TRANSMISSION PERFORMANCE OPTIMIZATION IN FIBER-WIRELESS ACCESS NETWORKS USING MACHINE LEARNING TECHNIQUES

A Thesis Proposal
Presented to
The Academic Faculty

by

Qi Zhou

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Electrical and Computer Engineering

Georgia Institute of Technology
December 2021

TRANSMISSION PERFORMANCE OPTIMIZATION IN FIBER-WIRELESS ACCESS NETWORKS USING MACHINE LEARNING TECHNIQUES

Approved by:

Dr. Matthieu R. Bloch, Advisor
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Xiaoli Ma
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. David V. Anderson
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. John Barry
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Shiwen Mao
Department of Electrical and Computer Engineering
Auburn University

Date Approved: December 2, 2021

To my family and friends.

ACKNOWLEDGEMENTS

I would like to take this great opportunity to express my sincere gratitude to all the people who made this dissertation possible.

First, I'm sincerely grateful to my advisor, Prof. Gee-Kung Chang, who provided me an excellent opportunity to start this PhD journey after a two-year career in the industry. His exceptional intellectual scope, research acumen has inspired and shaped me as a researcher who can think out of the box and perform innovative research. It is my greatest honor and pleasure to have worked with him as a PhD student.

Further, I thank my dissertation committee members. Prof. John R. Barry has given me the best digital communication lectures which form the essential theoretical foundation for my research. I would like to thank Prof. Matthieu R. Bloch, Prof. Xiaoli Ma, Prof. David V. Anderson, Prof. John Barry, and Prof. Shiwen Mao at Auburn University for their valuable time devoted in service to my dissertation committee, and insightful feedback that greatly improved the quality of this work.

I am certainly appreciative to my colleagues in the research group, for being the best teammates during my wonderful journey at Georgia Tech. Thanks to Dr. Daniel Guidotti, Dr. Feng Lu, Dr. Mu Xu, Dr. You-Wei Chen, Dr. Junwen Zhang, Dr. Chin-Wei Hsu, Dr. Zhensheng Jia, Dr. Jing Wang, Dr. Lin Cheng, Dr. Shuyi Shen, and soon to graduate fellow students, Yahya M. Alfadhli, Shuang Yao, Rui Zhang, Shang-Jen Su, Muhammad S. Omar, Xushuai Qu, as well as many other graduate students, and staff members.

Most importantly, I would like to thank my family for all their love and unconditional support. For my parents who gave me the best opportunity to achieve my educational aspirations and the greatest encouragement to start my PhD journey. Saving the best for the last, thanks to my beloved wife, Xi Chen. Your support, patience, and unswerving love are the pillars for the past 8 years of my life, and the foremost incentives to complete my PhD study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

3GPP	3rd Generation Partnership Project
5G NR	the fifth-generation New Radio
A-RoF/D-RoF	Analog/Digital Radio over fiber
BER	Bit-error rate
CDR	Clock data recovery
CU	Central unit
DBA	Dynamic bandwidth allocation
DMT	Discrete multi-tone
DNN	Deep neural network
DOCSIS	Data over cable service interface specifications
DPD	Digital pre-distortion
DSP	Digital signal processing
DU	Distributed unit
ECL	External-cavity laser
EDFA	Erbium doped fiber amplifier
eMBB	Enhanced mobile broadband
FDD	Frequency-division duplex
FEC	Forward error correction
IBA	Intelligent bandwidth allocation
IBFD	In-band full duplex
IMDD	Intensity modulation direct detection

LO	Local oscillator
mmWave	Millimeter wave
MZM	Mach-Zehnder modulator
NR	New radio
NOMA	Non-orthogonal multiple access
OCS	Optical carrier suppression
OFDM	Orthogonal frequency division multiplexing
OLT	Optical line terminal
ONU	Optical network unit
OOK	On-off keying
PAM	Pulse amplitude modulation
PAPR	Peak to average power ratio
PD	Photodetector
PIC	Parallel interference cancellation
PON	Passive optical network
PS	Probabilistic shaping
RAN	Radio access network
RL	Reinforcement learning
RoF	Radio over fiber
ROP	Received optical power
RRU	Remote radio unit
SIC	Successive/Self interference cancellation
SOI	Signal of interest

SSMF	Standard Single-mode fiber
SNR	Signal-to-noise power ratio
TDD	Time-division duplex
TIA	Transimpedance amplifier

SUMMARY

We are now in the era of alluvial global data flows from massive data contents are created by subscribers on various social media platforms such as YouTube, TikTok, Facebook and others, at an ever-accelerating pace. Additionally, the advent and early adoption of immersive experience, provided by such platforms as virtual reality (VR) and augmented reality (AR) enhances demand for ultra-low latency in the movement of massive volume of data. In response to this demand, fiber optics and wireless technologies have penetrated every capillary of modern, metropolitan community living to provide ubiquitous and seemingly instant communications. As an aggregator of fiber and wireless resources, the fiber-wireless access network is in urgent need of dramatic innovations so as to satisfy the requirements of high-capacity, ultra-low latency and ultra-reliability of communications in the new era. The radio access network (RAN) serves the purpose to connect the core network and the user equipment (UE). The base station (BS) connects the UE via wireless spectra, where millimeter-wave (mmWave) is adopted in 5G new radio (NR) to expand the wireless bandwidth. The link between the BS and the mobile core network is established by the optical fiber due to its low propagation loss and wide optical bandwidth. In a modern RAN, the multi-layer radio processing functions of a conventional BS are distributed to the central unit (CU), distribute unit (DU) and remote radio unit (RRU) through function splits.

The radio signals can be transmitted via the fibers in a RAN system based on two prevailing radio over fiber (RoF) schemes, namely, analog RoF (A-RoF) and digital RoF (D-RoF). In A-RoF, the analog radio signal is converted to the optical domain and

transported directly to or from the RRU antenna via optical fiber. On the other hand, D-RoF quantizes the radio signal and use robust optical modulation format, for example OOK to transmit the quantization bits over the fiber. At the RRU, the received bits require extra DSP to reconstruct the radio signal for wireless propagation. The A-RoF excels in terms of optical bandwidth efficiency over short distances while the quantized D-RoF, while less spectrally and power efficient, it generally obtains a much higher fidelity than its A-RoF counterpart which suffers from severe channel impairments, including nonlinear degradation, fiber dispersion, etc.

Considering the advantages and limitation of both RoF schemes, it is clear that both require further improvements if RAN implementation is to reach higher throughput and spectral efficiency.

The dissertation focuses on enhancing the transmission performance in the fiber-wireless access network through mitigating the vital system-level limitations of both D-RoF and A-RoF, with machine learning techniques being systematically implemented. Since D-RoF needs to substantially increase its optical transport capacity to deliver massive volume of digital radio signal quantization bits, the first task is to improve its utilization efficiency of the optical bandwidth. Compared with the conventional OOK modulation formats in the D-RoF, the advanced PAM8 modulation can triple the spectral efficiency. However, the intensity-modulation-direct-detection (IMDD) based optical link has a small dynamic range which can be easily overwhelmed by advanced modulated signals with large drive voltages. In this case, however, the signal quality is degraded by nonlinear impairments.

In this work, we design an efficient deep neural network (DNN) decoder to mitigate the resulting nonlinear degradations. Comparing with the conventional nonlinear Volterra series based methods, the DNN shows better performance and has a lower implementation complexity. To further improve the optical bandwidth efficiency, we also investigate the probabilistic shaping (PS) method which can customize the modulation entropy by tuning the constellation occurrence probability. With the PS-PAM8 modulation implemented, it is demonstrated that an 80-Gbps signal can be transmitted over 20-km standard single mode fiber (SSMF) with a BER at the FEC threshold of about 3.8e-3 at as minimum received optical power as possible. In addition, a neural network based digital predistortion (DPD) method is implemented to mitigate the nonlinear impairments over the whole link. The DPD function can be integrated within the transmitter which typically has more processing resources and power than the receiver in typical access networks. Through optical transmission experiments, it is shown that implementation of the DPD leads to a dramatic BER gain over competing linear pre-equalization schemes. Aside from DSP contributions, an optical lite coherent system is demonstrated to increase optical transmission capacity with significantly higher optical receiver sensitivity and longer transmission distance than conventional IMDD link. A -26-dBm sensitivity is experimentally achieved at 50-Gbps data rate over 100-km SSMF. It is worth noting that the D-RoF can employ the low-cost passive optical network (PON) infrastructure, where we design an intelligent dynamic bandwidth allocation (DBA) scheme bandwidth reinforcement learning (RL) to optimize the optical bandwidth utilization and reduce the latency.

The wireless spectrum is a scarce resource shared by different wireless systems and services such as mobile communication, wireless local-area networks (WLAN), and satellite positioning. As the density of active radio transmitters has been increasing exponentially, the signal interference becomes a prevalent and unavoidable issue in the 5G RAN. The exploitation of signals from different licensed systems and unlicensed transmitters creates a unprecedently complex interference environment which cannot be solved by conventional pre-defined network planning. In response to the challenges, a proactive interference avoidance scheme using reinforcement learning is proposed and experimentally verified in an mmWave-over-fiber platform. Except for the external sources, the interference may also arise internally from a local transmitter when the transmitting signal and receiving signal are co-time and co-frequency (i.e. transmission on the same channel at the same time). The self-interference (SI) from the local transmitter is overwhelming to the received signal of interest (SOI) from a remote radio station and cannot be removed by an RF filter since it is in band. Different from the conventional subtraction-based SI cancellation scheme, we design an efficient Dual-inputs DNN (DI-DNN) based canceller which simultaneously cancels the SI and recovers the SOI. Experimental results also indicate a better SOI recovery performance with capability on nonlinearity mitigation. The Non-orthogonal multiple access (NOMA) has been employed to utilize power-domain resources and increase the number of connected user equipment (UE) within the same time-frequency block. In contrast to the conventional orthogonal frequency-division multiple access (OFDMA), NOMA schemes allow multiplexing the UE signals in the power domain at the cost of larger inter-UE interference. Successive interference cancellation (SIC) is a common solution that

decodes the UE's data sequentially based on power-descending order. However, the SIC inevitably introduces high processing latency when more UE signals are packed in NOMA, while the inherent error propagation can significantly degrade the sensitivity of the subsequent UEs. In this case, parallel interference cancellation (PIC) which can jointly decode all the UEs' data, becomes a desirable solution. We propose and experimentally demonstrate a convolutional neural network (CNN) based PIC scheme to decode the NOMA signals which effectively avoids the error propagation with improved EVM performance.

CHAPTER 1. INTRODUCTION

1.1 Motivation

In the surge of the "information age", we have witnessed a revolution in the way we see, connect, and communicate due to the advent of vital technologies such as 8K videos with high refresh rate, AR/VR headset, and 3D project starline, to name a few. These emerging applications with massive data exchange require more efficient transport networks than ever before, which envisions ultra-broadband, instantaneity, and super-reliability of 5G and beyond radio access networks (RAN). Moreover, it must catch up with the explosive growth of mobile data traffic, which is expected to be 607 EB/month by 2025 and even 5016 EB/month by 2030 as shown in Figure 1.1 [1]. The underlying technologies to support the massive mobile data traffic rely on optical communication and networking due to their wide-bandwidth and low propagation loss. Optical fibers connect core networks to all kinds of wireless cells and serve as the primary medium in the mobile fronthaul (MFH) of the RAN. Besides, new wireless spectrum resources such as millimeter waves (mmWave) are adopted in 5G new radio (NR) to expand the wireless bandwidth. While most current technologies tackle the two mediums (fiber and wireless) in an isolated fashion, the topic of radio-over-fiber (RoF) based integrated fiber-wireless network focuses on optimizing and orchestrating the two media collectively to achieve higher bandwidth capability with simple and cost-effective design. At least that's the goal.

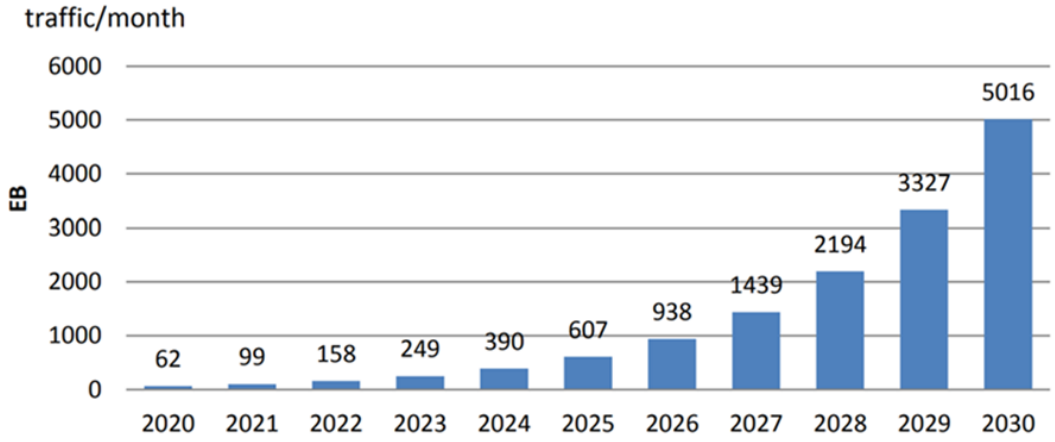


Figure 1.1 Global mobile traffic estimation from 2020-2030 [1].

There are two types of RoF systems, namely, analog RoF (A-RoF) and digital RoF (D-RoF). A-RoF significantly simplifies the architecture of the base station (BS), as it becomes purely analog without inclusion of any digital signal processing (DSP) functions. Besides, transmitting the analog radio signal instead of the digitized radio samples in the optical fiber, greatly saves the optical bandwidth. However, A-RoF is susceptible to nonlinear degradations due to the overwhelmed dynamic range of optical transceivers and power amplifiers [2]. The D-RoF transmission has a higher fidelity as it relies on robust modulation schemes such as OOK. Moreover, as D-RoF exploits similar optical transport technologies to the ones used in passive optical network (PON) and can be supported by a low-cost PON infrastructure with minor revisions [3]. Nevertheless, D-RoF has low bandwidth efficiency as large numbers of quantization bits digitized from radio signal are transmitted in the fiber link. For example, according to the technical specification defined by the Common Public Radio Interface (CPRI), transporting of 4-Gb/s wireless data via D-RoF technology requires 157.3-Gb/s of optical transmission capacity, for a typical configuration [4]. The fundamental challenge of D-RoF lies on improving the bandwidth efficiency. Therefore, advanced modulation formats with higher spectral efficiency and

effective schemes to recover the signals from the associated D-RoF transmission impairments deserve further investigations.

Aside from those fundamental challenges, due to small cells densification and intensive spectrum reuse, the 5G and beyond RAN also faces severe interferences from different sources including inter-cell interference, self-interference, etc. Moreover, the macro or small cells from different wireless operators mainly work independently such that conventional pre-defined network planning and resource allocation schemes will not work effectively to fix the interference in the heterogenous RAN [5]. The industry has long suffered from the expensive, unreliable, and inefficient interference management through inadequate power/frequency adjustments and human-centric interference hunting. Consequently, for the purpose of maintaining uncompromised user experience in the presence of accidental or malicious interference, a proactive self-organized interference mitigation scheme is inevitable.

