shrinivasan, Y. Shen, and P. R. Prucnal. Machine learning with neuromorphic photonics. *Journal of Lightwave Technology*, 37(5):1515–1534, 2019. Survey with quantitative speed/energy figures for photonic neuromorphic hardware.

[238] Mike Davies, Andreas Wild, Garrick Orchard, Yulia Sandamirskaya, Gabriel AF Sugiarto, Richard Wanka, Mihai Ionica, Schuyler Paul, Nigel Mathews, Phil Knag, et al. Efficient neuromorphic signal processing with loihi 2. In *2023 IEEE Workshop on Signal Processing Systems (SiPS)*, pages 254–259. IEEE, 2022.

[239] K. Padmaraju and K. Bergman. Resolving the thermal challenges for silicon microring resonator devices. *Nanophotonics*, 3(4–5):269–281, 2014. Review of thermal sensitivity, drift, and stabilization approaches in microring devices.

[240] A. N. Tait, T. F. De Lima, M. A. Nahmias, B. J. Shastri, and P. R. Prucnal. Broadcast-and-weight: An integrated network for scalable photonic spike processing. *Journal of Lightwave Technology*, 32(21):4029–4041, 2014. Introduces WDM broadcast-and-weight for scalable photonic SNNs.

[241] Carlos Ríos, Matthias Stegmaier, Peiman Hosseini, Di Wang, Torsten Scherer, C David Wright, Harish Bhaskaran, and Wolfram HP Pernice. Integrated all-photonic non-volatile multi-level memory. *Nature Photonics*, 9(11):725–732, 2015.

[242] Zhengyang Cheng, Carlos Ríos, Wolfram HP Pernice, C David Wright, and Harish Bhaskaran. On-chip photonic synapse. *Science Advances*, 3(9):e1700160, 2017.

[243] Daniel Brunner, Miguel C Soriano, Claudio R Mirasso, and Ingo Fischer. Parallel photonic information processing at gigabyte per second data rates using transient states. *Nature Communications*, 4(1):1364, 2013.

[244] Laurent Larger, Antonio Baylón-Fuentes, Romain Martinenghi, Vladimir S Udaltsov, Yanne K Chembo, and Maxime Jacquot. High-speed photonic reservoir computing using a time-delay-based architecture. *Physical Review X*, 7(1):011015, 2017.

[245] Wulfram Gerstner and Werner M. Kistler. *Spiking Neuron Models: Single Neurons, Populations, Plasticity*. Cambridge University Press, 2002.

[246] Pontus Johannisson and Magnus Karlsson. Perturbation analysis of nonlinear propagation in a strongly dispersive optical communication system. *Journal of Lightwave Technology*, 31(8):1273–1282, 2013.

[247] Andrea Carena, Vittorio Curri, Pierluigi Poggiolini, Gabriella Bosco, et al. Egn model of nonlinear fiber propagation. *Optics Express*, 22(13):16335–16362, 2014.

[248] Nikolaos Farmakidis, Bowei Dong, and Harish Bhaskaran. Integrated photonic neuromorphic computing: opportunities and challenges. *Nature Reviews Electrical Engineering*, 1(6):358–373, 2024.

[249] Anas Skalli, Mirko Goldmann, Nasibeh Haghighi, Stephan Reitzenstein, James A Lott, and Daniel Brunner. Annealing-inspired training of an optical neural network with ternary weights. *Communications Physics*, 8(1):68, 2025.