Machine learning (ML) technology is known as an essential driving force on the ongoing Fourth Industrial Revolution. The ML can play a crucial role in 5G and beyond RAN, thanks to its capability to model systems without requiring closed-form equations. Meanwhile, certain heuristic or brute-force algorithms can be replaced by ML with better performance and lower inference complexity. Moreover, it is envisioned that ML will ultimately enable real-time fault diagnosis and autonomous zero-touch network management [6]. Recently, both academia and industry put together heroic efforts in delivering new ML algorithms and ML training/inference hardware, significantly reducing the implementation cost, and strengthening the effectiveness/robustness of ML model. This wide acceptance of ML has inspired telecom operators to improve the

availability and accessibility of their operational data, which can be fully mined and exploited by ML algorithm developers. This dissertation presents an attempt to tackle the essential challenges in the fiber-wireless networks through experimental implementation of novel ML algorithms and validation on their performance over conventional baseline methods.

The objective of the dissertation is to experimentally investigate and validate machine learning techniques for transmission performance optimization in radio-over-fiber (RoF) based fiber-wireless access networks. Since the D-RoF is the most widely accepted RoF transport technology in modern RANs, we aim to alleviate the capacity crunch of the D-RoF using advanced modulation schemes and neural networks based nonlinear equalization/pre-distortion algorithms. Besides, we improve the efficiency of the PON infrastructure in terms of capacity and latency to facilitate low-cost D-RoF deployments, where a lite coherent system and an intelligent DBA algorithm are experimentally validated. Moreover, the dissertation focuses on mitigating the complex interference in 5G and beyond RANs. The interference mitigation performances are experimentally evaluated based on millimeter-wave-RoF platforms. Specifically, a proactive interference avoidance scheme using reinforcement learning and an effective method for simultaneous self-interference (SI) cancellation and signal-of-interest (SOI) recovery using a dual-inputs DNN (DI-DNN) are experimentally demonstrated.

1.2 Background and Challenges

1.2.1 Fiber-Wireless Network Architectures and mmWave Integration

As an effective integration of optical fiber and wireless communications, the fiber-wireless network shown in Figure 1.2 takes advantage of both technologies, such that it poses benefits of long transmission distance and large communication bandwidth while supporting high mobility and seamless coverage. These valuable features, pioneered by Prof. Chang in 2005, have catapulted the fiber-wireless network to the status of key enabling technology for mobile fronthaul access networks.

The most widely deployed MFH architecture in fiber-wireless network is the D-RoF which digitally delivers the radio waveform through the optical fiber after converting the waveform into digital bits followed by intermediate frequency carrier modulation, optical conversion and fiber transmission. Since optical and wireless modulation are occupying different physical domains, optical fiber transmission performance can be optimized separately to improve the fidelity of the digitized radio signal without affecting the wireless standards. For example, low-cost commercialized facilities and products intended for the PON industry, as shown in Figure 1.2, can also fit the needs of D-RoF with minor revisions. The commonly implemented standard for the D-RoF is CPRI which replaces a copper coaxial cable link between a radio transceiver and a base station such that allowing a remote and more stable connection. A CPRI link transmits the digitized IQ samples of baseband radio signals. Although it is a simple technology that has been widely adopted since 2G, the transmission of digitized IQ samples requires a large optical bandwidth, making it unscalable for 5G services with extended wireless bandwidth, carrier aggregation and massive MIMO. For example, to achieve a wireless data rate of 1 Gb/s with 8x8 MIMO and 3 sectors, a 147.5 Gb/s of optical rate is required based on the CPRI interface [7]. The CPRI rate requirement can

be reduced through data compression by downsampling [8], nonlinear companding [9], etc. The downsampling method acquiescently sets an over-sampling ratio to remove the OFDM oversampling dependency, while the nonlinear companding can utilize OFDM amplitude distribution via compression algorithms [10]. Although the optical bandwidth demand can be reduced by up to 50%, the compressed CPRI still requires a significant capacity.

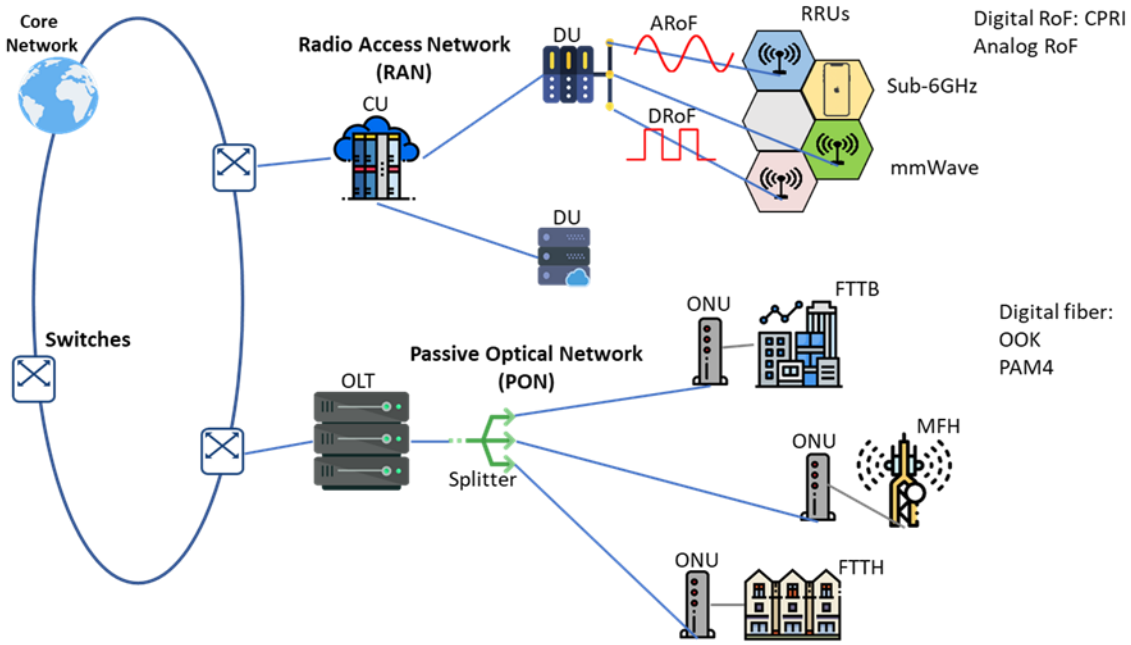


Figure 1.2 Overview of fiber-wireless access network including RAN and PON.

In 5G NR, to disaggregate the conventional BBU and radio tower functions, the 3GPP has defined a series of function split options targeting for use cases with different capacity and latency requirements. As shown in Figure 1.3. the split points of the last three options lie in the PHY layer, where the option 8 represents the conventional D-RoF link with CPRI encapsulation protocol. The option 6 and 7 shift part of the PHY functions to the RRU such that the fronthaul data traffic requirement can be alleviated.

Other higher functional splits are like a distributed RAN configuration, which loses the coordination benefits of C-RAN and reduces the overall network capacity. Due to the extra complexity associated with implementing vendor-specific functional splits, the CPRI-based option 8 is still mostly deployed in spite of the low bandwidth efficiency. Therefore, it's urgent to improve the spectral efficiency of D-RoF for a higher transmission capacity.

Figure 1.3 Logical architecture and function split options of 5G NR.

The A-RoF is envisioned by researchers to be a promising option for MFH with low complexity and latency. The radio signal directly modulates an optical carrier and keeps the optical bandwidth the same as electrical bandwidth, leading to a higher spectral efficiency compared with the D-RoF. According to the radio frequency transmitted in the fiber, two A-RoF types can be defined as shown in Figure 1.4: RF over Fiber (RFoF) and IF over Fiber (IFoF) [11]. In the first type, a modulated RF carrier frequency is converted to light and introduced into an optical fiber so that frequency upconversion is not necessary at the RRU. The RFoF has a better spectral efficiency than the D-RoF, but the optical fractional bandwidth is small as only a tiny portion of the optical bandwidth is utilized by the modulated signal. The other type of radio over fiber transmission carries a modulated intermediate radio frequency (IF) wave, such that carrier aggregation is achieved through frequency division multiplexing (FDM), thus improving optical bandwidth utilization. It is noted that, the IF signal combining/splitting can be achieved

via either analog multiplexer or DSP [12]. The drawback of the IFoF is that dedicated signal upconverters (i.e., mixer) and downconverters (i.e., envelope detector) are required at the RRU. Both means of ARoF provide high spectral efficiency and low RF frontend complexity. With the purely analog processing at the RRH, a reduced latency is also possible. However, the analog radio signal transmitted in the fiber requires a transmission link with a high dynamic range, otherwise the radio signal transmission quality will be severely degraded. Therefore, the pragmatic deployment of A-RoF is still pending on low-cost optical and electrical components with high linearity.

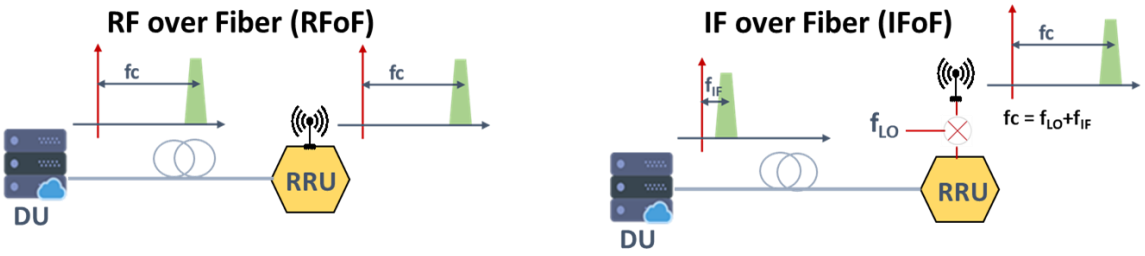


Figure 1.4 Illustration on transport schemes for RFoF and IFoF.

Considering the scarcity of the spectral resources in the sub-6 GHz spectrum, mmWave band are finding their way in the 5G RAN to provide much needed additional bandwidth resources. Recent amended FCC regulations allocate 10.85 GHz of additional spectrum for wireless broadband, including a 3.85 GHz of licensed flexible use spectrum in the 28 to 40 GHz bands and a license-free spectrum ranging from 64 to 71 GHz [13]. Additionally, according to the Further Notice of Proposed Rule Making (FNPRM), 24 to 25, 32, 42, 48, 51, 70 and 80 GHz will also be accessible [14]. The mmWaves can be used to provide high directionality through massive MIMO beamforming. Directionality can provide array gain resulting, high spatial signal density and low inter-cell interference

[15]. An unfortunate characteristic of mmWaves is high atmospheric propagation loss which is compensated by array gain, small cells and frequency reuse. Moreover, with the short wavelengths, the mmWave antenna can be relatively compact and more suitable for integration in a user device.

The mmWave generation in electrical domain is a challenging and costly process due to the inherent bandwidth limitation of the conventional electronics. Therefore, photonics-based generation of mmWave has become an attractive solution to circumvent electrical bandwidth limitations. The basic way to generate mmWave by optical means involves the mixing of two coherent light bundles of different wavelengths such that the difference in wavelengths corresponds energetically to the desired mmWave frequency [16]. The photonics-based generation of mmWave is particularly suited because on wavelength can be continuous wave (CW) while the other may carrier modulation. However, it's limited to the downlink application as it's not feasible to be integrated into a compact user device for uplink transmission. A wide variety of investigations on applying the photonics-based mmWave generation into fiber-wireless networks have been reported. For the V-band (40-75 GHz) frequencies, optical external modulation based methods using optical carrier suppression (OCS) and optical frequency comb (OFC) have been demonstrated in [17-18]. To target tens of Gbps wireless data rate, W-band carrier ranging from 75 GHz to 110 GHz can also be realized using an optical heterodyne detection scheme with an optical local oscillator (LO) [19]. Therefore, photonics-based mmWave generation and propagation are extensively employed in experiments discussed in this dissertation.

1.2.2 Advanced Signal Recovery and Modulation

The modern RAN requires the deployment of high-capacity and lost-cost fiber optical communication infrastructures. Among the available systems, intensity modulation/direct detection (IMDD) is a preferred scheme due to its low power consumption, small footprint, and low cost [20]. On the other hand, those features of the IMDD scheme bring in some drawbacks as well, including limited transmitting power and small channel bandwidth. PAM modulations like OOK and PAM4 are widely used in IMDD since they are simple and robust [21]. Besides, the PAM8 has recently been adopted in short distance optical fiber communication to further increase the transmission capacity [22]. However, the PAM modulation lacks sufficient flexibility in adjusting spectral efficiency, which makes it incapable of fully utilizing the channel resources (i.e., bandwidth and SNR). Originated from coherent optical communications, probabilistic shaping (PS) has become a popular approach to reduce the gap in SNR between the capacity of the optical communication system and the Shannon limit [23-25]. Through shaping the occurrence probability of the constellation points in a modulation scheme, or in other words, increasing the possibility to transmit lower amplitude symbols, we can increase the minimum Euclidean distance among all the constellation points at a given average signal power. Besides, the spectral efficiency of PS modulation can be continuously adjusted by varying the constellation probabilistic distribution. Nevertheless, the PS shaped signal is associated with a higher PAPR, which can easily exceed the small dynamic range of the low-cost IMDD based MFH link, resulting in severe nonlinear impairments. In this case, commonly used linear equalization schemes (i.e., least mean square algorithm (LMS)) can not suffice. Nevertheless, DNN is expected to be able to eliminate the nonlinear impairments efficiently [26-28] because of its superior modeling

capability owing to the multi-layer architecture and nonlinear activation function in each neuron. Compared with conventional Volterra nonlinear equalizer, the DNN is expected to show an improved performance with reduced inference complexity. Moreover, implementing the DNN into the fiber-wireless network is more practical now because of the explosive improvements in efficient nonlinear multi-variable optimization algorithms and the low-cost parallel computing hardware (i.e., graphic processing unit (GPU) and tensor processing unit (TPU)) [29-30]. Although there are some concerns about the DNN complexity, the training time of the DNN to adapt dynamic channel environment can be significantly reduced through transfer learning once the initial training is complete. Moreover, the online querying of the trained DNN, i.e., nonlinear equalization, just involves multiple matrix multiplications with a low time complexity.

Except for the receiver-side equalization, nonlinear degradations can also be mitigated through transmitter-side digital predistortion (DPD) with simple operation and high flexibility [31-32]. The DPD can be implemented in the transmitter of the fiber-wireless network to take advantages of the centralized DSP resources. This scheme also reduces the processing burden and simplifies the receiver design, which allows a cost-effective implementation of nonlinear compensation. Through pre-distorting the transmitted signal, the output of a nonlinear system will be linearized. Till now, the DPD scheme is widely applied in modern transmitters to operate the power amplifier (PA) at a higher efficiency through expanding the overall linear range with a pre-distorted baseband signal [33-34]. Most DPD methods are based on finding the optimal coefficients of nonlinear kernels of the Volterra series. However, the performance of the Volterra-based model can be limited under severe nonlinearities due to high estimation

error on the high-order kernels [35]. Alternatively, the neural networks are well known for the capability to learn arbitrary nonlinear function based on the universal approximation theorem [36]. Considering the DPD is inherently a nonlinear function, it is therefore desirable to utilize the neural networks for the DPD implementation.

1.2.3 Overview of Passive Optical Network supporting D-RoF Implementation

Delivering more bandwidth/capacity has been a top research focus on passive optical networks to support the dramatic data traffic of D-RoF and other fiber to the x (FTTX) applications. To meet this trend, the ITU-T/Full Service Access Network (FSAN) have been developing the standardization of the next-generation passive optical network stage 2 (NG-PON2) since 2015 [37]. As illustrated in Figure 1.5, NG-PON2 is a PON system that exploits time- and wavelength- division multiplexing (TWDM). The optical line terminals (OLT) use different wavelength pairs, while each OLT communicates with multiple optical network unit (ONU) via time-division multiple access (TDMA). Also, NG-PON2 is compatible with the legacy optical distribution network (ODN) with power-splitter based fiber infrastructure. It's detailed that the NG-PON2 requires 40Gb/s/λ or beyond data rate [38]. To comply with the new standard, approaches using advanced modulation formats such as PAM4, Duobinary and discrete multi-tone (DMT) are proposed [39-40]. Compared with the conventional non-return-to-zero (NRZ) signal, data rate can be increased with the same bandwidth. However, those intensity-modulation-direction-detection (IMDD) based methods have a comparatively low modulation index and degraded transmission performance due to the creation of optical double sidebands, as well as the fiber/modulator induced dispersion. Also, the optical phase information is lost during square-law detection in the photodetector (PD). Those drawbacks result in

low receiver sensitivity (>-17dBm), which makes it infeasible to satisfy the critical PON link budgets. Besides, the IMDD scheme has a very limited fiber dispersion tolerance (<30ps/nm), preventing it from being implemented in the desired O+ band (>1320 nm) [41]. Optical coherent detection, on the other hand, utilizes an optical local oscillator (OLO) serving as an optical domain downconverter. This method enables linear optical field detection, which offers significant benefits for digital fiber dispersion compensation and electrical/optical components induced impairment mitigation. Also, a huge signal gain is obtained proportional to the magnitude of the OLO. However, the coherent detection hardware is associated with high-cost from the expensive polarization-diversity hybrid, balanced-photodetector array, etc. On top of this, the PON is sensitive to cost, which limits the massive deployment of coherent detection. Therefore, a lite coherent system combining the benefits of both IMDD and coherent is designed and validated in this dissertation.

The importance of low latency and high reliability are increasing for future PONs, as they will be asked to deliver time critical services like 5G mobile X-haul. Thus, new deterministic and reliable latency management approaches are necessary. For instance, a 1-10 ms one-way latency is required for F1 mobile fronthaul interface, while this number reduces to 100 to few 100 μsec if we move to a lower layer function-split of mobile fronthaul [42]. As a point-to-multi-point system, PON has been one of the dominant architectures to provide bandwidth sharing among different types of services [43]. In general, dynamic bandwidth allocation (DBA) is used in PON to allocate traffic bandwidth in upstream based on the instantaneous demands and requests from users (ONUs). Different DBA algorithms or strategies have been proposed to support the

upstream bandwidth sharing [44-45], however, most of the algorithms are based on a fixed strategy and have no feedbacks from the network environment changes and use-scenario requirements upgrading. Theoretically, different DBA algorithms would be suitable for different use-scenarios or traffic conditions. In addition, the "optimal" DBA scheme for the same network can vary from time to time as the traffic load throughout a day or week can change dramatically. In addition, different users/services may have distinct latency requirements. When the traffic load from each user changes, the corresponding network latency also changes. As mentioned above, many emerging services require more deterministic and reliable latency. Therefore, an intelligent bandwidth allocation that can perceive or sense the network environment changes proactively and correspondingly update its bandwidth allocation policy smartly to manage the latency for different users can be very attracting. Besides, machine-learning based methods with promising performances have been reported on bandwidth and resource allocation in wireless and mobile access networks [46]. It is impactful to investigate machine learning based latency management algorithms in PONs.

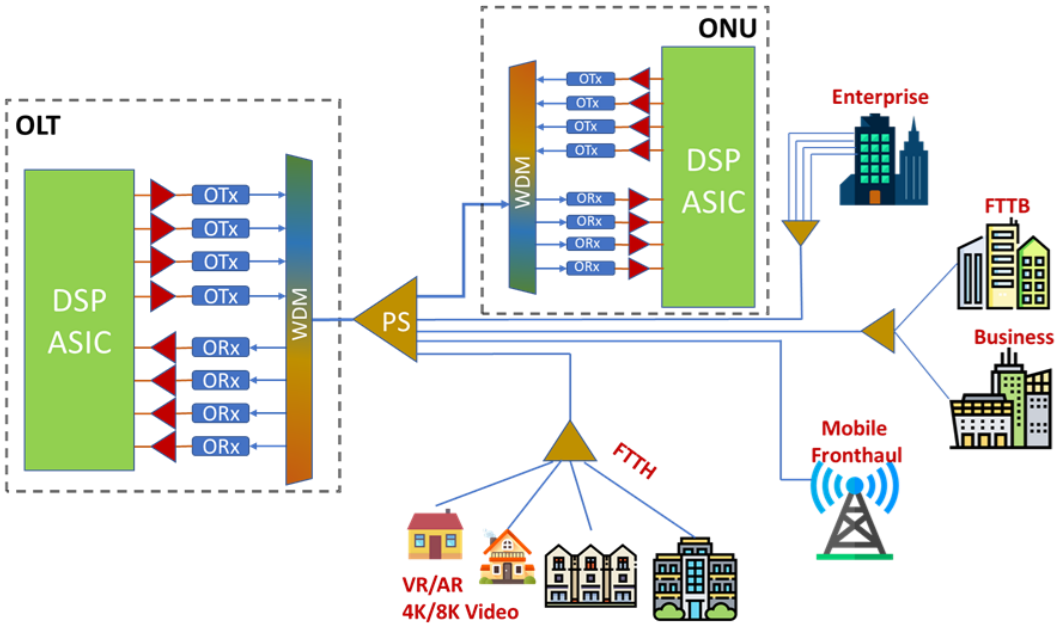


Figure 1.5 Advanced PON architecture.

1.2.4 Interference Avoidance and Cancellation

The scarcity of wireless spectrum has led to intensive frequency reuse and cells densifications in 5G and beyond RANs. Along with the exponentially increasing number of active radio transmitters, the interferences among signals from various sources has become prevalent and unprecedently complex. Generally, as illustrated in Figure 1.6, the interference in RANs may have different origins such as inter-cell interference from a marco/small cell or jamming station, etc. [47-48]. To handle these interferences properly for a higher transmission quality, two steps are typically involved. The first step is interference detection, which detects the presence and characteristics of the interference. The second step is interference mitigation, which can be conducted via interference avoidance or cancellation [49-50]. In a practical system, interference cancellation is hard to accomplish in real time for most of use cases, as it requires the full knowledge of the

interference. On the other hand, interference avoidance can be realized via simpler schemes, for example, frequency shifting or spatial steering. The natures of the macro/small cells from different carriers are mainly independent operations, such that conventional static network planning or allocation is not capable of avoiding interference in a protean network. For fast reaction to environmental change and minimizing the interference caused impairments, a self-organized autonomous interference avoidance scheme is paramount [51]. Reinforcement learning (RL) has attracted tremendous interest and demonstrated its superior performance in many optimizing and strategy-selecting applications, such as the AlphaGo, network congestion control, etc. [52-53]. The RL agent can interact with the environment and obtain feedbacks on the actions it executes. In addition, the agent stores a value network in the form of a Q-table or a neural network, such that the agent can make the most beneficial decision at a certain state based on its previous trial and error. Unlike the supervised machine learning, which requires a large dataset and is time/resource consuming due to the offline training, the RL belongs to the online learning. The agent evolves and becomes more intelligent through the real-time interactions with the environment. These excellent characteristics make RL a promising candidate for interference avoidance. Therefore, we design and experimentally demonstrate a proactive real-time interference avoidance scheme using Q-table based RL.

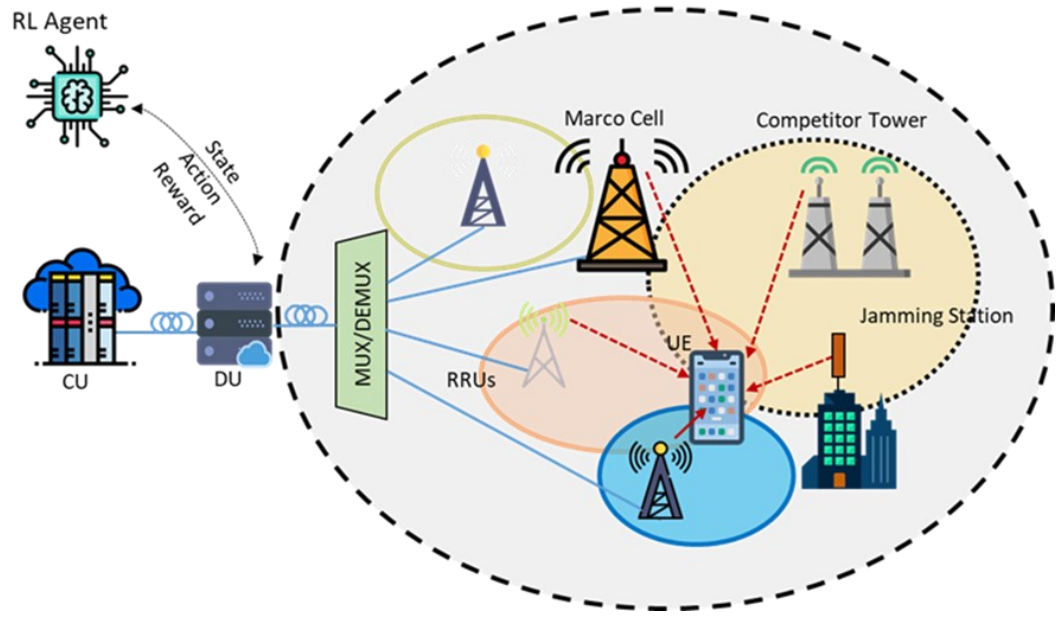


Figure 1.6 Illustration of complex interferences in advanced RANs.

Moreover, to satisfy the dramatically increasing data traffic demand, unlike conventional time-division duplex (TDD) and frequency-division duplex (FDD), in-band full duplex (IBFD) communication has drawn tremendous interest as it uses the same physical resource for bi-directional transmission and could be deployed as a standalone technique or complementing mmWave to increase channel capacity [54]. Inevitably, severe self-interference (SI) from downlink signal will overwhelm the signal-of-interest (SOI), i.e., uplink signal, at the base station receiver due to the approximation of the collocated transmit antenna and receive antenna. Enabling the IBFD relies on efficient cancellation schemes to suppress the Tx SI. The SI cancellation scheme is mainly categorized into passive suppression and active cancellation as summarized in Figure 1.7. In passive suppression, the SI signal is suppressed at the Tx-Rx air interface using antenna configuration and design to lessen coupling between Tx and Rx [55]. In active cancellation, the SI signal is cancelled via subtracting a reconstructed SI copy from the

known Tx signal. The active cancellation can be further divided into analog and digital cancellation, based on the processing domain (analog or digital) of where the SI is subtracted. Various analog interference cancellers based on electronics or optics have been reported in [56-57]. In [57], a 30-dB cancellation depth over 5.5-GHz bandwidth was demonstrated based on an optoelectronics canceller. However, the analog canceller is typically costly and not efficient or scalable to remove the SI after a non-ideal channel response. In this case, digital cancellation can be implemented to cancel the remaining SI components after the mitigation of the dominant SI by an analog canceller.

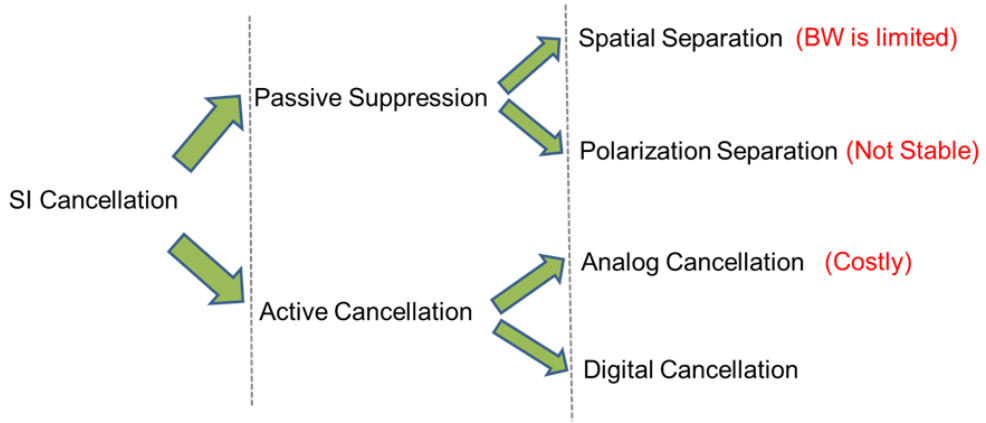


Figure 1.7 Summary of prevailing self-interference cancellation schemes.

The basic idea of the conventional digital SI cancellation is to estimate the channel coefficients between the local transmitter and the receiver. Based on the estimated coefficients, an SI cancellation signal is constructed to subtract the SI from the received signal. The bottleneck of the digital cancellation is the limited quantization dynamic range of an analog-digital-converter (ADC), which requires a low SI-to-SOI power ratio, otherwise the quantization noise of the SOI will increase dramatically. On the other hand, due to the highly directional beams and reduced signal travel range of mm-wave bands,

the SI to SOI power ratio becomes much higher compared to sub-6GHz cases, ensuring a more applicable and reliable digital cancellation for a full-duplex RAN. Nevertheless, previous reported digital cancellation schemes [58-59] lack the analysis and consideration of SI cancellation and SOI recovery when the channels exhibit nonlinearities. In those cases, the remaining nonlinear SI after conventional analog and digital cancellation is still inevitable and degrades the SOI quality significantly. To address those challenges, ordinary neural networks are introduced to construct a non-linear SI cancellation signal [60-61], which has less computation complexity at the inference step (i.e., after training is completed) than nonlinear polynomial-based methods (Volterra series). However, those reports assume that the local receiver is operating linearly such that the SI can be removed by simple subtraction. In realistic cases, any receiver nonlinearity will cause inaccurate SI channel estimation such that the subtraction-based digital cancellation schemes might fail to suffice. As a result, we propose a novel dual-inputs DNN (DI-DNN) to cancel the SI and recover the SOI simultaneously.

Non-orthogonal multiple access (NOMA) has been investigated to utilize power-domain resources and further increase the connectivity to end user equipment (UE) [62]. In contrast to the conventional orthogonal frequency-division multiple access (OFDMA), NOMA schemes allow multiplexing the UE signals in the power domain at the cost of larger inter-UE interference. Successive interference cancellation (SIC) is a common solution that decodes the UE's data sequentially based on power-descending order. However, the SIC inevitably introduces high processing latency when more UE signals are packed in NOMA, while the inherent error propagation will significantly degrade the sensitivity of the subsequent UEs [63]. In this case, parallel interference cancellation (PIC)

which can jointly decode all the UEs' data, becomes a desirable solution. On the other hand, the neural network has proven its efficacy in our work on simultaneous SI cancellation and SOI recovery, which follows a similar decoding process as NOMA except the characteristics of the SI are known. Therefore, the learning-based architecture and deep modelling capacity of the neural network make it a promising solution to facilitate the PIC in the NOMA decoding.

1.3 Dissertation Organization

The dissertation aims to enhance the transmission performance in the fiber-wireless access network through mitigating the vital system limitations of both A-RoF and D-RoF, with machine learning techniques being systematically implemented. Figure 1.8 illustrates the topics covered in this dissertation. The first thrust is improving the spectral efficiency for the optical transmission in the D-RoF to support the delivery of massive number of bits from digitized radio signals. Advanced digital modulation schemes like PAM8, DMT, and probabilistic shaping are investigated and implemented, while they may introduce severe nonlinear impairments on the low-cost optical IMDD based D-RoF link with a limited dynamic range. An efficient deep neural network (DNN) equalizer/decoder to mitigate the nonlinear degradation is therefore designed and experimentally verified. Besides, we design a neural network based DPD to mitigate the nonlinear impairments from the whole link, which can be integrated into a transmitter with more processing resources and power than a receiver in an access network. More than that, we demonstrate a lite coherent system and an intelligent DBA algorithm to further improve the efficiency of PON, which is a readily available infrastructure to facilitate a low-cost D-RoF implementation. Another thrust is to proactively mitigate the

complex interferences in RANs. The composition of signals from different licensed systems and unlicensed transmitters creates an unprecedently complex interference environment that cannot be solved by conventional pre-defined network planning. In response to the challenges, a proactive interference avoidance scheme using reinforcement learning is proposed and experimentally verified in a mmWave-over-fiber platform. Except for the external sources, the interference may arise internally from a local transmitter and share the same time and frequency block as the SOI. Different from the conventional subtraction-based SI cancellation scheme, we design an efficient dual-inputs DNN (DI-DNN) based canceller which simultaneously cancels the SI and recovers the SOI. We also extend a similar scheme for the PIC in NOMA, which effectively cancels the inter-user inference and alleviates the error propagation.

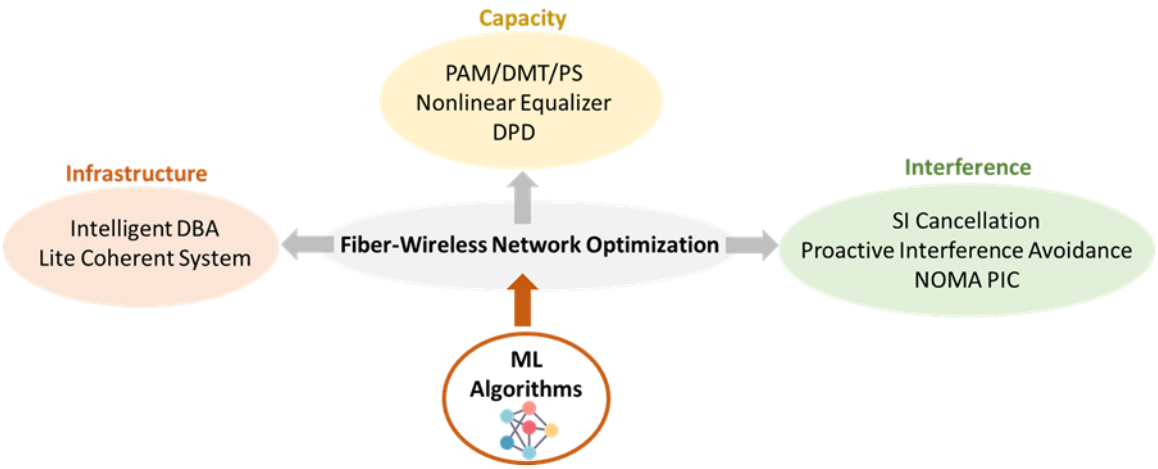


Figure 1.8 Topics and organizations of the dissertation.

After the introduction of motivations and research backgrounds in Chapter 1, Chapter 2 demonstrates the experimental implementations on advanced modulation schemes and nonlinear compensation algorithms. We show that the PAM8 has a higher

spectral efficiency and supports longer fiber transmission distance at the same data rate compared with conventional modulation schemes like PAM4. Probabilistic shaped PAM8 is also verified to be flexible in adapting the channel frequency response. Those schemes are vulnerable to nonlinear degradations, such that we design an efficient DNN equalizer to mitigate the nonlinear impairments. Performance comparison between conventional linear and Volterra based nonlinear equalizer is quantified experimentally. Considering the abundant DSP resources in the transmitter of an access network, we systematically implement a neural network based DPD to perform nonlinear compensation for the whole optical link with PS-DMT modulation. The neural network based DPD shows obvious gain over linear pre-equalization schemes.

In Chapter 3, we investigate the low-cost PON infrastructure and improve its efficiency in terms of capacity and bandwidth allocations to support the D-RoF implementations. A lite coherent system is demonstrated with high receiving sensitivity and exoneration on the dispersion-induced power fading. Besides, an intelligent DBA algorithm based on reinforcement learning is verified to improve the bandwidth utilization than what the commonly deployed Interleaved Polling with Adaptive Cycle Time (IPACT) can achieve.

In Chapter 4, we focus on solving the complex interferences in the 5G and beyond RANs. Firstly, we design a proactive frequency-shifting based interference avoidance scheme through SARSA reinforcement learning. The rough frequency location of the multi-band interferences can be detected via the sub-EVMs. We experimentally validate that the proposed scheme could avoid both the static and dynamic multi-band interferences in real time. Besides, a simultaneous SI cancellation and SOI recovery

scheme is demonstrated. This scheme is based on a DI-DNN instead of the conventionally implemented subtraction-based SI cancellation schemes. Experimental results show the DI-DNN can effectively cancel the SI while recover a clean SOI in both frequency and constellation domains. Moreover, a neural network based PIC scheme is experimentally validated, outperforming the SIC and avoiding the error propagations.

The contributions of this dissertation are summarized in Chapter 5, along with some pointers for future research opportunities and directions.

CHAPTER 2. ADVANCED SIGNAL RECOVERY AND MODULATION

In this chapter, we demonstrate an efficient DNN-based scheme for multi-level signal recovery and nonlinear impairments mitigation in the D-RoF based MHF. Advanced modulation formats are also integrated to improve the spectrum efficiency. Specifically, we investigate the performance of PAM8 and demonstrate its advantages over PAM4. We also take advantage of the flexibility of PS to dynamically adapt to the optical channel response. The improvement on multi-level signal recovery and nonlinear compensation are experimentally verified. Moreover, to better utilize the centralized DSP resources in RANs, a DPD scheme based on the convolutional neural network (CNN) is designed and validated in an IMDD-based optical link.

2.1 Enhanced Multi-Level Signal Recovery

Budget-friendly high-data-rate and long-distance MFH technologies are in an urgent need to support several orders of magnitude higher wireless throughput. Among the available connection technologies, the directly modulated laser (DML) based IMDD system is popular because of its low-cost, robustness and easy implementation [64]. OOK and PAM4 modulation are widely used in IMDD systems. Nevertheless, PAM8 is a promising substitute with a higher spectral efficiency. Besides, compared with PAM4 and OOK, the PAM8 suffers less from the power-fading induced by fiber dispersion, which makes the DML based IM-DD system with PAM8 modulation an ideal candidate for high-throughput and long-distance MFH transmission.

The supervised machine learning has two major categories, the first one is regression. Due to the excellent nonlinear modeling ability, it has been implemented as

an equalizer at the receiver end, which solves a regression problem [65]. However, a hard/soft decoder is required afterward. The DNN has multiple hidden layers and a much better model representation capability. It's much more powerful and resource saving compared with a shallow and wide neural network. The other category of supervised learning is classification. The input layer is still taking previous and subsequently received samples; however, the outputs are decoded samples, which simplifies the equalization-decoding process and eliminates the need for the extra decoders. In this section, we propose and experimentally demonstrate a 60-Gb/s PAM8 transmission over 30-km standard single mode fiber (SSMF) using DNN decoder for enhanced MFH transmission. A pace-setting data-rate and transmission-distance product at 1800-Gbps·km based on the PAM8 modulated 14-GHz class DML IM-DD link is also demonstrated. Our related work has been published in IEEE PTL [66].

2.1.1 Principles and System Setup

Figure 2.1(a) illustrates next generation fronthaul interface (NGFI) structure. The NGFI consists of the CU, DUs, and RRUs. In the CU, a powerful cloud GPU, is implemented to provide initial DNN model training and pass it to DUs to save cost and resources at the DU side. Figure 2.1(b) is the proposed system setup for fronthaul II transmission. The modulator driver is used to provide sufficient driving voltage onto the DML, A 14-GHz class DML is employed as the transmitter to convert the driving signal into the optical domain. The optical signal is then propagated through a 30-km SSMF and detected by a PIN receiver in the DU. To improve the spectral efficiency and extend the power-fading limited transmission distance at a certain data-rate, we implement PAM8 as the modulation scheme. Though a clear benefit of PAM8 is that it has more signal levels than the traditional and widely utilized OOK and PAM4, additional signal levels cause severe degradation of signal to noise ratio (SNR). To alleviate the impact of a lower SNR, a higher driving voltage is necessary. However, this requirement imposes strong

nonlinear effects due to the limited dynamic range of the modulator driver and the DML. In this case, the commonly used linear least mean square algorithm (LMS) is not able to eliminate the nonlinearities. This is where a DNN has the advantage, thanks to its ability to accommodate a massive number of parameters, Even the most complex function can be explicitly represented by a DNN. In addition, the increasingly efficient nonlinear optimization algorithm and the exponentially growing capabilities and widening use of parallel processing through GPUs and tensor processing unit (TPU) make DNN more practical to be implemented into the MFH applications.

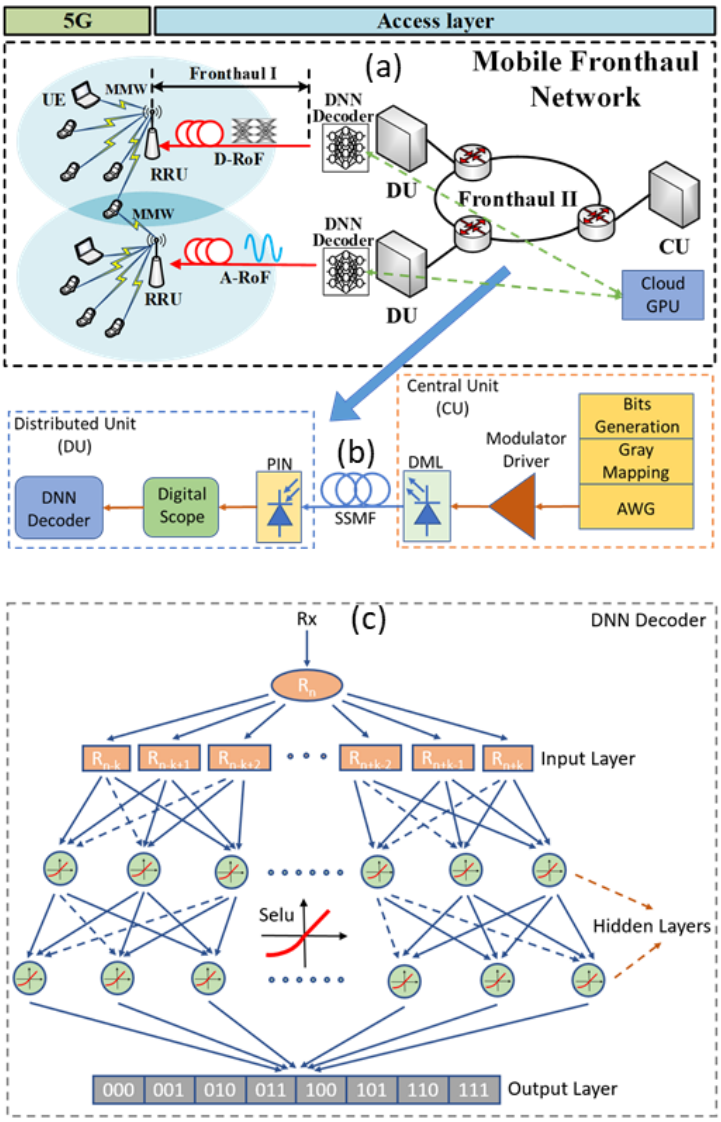


Figure 2.1 (a) NGFI structure; (b) Experimental system setup; (c) DNN decoder structure.

Figure 2.1(c) shows the structure of the proposed DNN decoder. We take the input sample with its 64 previous and 64 subsequent samples as the input layer. Besides, we have two hidden layers; each layer has 1024 neurons, the output layer consists of 8 nodes, each node corresponds to a PAM-8 level as shown in the gray boxes. At each hidden layer, we use the Selu activation function; the Selu function is defined as:

$$Selu(x) = \lambda \begin{cases} x & if\ x > 0 \\ \alpha e^x - \alpha & if\ x \leq 0 \end{cases}$$

(2.1)

Since the gradient is the constant λ when the input x>0, this function can effectively eliminate the gradient vanishing problem. A vanishing gradient will cause the DNN to become incapable of converging because the gradient will not be able to decrease at an upper hidden layer. On the other hand, the gentle gradient when x $\leqslant$ 0, minimizes the variance of the input and effectively eliminates gradient exploding problem. At the output layer, the activation function is the Softmax function. The function is given by:

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}}$$

(2.2)

Instead of just selecting the maximum value, the Softmax function outputs all the input values with respective probabilities to be selected. The probability of each input can be derived from the formula (2.2); larger input values will have a higher probability at the output. The proposed DNN loss function is categorical cross-entropy as defined below:

$$J(w) = -\frac{1}{N}\sum_{n=1}^{N}[y_n log\hat{y}_n + (1 - y_n)\log(1 - \hat{y}_n)]$$

(2.3)

The combination of the Softmax function and categorical cross-entropy make the loss calculation easier, which enables a more convenient way to derive the partial derivatives for back-propagation [67]. Moreover, we implement the Adamax algorithm proposed by [68], an algorithm for first-order gradient-based optimization of stochastic objective functions, based on adaptive estimates of lower-order moments. This new algorithm implements momentum during each training epochs. The momentum acts as an accelerator to find the global minimum instead of limiting the scope to the local minimum. Though the proposed DNN provides a sufficient number of parameters to find the correlations between the received symbols and the transmitted symbols, overfitting can result because of the limited size of the training set. To alleviate overfitting, we integrate the dropout layers in the DNN as shown by the dashed blue lines in Fig.1(b), which deactivate 20% of neurons during each epoch. In addition, a Max norm kernel constraint is set with the dropout layers to avoid possible strong connections between the neurons of different layers. Besides, an early stopping mechanism is implemented to monitor the accuracy of the validation. The validation set is a subset of the training set. We use it to measure the DNN's generality at each epoch instead of using it for model training. When the validation set accuracy doesn't improve over a certain number of epochs, we stop the training, and the model becomes the most generalized.

2.1.2 Experimental Results

In the experiment, we generate a pseudo-random 8-level sequence, which is downloaded to an arbitrary waveform generator (AWG) with 65 GSa/s sampling-rate. The initial source sampling rate is set to 20 GSa/s which is equivalent to a 20-Gbaud PAM8 signal. The output signal voltage of the AWG is 180 mV, which is amplified afterwards by a 30-dB fixed gain modulator driver (SHF 810) used to modulate a 14-GHz

class DML. The driving current of the DML is set to 150 mA to provide sufficient optical power for long distance transmission. The output wavelength of the DML is 1543.2 nm with 9.5-dBm optical power. The output signal of the DML is then propagated through a 30-km SSMF and detected by a PIN receiver (Finisar XPDV2020R). At the receiver end, an 80-GSa/s real-time oscilloscope is used to sample the analog signal with 8-bits quantization precision. The sampled data is used for training DNN decoder model and calculating BER offline. Firstly, we resample the data to 20 GSa/s which matches the source sampling rate, then we perform a correlation between the transmitted signal and resampled signal to achieve synchronization. The first 64k synchronized data is fed to the DNN for training, 20% of them are used as the validation set during each epoch. After the validation set accuracy stops improving, after a certain number of epochs, the program will stop the training automatically and save the best DNN model. The best model is used to decode the resampled data. This DNN decoder integrates the equalization and decoding process, the output is one of the PAM8 levels. The signal recovery logic is greatly simplified from the initially two steps to one single step. The BER is calculated by comparing the difference between the original and recovered samples. Figure 2.2(a) demonstrates the BER performance over different data rate. We scan the source sampling rate from 15 GSa/s to 25 GSa/s with 1 GSa/s per step, which corresponds to data rates from 45 Gb/s to 75 Gb/s using PAM8 transmission. From Figure 2.2, it can be inferred that the system reaches 2.69×10^{-3} BER at 60 Gb/s over 30 km, which is below the 7% overhead forward error correction hard decision (FEC-HD) threshold of 3.8×10^{-3}. This is the first demonstration of 1800-Gbps·km data-rate transmission-distance product based on a 14-GHz class DML IM-DD system without dispersion compensation. We also compare the performance difference between the DNN decoder and the Volterra series based nonlinear equalizer + Hard Decision (VNLE-HD). The latter receiver equalizer has 17-tap second order Volterra series and 128-tap linear filter. The DNN outperforms it with a huge margin, i.e., a 14.7 Gb/s data rate boost (30.6% of original data rate) is

resulted at the FEC-HD threshold. In addition, we measure the system sensitivity based on both DNN and VNLE-HD. The results are summarized in Figure 2.2(b). The data rate is fixed at 60 Gb/s using PAM8 transmission and propagates through a 30-km SSMF.

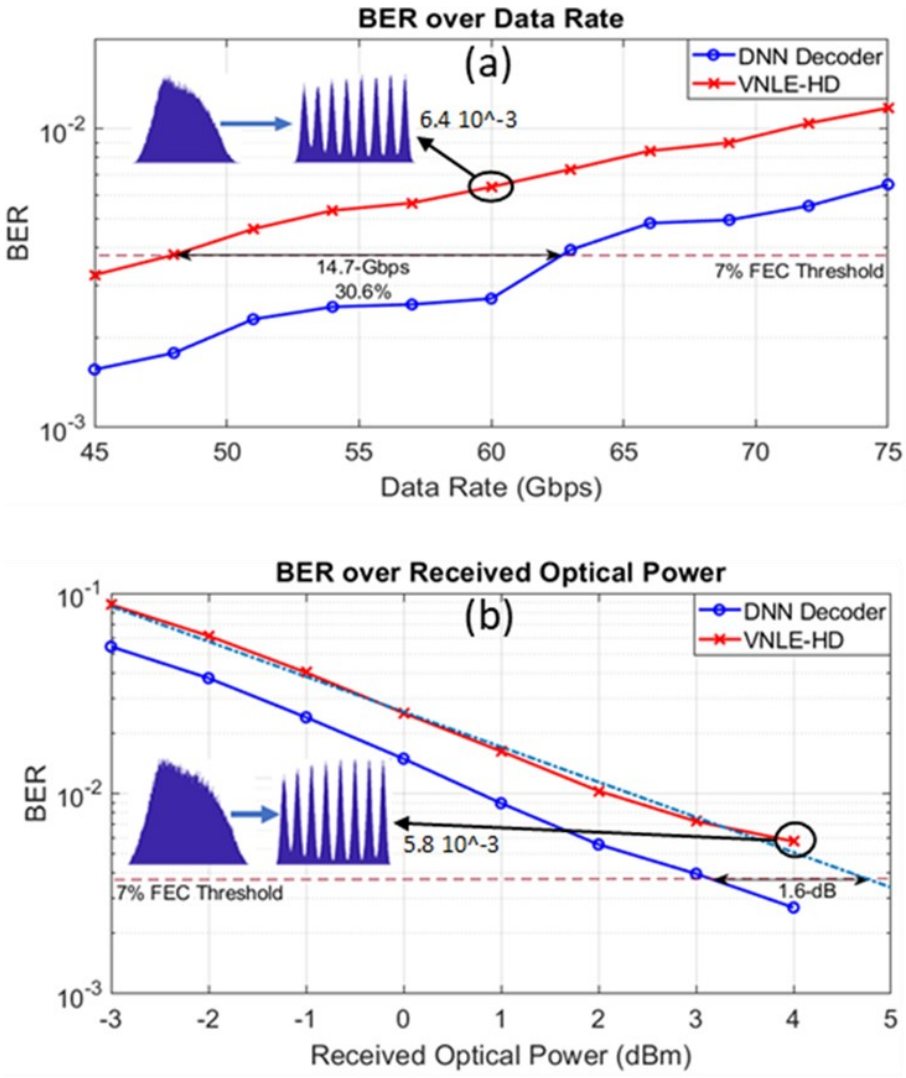


Figure 2.2 (a) BER comparison over data rate. (b) BER comparison based on received optical power.

The received optical power (ROP) is varied from -3 dBm to 4 dBm. Here, the DNN also shows superior sensitivity performance over VNLE-HD. The VNLE-HD can never reach the FEC-HD threshold at 60 Gb/s over 30 km even at the highest ROP. For a

better comparison, we use a minimum mean square error algorithm (MMSE) to fit the BER over the entire range of ROP used in this experiment. The performance of a VNLE-HD is indicated as the dash-dot blue line in Figure 2.2(b). As shown by the double arrow in the same figure, at 7% overhead FEC threshold, a 1.6 dB sensitivity gain is obtained over the VNLE-HD. The insets for PAM8 recovery illustration in Figure 2.2 are the received signal histograms before and after the VNLE.

We also measure the RF response of the full link, which includes the modulator driver, DML, 30-km SSMF and the PIN receiver. The result is illustrated by the red curve in Figure 2.3. Due to dispersion, deteriorated double-sideband (DSB) beating and DML chirp, a frequency dip is observed at 15.2 GHz, which causes a strong signal degradation around the frequency dip. The demonstrated link data rate of 60 Gb/s over 30-km transmission using PAM8 is not trivial. Using PAM4 it can be even more difficult to transmit over 30 km with a dispersion limited system. As shown by the black curve in Figure 2.3, 60-Gb/s PAM4 has strong signal spectrum leakage after 15.2 GHz, which right at the dip in the RF frequency response. Therefore, the power fading causes significant information entropy loss, which causes extreme difficulty to recover the PAM4 signal. PAM8 modulation breaks this barrier since the source spectrum, as illustrated by the dashed blue curve, is entirely contained for frequency below the dip at 15.2 GHz. This happenstance is key to support 60 Gb/s transmission. The main limitation of using PAM8 is related to SNR requirement which is even worse when system power budget is limited. Here, the power budget is 6.5 dB, which is limited by the low responsivity of the PIN receiver. Using the APD receiver instead will significantly boost the system power budget. Furthermore, optical dispersion compensation could be implemented to increase dispersion tolerance in MFH.

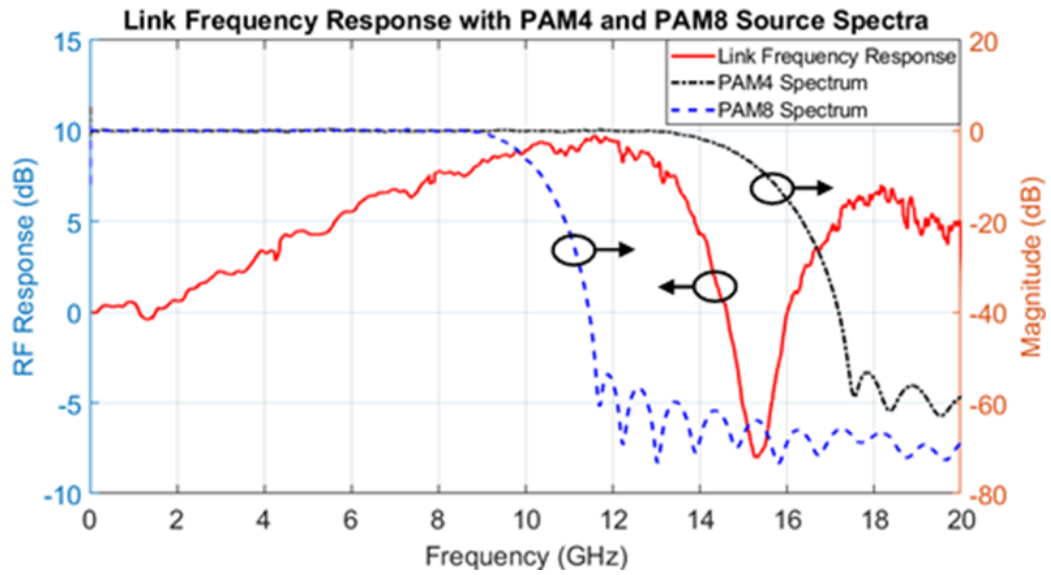


Figure 2.3 Link frequency response measurement and PAM4/PAM8 source spectra.

Training DNN to high accuracy from scratch is resource consuming. However, once we get an initial optimized model, only a small number of training epochs are required to regain the optimal model over a limited range of dynamic channel conditions. To verify the stability of the DNN decoder, we arranged to receive data every 6 minutes for a consecutive period of 3 hours. Figure 2.4 illustrates the BER over time with only 20 training epochs for each point. We also include the VNLE-HD for comparison. For the DNN decoder, the BER varies between $2.7\text{x}10^{-3}$ to $3.5\text{x}10^{-3}$, all below 7% FEC-HD threshold. However, for VNLE-HD, BER falls well outside this range scoring values from $6\text{x}10^{-3}$ to $7\text{x}10^{-3}$. These results support our contention that a DNN can be trained quickly for the circumstances at hand once an initial optimized model is obtained. The computational complexity of the DNN can be quantified by the number of hidden layers [69]. Here, the computational complexity of DNN is $O(n^2)$, which is comparable to that attributed to a VNLE-HD. Rapid DNN training is critical for MFH transmissions because (1) the DU doesn't have to commit too many scarce resources for DNN training to recover a signal since the initial training can be performed by the cloud GPU at the CU

and then pass the trained DNN down to the DU as a baseline. At the DU, the DNN can be refined to suit with minimal training epochs. The DNN model is robust enough to take care of the expected dynamic channel variations.

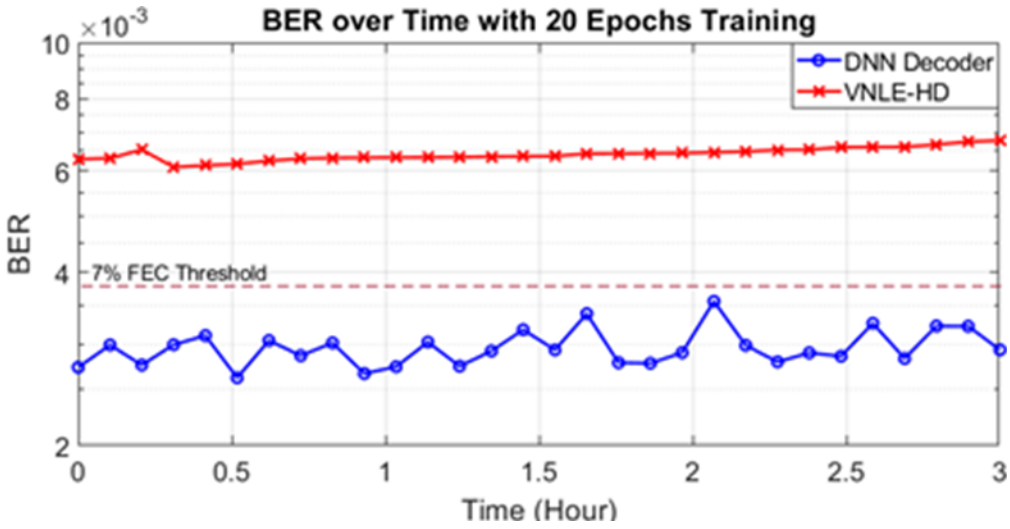


Figure 2.4 BER over time performance verification for DNN decoder.

2.1.3 Summary

In this section, we have proposed and experimentally demonstrated a DNN decoder for MFH transmission. A record-high data-rate-x-transmission-distance product of 60 Gb/s over 30 km based on a 14-GHz class DML IMDD system is achieved. The experimental system utilizes the PAM8 modulation, which is more spectrally efficient than traditional PAM4 and OOK modulations. Although PAM8 is less sensitive to power fading than PAM4 + OOK at the same data-rate, PAM8 leads to stronger nonlinearity and a higher SNR requirement due to higher chirp. This potential problem is resolved thanks to the effectiveness of the multiple parameter data fitting capability of the DNN decoder and the rapid training for a dynamic environment once the primary DNN is trained offline. The DNN decoder utilizes the limited PAM8 SNR to reach the possible minimum BER. Moreover, the system's BER test over different ROP and data rate is analyzed and

compared with the VNLE-HD, and the superiority of the DNN decoder over VNLE-HD is demonstrated. Though the DNN training from scratch would be prohibitive, a much-reduced training epoch need be involved at the DU once an initial optimized model is obtained offline at the cloud GPU in the CU. Our scheme provides a promising low-cost enhanced-performance solution for MFH transmission.

2.2 Integration of PS-PAM8 and DNN to Mitigate Capacity Crunch

To satisfy the dramatic bandwidth requirements from data-thirsting 5G services, advanced modulation format with higher spectral efficiency and flexibility are necessary in RANs. PS-PAM8 provides the flexibility to adapt the channel response and an improved entropy. However, it can easily overwhelm the limited dynamic range of an optical link. Therefore, we efficiently take advantage of both PS to utilize the channel resources and DNN on nonlinear signal equalization/decoding, and experimentally demonstrate a capacity-approaching transmission in the D-RoF. A proof-of-concept experiment is conducted to verify the efficacy of the proposed scheme with an 80-Gbps PS-PAM8 signal over 20-km standard single mode fiber (SSMF) transmission based on SD-FEC. Our related work has been published in [70].

2.2.1 Operating Principles and Experimental Setup

The Figure 2.5(a) shows the experimental setup of the proposed scheme. Firstly, a uniformly distributed binary bit sequence is randomly generated and fed into a distribution matcher (DM) for PS. The implemented DM is based on the constant composition distribution matcher (CCDM) [71]. The PS symbols output by the CCDM follow the Maxwell-Boltzmann distribution with flexible entropy through adjusting the constellation point occurrence probability as given by:

$$P(a_i) = \exp(-\lambda {a_i}^2) \Big/ \sum_i (-\lambda {a_i}^2)$$

(2.4)

where a_i is the amplitude of a constellation point, and λ is a rate parameter. The shaped symbols are then PAM mapped and converted to an analog signal as shown by inset (i) through a 64-GSa/s arbitrary waveform generator (AWG, M8195A). The modulator driver is used to provide sufficient signal voltage swing to drive the DML (SCMT-100M11G), while the DML is used as the transmitter in the CU to convert the driving signal into the optical domain with a center wavelength at 1548.3 nm as illustrated in inset (ii). The optical signal propagates through a 20-km SSMF and detected by a PIN receiver (SCMR-100M11G) in the DU. The detected signal is digitized using an oscilloscope with 80-GSa/s sampling rate (DSOZ254A) for the following digital signal processing. Though the PS-PAM8 signal with an information rate of 2.5 bits/symbol has a larger minimum Euclidean distance than the uniform PAM8 modulation at a certain average signal power, it also introduces a higher PAPR. The peak of the signal can easily exceed the small dynamic range of the low-cost electrical and optical components in the MFH and make it vulnerable to nonlinear impairments. Figure 2.5(b) shows the structure of the implemented DNN decoder. We take the input symbol with its 80 previous and 80 subsequent symbols as the input layer. The DNN consists of 2 hidden layers with 1024 neurons at each layer. There are 8 neurons at the output layer, each one corresponds to a constellation point of the PAM8 modulation. We use the Selu function as the activation function at the hidden layers which can effectively eliminate the gradient vanishing and gradient exploding problems that commonly result in DNN's incapable of convergence.

The activation function at the output layer is Softmax given by: $\sigma(z)_j = e^{z_j}/\sum_{k=1}^{K} e^{z_k}$. The Softmax function is employed to calculate the probability of each DNN output values, while categorical cross-entropy defined as: $J(w) = -1/N \sum_{n=1}^{N}[y_n log\hat{y}_n + (1 - y_n)\log(1 - \hat{y}_{n)}]$, is implemented for loss calculation. In the combination of the Softmax function and the categorical cross-entropy, the calculation of the partial derivative for back-propagation is more concise. Besides, we utilize the Adamax algorithm to train the DNN, which can efficiently avoid the local optima. Another severe problem associated with DNN is overfitting. We alleviate it through integrating dropout layers in the DNN as shown by the dashed blue arrows in Figure 2.5(b), which deactivates 20% of neurons during each training epoch. We also divide the received symbols into 3 sets, namely, training set, validation set, and testing set with a ratio of 0.6, 0.1, and 0.3, respectively. The validation set is used to measure the validation accuracy over the whole training process. When the validation set accuracy is not improving over 100 epochs, we will stop the training to avoid severe overfitting.

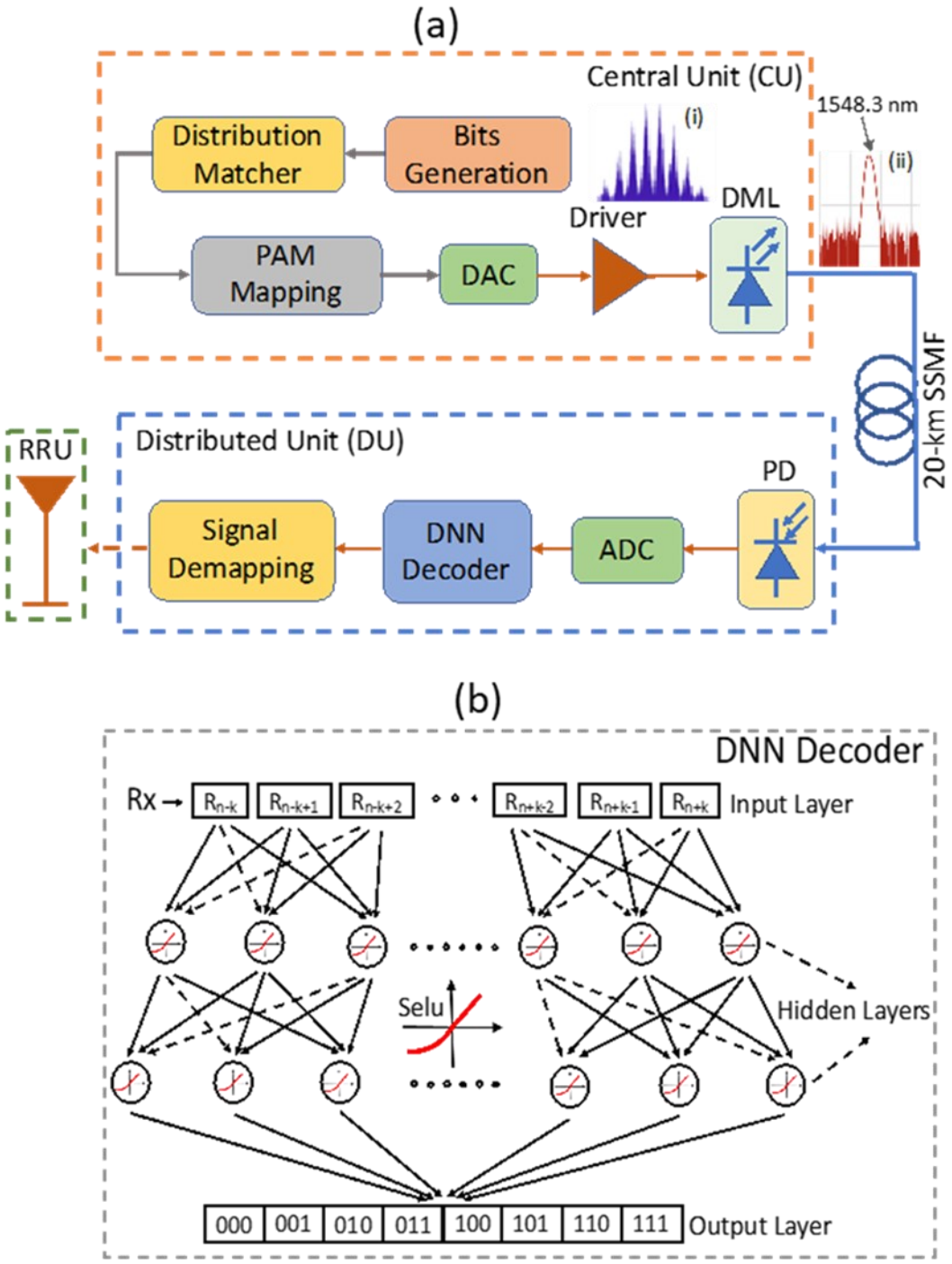


Figure 2.5 (a) Experimental setup; (b) structure of the DNN decoder.

2.2.2 Experimental Results and Evaluations

The frequency response of the end-to-end system configured as Figure 2.5(a) is measured by a 20-GHz vector network analyzer as shown by the red curve of Figure 2.6(a). The measured S21 response is in between the input to the modulator driver and the PIN receiver output. As observed, the S21 response is relatively flat below 16 GHz, while it degrades dramatically beyond that frequency. For comparison, we also plot the spectra of PS-PAM8, PAM8 and PAM4 with a bit rate of 80 Gb/s in Figure 2.6(a) as indicated by blue, purple, and green curve, respectively. Due to the low spectral efficiency of PAM4 modulation, the spectrum of PAM4 signal is far beyond the bandwidth of the system. However, since we can flexibly set the entropy of PS-PAM8 signal, the PS-PAM spectrum closely matches the system frequency response, which maximizes the bandwidth utilization. As for the PAM8 signal, though it is within the bandwidth limit of the system, the Euclidean distance between adjacent constellations is too small to reach acceptable BER performance because of the limited SNR in the low-cost IMDD system. To set up a baseline in reference to the discussed modulation schemes, we firstly measure the optical back-to-back (BtB) transmission performance of 80-Gb/s signals based on IMDD with an optimized LMS post-equalization. During the measurement, the received optical power (RoP) is scanned from -4.5 dBm to 4.5 dBm with 1-dB increment. The measured BER curves for PS-PAM8, PAM8 and PAM4 are shown in Figure 2.6(b) as blue, red, and yellow solid curves, respectively. The SD-FEC and HD-FEC thresholds are also plotted as reference lines. For the PAM4 modulation, the BER cannot go below either of the thresholds for all the measured RoP values due to the insufficient system bandwidth. In the PAM8 modulation case, the BER becomes

lower as the RoP increases and passes the HD-FEC threshold at -0.6-dBm RoP. However, the PS-PAM8 achieves a -3.4-dBm receiver sensitivity and demonstrates a prominent 2.8-dB gain over PAM8 modulation at the HD-FEC threshold.

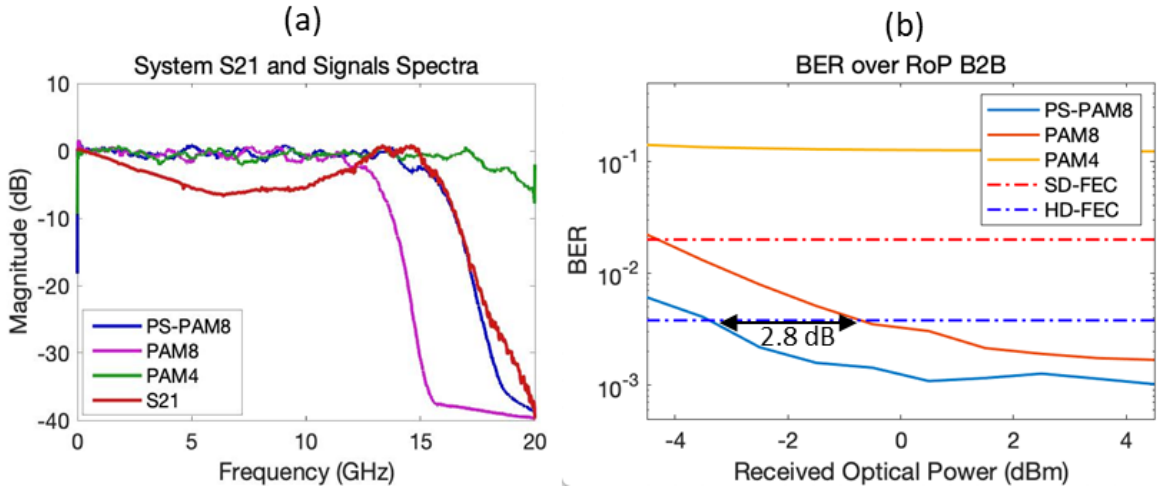


Figure 2.6 (a) Measured system S21 response and signals spectra. (b) BER over RoP at B2B case.

In addition, the BER curves after 20-km SSMF transmission are also measured as shown in Figure 2.7(a). The PS-PAM8 demonstrates a remarkable 4.1-dB gain compared with PAM8 at SD-FEC threshold. As discussed in the former section, the higher power efficiency of the PS signal is at the cost of increasing the PAPR. In the low-cost MFH system, the high PAPR can easily exceed the limited dynamic range of the transmission link and resulting strong nonlinearity. The DNN decoder is an efficient method for mitigating nonlinear impairments. To verify the performance, we randomly generate a 100k symbol length PS-PAM8 signal in the CU. At the DU, after resampling and resynchronization, we take the first 60% of the received symbols as training set, the following 10% as validation set, and the remaining 30% is test set for the DNN. We

monitor the validation-set accuracy while training the DNN, once validation-set accuracy stops improving over 100 epochs, the algorithm will stop the training automatically and save the most current DNN model parameters. The trained DNN decoder will then decode the test-set data. Figure 2.7(b) demonstrates an experimental verification on the DNN's performance. The blue curve is the baseline BER performance over RoP with the LMS post-equalization, while the red curve indicates the performance of the DNN decoder. At the SD-FEC threshold, the DNN decoder demonstrates a noteworthy extra 3.2-dB gain over the baseline, which proves the DNN's efficacy on eliminating the probabilistic shaping's drawback of nonlinear impairment, and on significantly improving the overall transmission performance from the CU to DU.

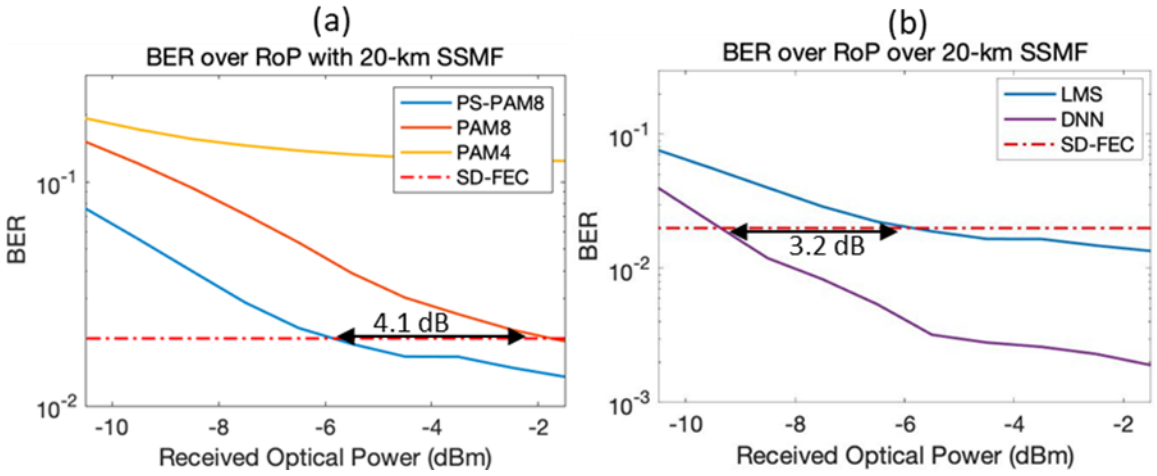


Figure 2.7 BER over RoP (a) 20-km SSMF transmission. (b) With and without DNN.

2.2.3 Summary

We propose and experimentally demonstrate a capacity-approaching transmission in 5G MFH based on PS-PAM8 modulation and DNN decoder. The PS-PAM8 signal is power efficient and is flexible in information rate to adapt to varying channel conditions.

However, probabilistic shaping introduces a higher PAPR and makes the PS signal vulnerable to the limited dynamic range of the hardware components in 5G MFH resulting in severe nonlinear impairments. We subsequently implement a DNN decoder for nonlinear compensation to significantly improve the overall transmission performance from the CU to DU with a 3.2-dB extra gain at SD-FEC threshold. Adding up the 4.1-dB gain from PS-PAM8, a 7.3-dB gross gain is realized compared to conventional uniform PAM modulations with linear post-equalization. Our scheme offers a promising solution to mitigate capacity crunch in the D-RoF MFH.

2.3 DPD Enhancement by CNN for PS-DMT Transmission

To circumvent the bandwidth limitation, high spectral efficiency modulation schemes like discrete multi-tone (DMT) has been deployed [72]. The DMT can provide higher tolerance to chromatic dispersion, compatibility to low-cost IM/DD systems, and flexibility for multiple access. Moreover, the probabilistic shaping (PS) has been developed for adoption in access network to mitigate capacity crunch [70] as it can customize entropy and enhance sensitivity by tuning the distribution probability of the constellation. However, both DMT and the PS are subject to high peak to average power ratio (PAPR), which easily overwhelms the limited dynamic range of the low-cost IMDD based optical link even under regular operation condition, such that the nonlinear impairments will be inevitably present in transmission systems. Several nonlinearity mitigation techniques including receiver-side equalization and transmitter-side digital predistortion (DPD) are available as potential solutions. Among them, the DPD is the most effective and precise scheme with simple operation and high flexibility. In this section, we efficiently integrate convolutional neural network (CNN) based DPD for

nonlinearity mitigation and PS-DMT to exploit the limited channel capacity. A 68.2-Gb/s net data-rate transmission over 15-km standard single-mode fiber (SSMF) is experimentally demonstrated using 11G-class devices. Besides, the proposed scheme attains 1.1-dB sensitivity gain over conventional linear pre-equalization in a proof-of-concept experiment. The contents of this section are derived from our work published in OFC [73].

2.3.1 Operating Principles

The PS sequence for DMT modulation is generated via a constant composition distribution matcher (CCDM) [71]. The output sequence follows Maxwell-Boltzmann distribution, and the entropy can be adjusted through tuning the rate parameter λ as given by: $P(a_i) = \exp(-\lambda {a_i}^2) \,/\, \sum_j (-\lambda {a_j}^2)$, where a_i and a_j are available symbols from the alphabet $\{ai\}$. In this case, the symbol with a higher amplitude will have a lower occurrence probability to save signal power. For the DPD part, we implement indirect learning architecture, which reduces half of the complexity for system identification than direct learning. As shown in Figure 2.8, the discrete-time post-distortion function is firstly estimated and then copied to the pre-distortion function. The post-distortion function is equivalent to the inverse of the system transfer function with the sampled output y_n as the input and the post-distortion output:

$$\hat{d}_n = \widehat{H}_{post}^{-1}(y_n) \tag{2.5}$$

The $\hat{H}_{post}^{-1}$is derived through minimizing the error defined as: $e_n = d_n - \hat{d}_n$, while the predistortion output d_n is equal to input sample x_n at the beginning of the learning process. The pre-distortion function then copies the post-distortion function such that:

$$y_n = \hat{H}_{CH}(\hat{H}_{pre}^{-1}(x_n)) = G \cdot x_n \tag{2.6}$$

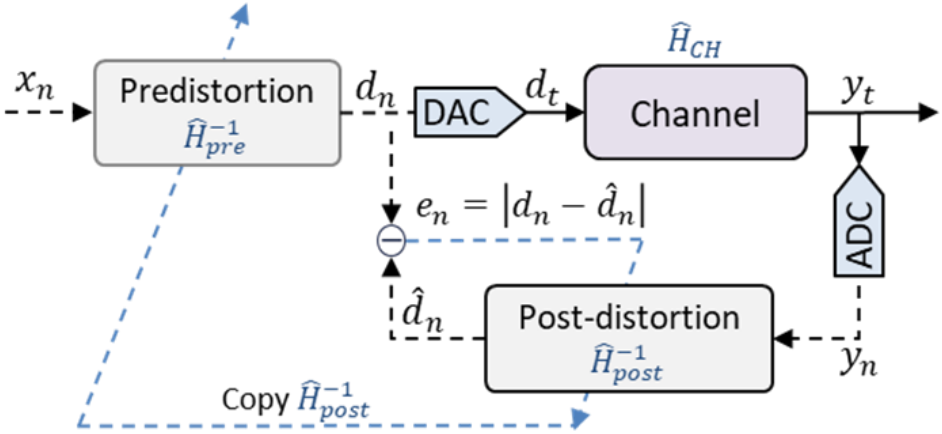


Figure 2.8 Indirect learning process.

where G is a scalar corresponding to the average gain of $\hat{H}_{CH}$. In this case, the channel output is linearized. The regression process to derive the $\hat{H}_{post}^{-1}$ can also be realized via neural networks such as CNN and multilayer perceptron (MLP), converting the problem to hyper-parameter optimization. Here, CNN is implemented due to its lower model and computation complexity from kernel coefficients sharing. The architecture of CNN is demonstrated in Figure 2.9. The φ_y is the 1-D tensor consists of the currently received sample y_n and its pre-cursor/post-cursor samples. There are two 1-D convolutional layers with 10 different kernel filters for each layer followed by batch normalization to suppress overfitting and max pooling for downsampling. The output of the Conv2 is flattened and

serves as the input to a multilayer feedforward neural network with two hidden layers. The first hidden layer has 20 neurons while the second has 10 neurons. The ReLU function is used at each layer for activation, while the mean square error (MSE) is used as the loss function.

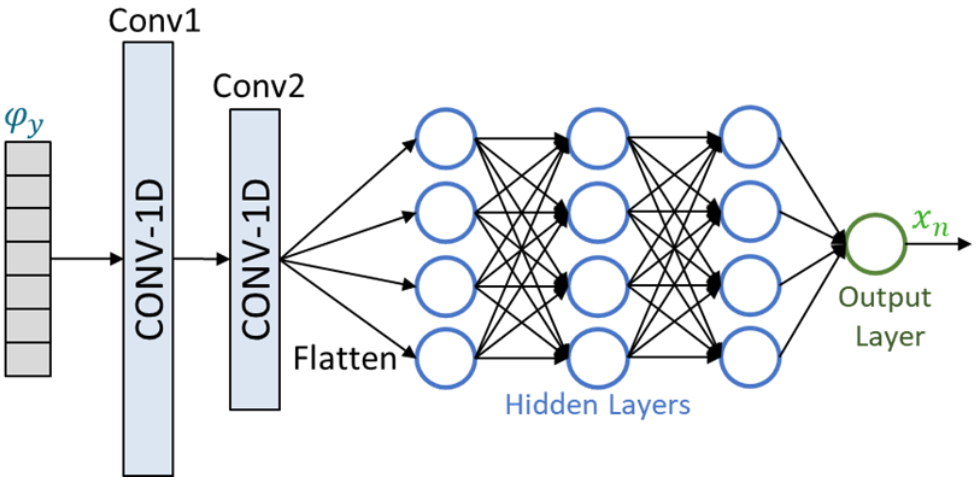


Figure 2.9 CNN architecture for DPD implementation.

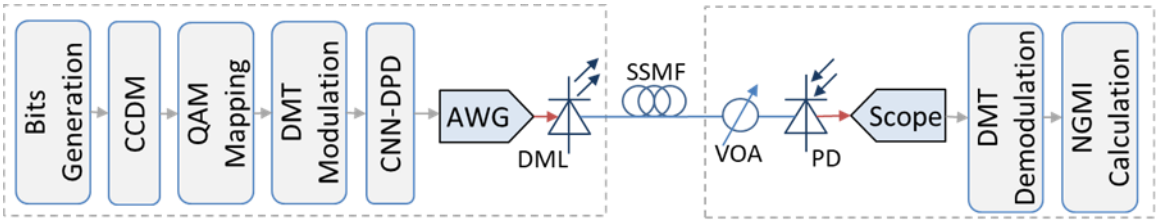


Figure 2.10 Experimental setup for DPD implementation.

2.3.2 Experimental Setup and Results

Figure 2.10 depicts the setup of the proof-of-concept experiment. A uniformly distributed binary bit sequence is randomly generated by MATLAB, which is then mapped to PS-64QAM through CCDM and QAM modulation. The entropy is set to 5.6 for optimized channel SNR utilization. The same entropy applies to all the data subcarriers without

using bit loading to reduce processing complexity. The DMT modulation process is followed, with 4096-point FFT, 15.36 MHz subcarrier spacing, and 1026 active data subcarriers. The DMT setting is determined by the measured end-to-end S21 response of the system as shown in Figure 2.11 to exploit the limited bandwidth. Although the 3-dB bandwidth is only 2.8 GHz, the S21 response is not dropping severely until 15.4 GHz. The DMT signal is then digitally pre-distorted and converted to an analog signal using a 64-GSa/s arbitrary waveform generator (AWG, M8195A). An 11G-class directly modulated laser (SCMT-100M11G) with pre-amplifier and transimpedance amplifier (TIA) is implemented at the transmitter to convert the signal into the optical domain with 1548.2 nm center wavelength and 8.56 dBm output power. The optical signal transmits over a 15-km SSMF and is detected by an 11G-class photodetector (SCMR-100M11G) at the ONU. We employ an 80-GSa/s oscilloscope (DSOZ254A) to capture and digitize the detected signal for receiver end processing including iFFT, zero-forcing equalization, etc.

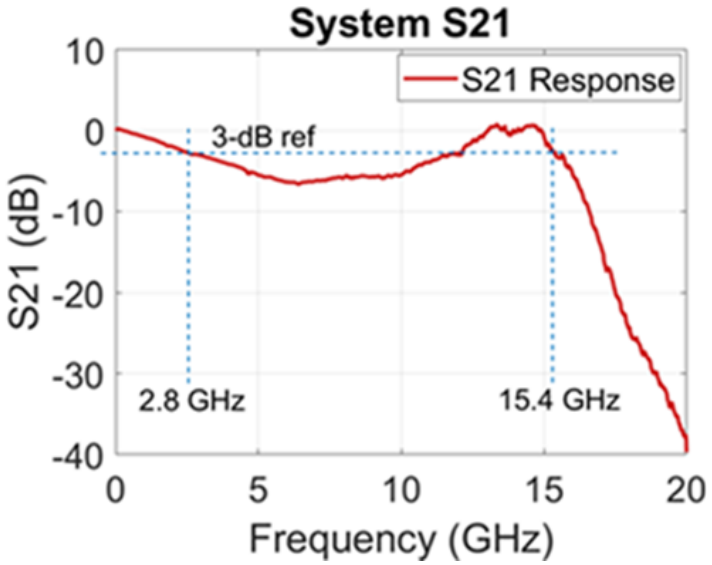


Figure 2.11 Measured end-to-end system S21 response.

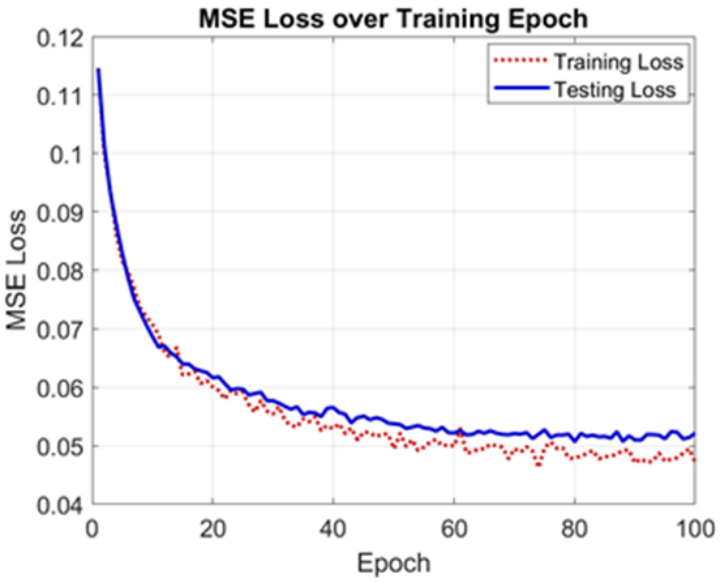


Figure 2.12 Training and testing MSE loss over training epoch.

Figure 2.12 shows the training and testing loss of the CNN during the indirect learning process using PyTorch platform. Among the 557036 received digitized samples, the first 70% percent is used as the training set while the remaining 30% percent is separated off as the testing set. The testing loss drops along with the training loss with only a small deviation, justifying there is no overfitting issue. The trained CNN is then applied in the DPD process and newly captured data are used for performance evaluation and comparisons. Figure 2.13 compares the received spectra of CNN-DPD, linear pre-equalization, and without DPD. The CNN-DPD pre-compensates both the linear and nonlinear impairments, up to 4.3-dB more nonlinearity mitigation is observed compared to linear pre-equalization, while the latter only equalizes the linear response. To evaluate the realistic transmission performance, we measure and calculate the normalized generalized mutual information (NGMI) serving as a pre-FEC metric, over received optical power (RoP) sweeping from -9.8 dBm to 0.2 dBm. Figure 2.14 (a) demonstrates the back-to-back (B2B) measurement. The NGMI difference at low RoP is small due to

the inaccurate parameters' estimation under low SNR. However, at 0.89-NGMI threshold corresponding to 20% overhead (OH) SD-FEC [81], the CNN-DPD shows 0.8-dB sensitivity gain over the linear pre-equalization and 1.7-dB gain than the without DPD case. The 15-km SSMF transmission performance comparison is demonstrated in Figure 2.14 (c), where 1.1-dB and 3.9-dB gain are obtained, respectively. Therefore, the sensitivity is improved from -2.6 dBm to -6.5 dBm with the help of the CNN-DPD. Besides, a clear PS-64QAM-5.6 constellation is recovered at 0.2-dBm RoP as shown in the inset with the warmer color indicates a higher constellation-point density. The net spectral efficiency (SE) of the PS-64QAM-5.6 at 20% OH SD-FEC can be calculated using:

$$SE = H_{raw} - \frac{OH}{100\% + OH} \cdot \log_2 M = 5.6 - \frac{20\%}{100\% + 20\%} \cdot \log_2 64 = 4.6$$

Considering the cyclic prefix OH is 1/16, the effective baud rate of the DMT signal is derived as 14.83 GBaud. As a result, the net data rate reaches 68.2 Gb/s with -6.7 dBm sensitivity.

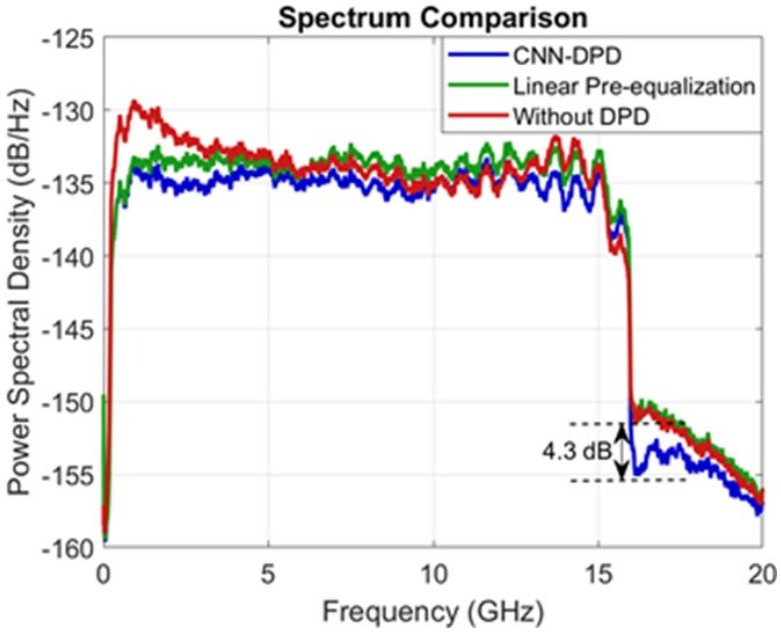


Figure 2.13 Received spectra comparison.

2.3.3 Summary

To the best of our knowledge, this is the first theoretical investigation and experimental demonstration on combining DMT and PS to exploit the channel capacity in an access network and integrating CNN-DPD to mitigate the nonlinear impairment arising from the high PAPR. The system can mitigate up to 4.3-dB nonlinearity which is measured and confirmed in the spectra comparison. Moreover, a net 68.2-Gb/s transmission over 15-km SSMF is experimentally demonstrated with 1.1-dB sensitivity gain compared to conventional pre-equalization schemes.

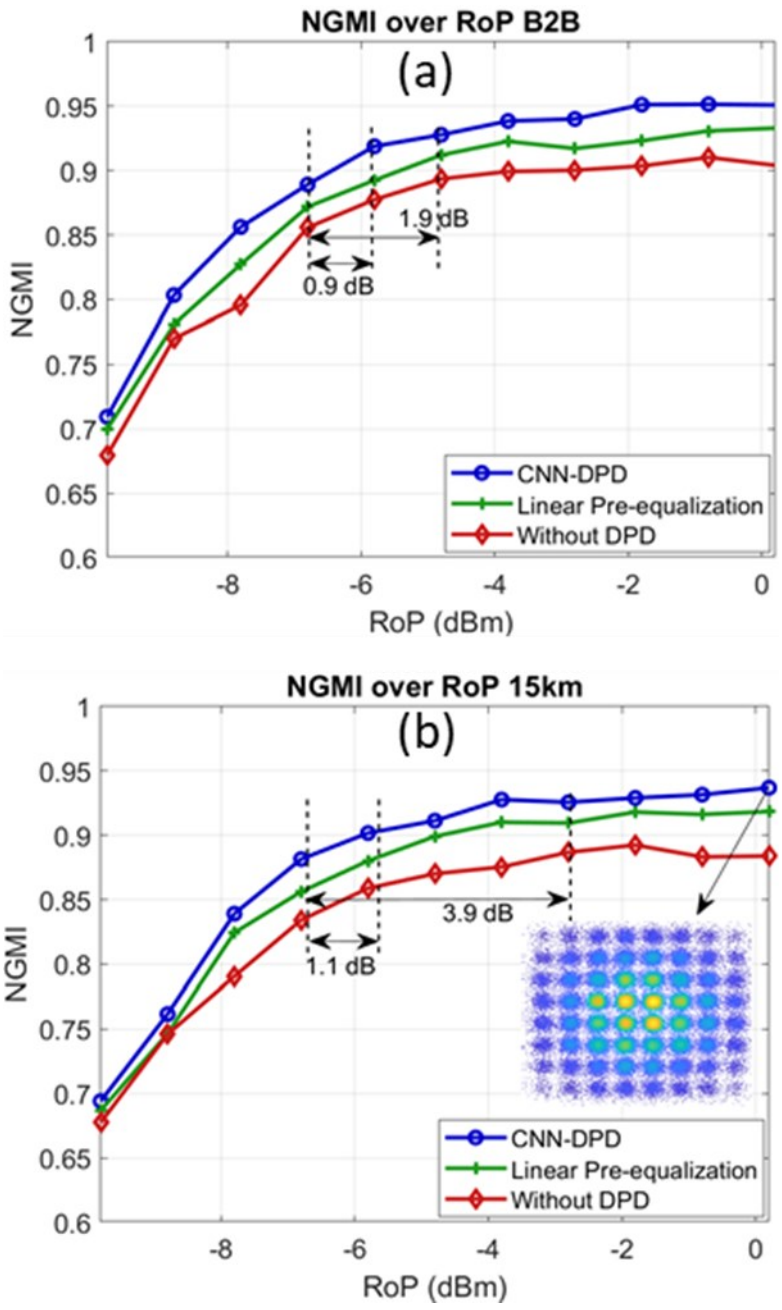


Figure 2.14 NGMI over RoP: (a) at B2B connection; (b) at 15-km SSMF transmission.

CHAPTER 3. ENHANCED PON SUPPORTING D-ROF TRANSMISSION

The PON infrastructure is an effective solution to realize economical RAN configuration and MFH transport as it allows the sharing of optical fibers and transmission equipment [74]. However, the conventional PON bandwidth will become deficient in supporting the increasing mobile network traffic in the near future. In addition, the latency in the upstream transmission of PON, due to the DBA, remains a challenge for satisfying the fronthaul latency tolerance. In this chapter, we propose a lite coherent transceiver design to improve the PON capacity and an intelligent DBA algorithm with higher bandwidth utilization to reduce the latency. Corresponding simulation and experimental results verify the effectiveness of the proposed scheme while build up the applicability of PON.

3.1 Lite Coherent Optical System

In this section, we propose and investigate a novel lite coherent receiver based PON system. The lite coherent receiver consists of only one single PD, one OLO, one analog to digital converter (ADC) and an optical coupler, which significantly simplifies the components and architecture of the conventional coherent receiver. By tuning the wavelength of the OLO, arbitrary wavelength channel selection is feasible. Also, we achieve bits stream I-Q modulation and upconversion digitally, with the optical intensity modulator converting the signal into optical domain, which cuts the high cost and eliminates the I-Q imbalance associated with the optical I-Q modulator. Based on the

proposed system built up by 10-G class electrical and optical components, we for the first time, demonstrate a symmetric 50Gb/s/λ 16-QAM transmission over 100-km standard single mode fiber (SSMF). The prototype inherits the benefits of coherent detection without the high-cost as well as facilitates the realization of promising applications like OLT-less inter-ONU communication. The contents of this section are derived from our work published in OFC [75].

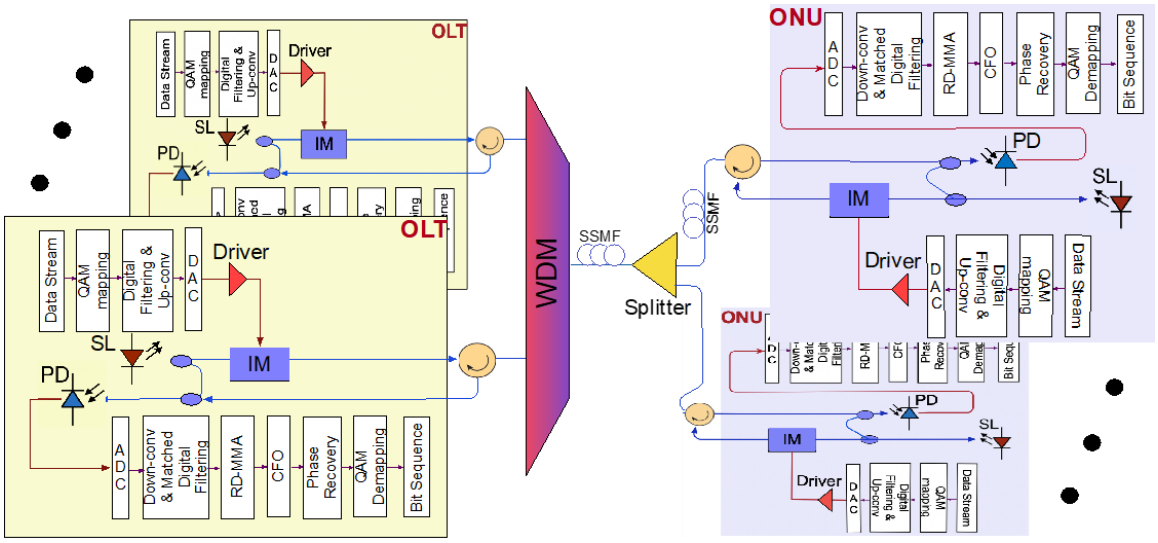


Figure 3.1 Experimental setup and framework of the prototype lite coherent access system.

3.1.1 System Setup and Principles

Figure 3.1 demonstrates the framework and experimental setup of the prototype system. An external cavity laser (ECL, PPCL100) referred as ECL1, with 13.5-dBm maximum output power is implemented in the OLT to serve as both the seed light (SL) to the optical intensity modulator for the downlink and as an OLO to the photodetector for the uplink. The OLT is connected to multiple ONUs through some optical power

splitters. In the downlink, The PRBS-15 bits sequence is first mapped to 16-QAM constellation using Gray coding. We then upconvert the digital I-Q modulated signal to an intermediate frequency (IF) and extract its real part for digital to analog conversion (DAC) using a 65 GSa/s arbitrary waveform generator (AWG, Keysight M8195A). The output of the AWG is amplified by the modulator driver (Picosecond Model 5865) with a 12-GHz 3-dB bandwidth to generate sufficient swing to drive the optical intensity modulator. We bias the JDSU 10-Gb/s intensity modulator to its null point where the best field-wise linearity locates, so that the optical carrier is suppressed while the modulation index is the highest. The modulated optical signal propagates through a 100-km SSMF. At the ONU side, an optical attenuator is implemented in front of the lite coherent receiver to simulate the path loss from power splitters. Another ECL, referred as ECL2 is implemented at the ONU serving the same functionalities as it is in the OLT. The ECL2 is set to be roughly 14-GHz away from the ECL1 and combines with the received optical signal (ROS) using a 3-dB optical coupler for lite coherent detection. After the square law detection in a PD (MITEQ-SLR) with 9.8 GHz 3-dB bandwidth, the output photocurrent IPD(t) is proportional to:

$$I_{PD}(t) \propto \{E_{OLO} \cdot \cos(\omega_{OLO} \cdot t) + E_{ROS}(t) \cdot \cos[\omega_{ROS} \cdot t + \emptyset(t)]\}^2$$

$$\propto \frac{1}{2}E_{OLO}^2 + \frac{1}{2}E_{ROS}^2(t) + 2E_{OLO} \cdot E_{ROS}(t) \cdot \cos(\omega_{OLO} \cdot t) \cdot \cos[\omega_{ROS} \cdot t + \emptyset(t)]$$

$$\propto E_{OLO} \cdot E_{ROS}(t) \cdot \{\cos[\omega_{IF} \cdot t + \emptyset(t)] + \cos[(\omega_{ROS} + \omega_{OLO}) \cdot t + \emptyset(t)]\}$$

(3.1)

Here, the E_{OLO}^2 is the average DC energy flux filtered out by the PD's internal capacitor. Since the OLO magnitude is much greater than the ROS, the $E_{ROS}^2(t)$ term can be negligible. Besides, the high frequency component at frequency $\omega_{ROS} + \omega_{OLO}$ is out of PD's bandwidth. In this case, the dominant term in photocurrent is $E_{OLO} \cdot E_{ROS}(t) \cdot cos[\omega_{IF} \cdot t + \emptyset(t)]$. ω_{IF} is the frequency difference between ω_{ROS} and ω_{OLO}. It can be seen the OLO provides huge gain and enables linear detection of ROS E-field. The ADC function onto the PD's output is performed by an 80 GSa/s oscilloscope (DSOZ254A).

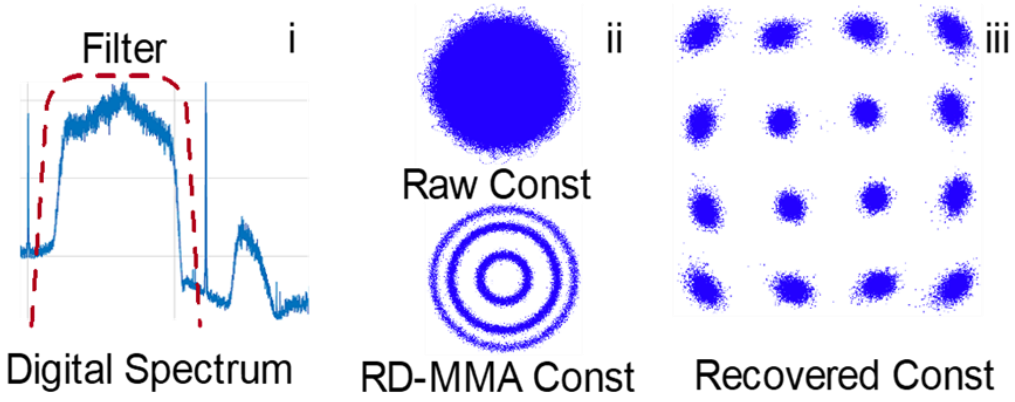


Figure 3.2 DSP process of the lite coherent receiver.

The inset (i) in Figure 3.2 shows the spectrum of the digitized samples, where the lower sideband of the signal is filtered out using the minimum-order FIR filter and downconverted to baseband. We design a radius decision based multi-modulus algorithm (RD-MMA) with 30 equalization taps. The constellations before and after the RD-MMA are demonstrated in inset (ii). The carrier frequency offset estimation and compensation are performed through finding the frequency at the maximum of $|FFT[s^4(t)]|$. Here, s(t) is the signal after RD-MMA, and FFT is fast-Fourier transformation. We apply the phase recovery algorithm afterwards. Since the phase noise frequency is relatively low in our

system, we employ 80-symbol length sliding window. In this case, any residual phase components after the fourth power of the constellation points on the centre $\sqrt{10}$ radius ring will be eliminated by each other. The recovered constellation is illustrated in inset (iii). After decoding, the BER is calculated by comparing the difference between the original and recovered bits sequences.

3.1.2 Experimental Results

Figure 3.3(a) demonstrates the BER performance over received optical power (RoP), the data rate is set to the fixed 50 Gb/s using 16-QAM modulation. The sensitivity of the prototype reaches -26.4 dBm at 7% forward-error-correction hard decision (FEC-HD) BER threshold. The link-budget reaches 40.1 dB with 13.7 dBm maximum transmitting power at the OLT. We also compare the optical back-to-back (B2B) with the 100-km transmission performance, and no obvious penalty is observed. The sensitivity over data rate result is shown in Figure 3.3(b). We choose 2.5-GBaud, 5-GBaud, 8-GBaud, 10-GBaud and 12.5-GBaud as the test points, which correspond to 10-Gb/s to 50-Gb/s using 16-QAM modulation. As observed, the higher sensitivity is obtained at lower data rate, which enables flexible data rate configuration for the ONUs requiring different link budgets. Since the OLTs and ONUs use the same optical/electrical components and architecture, the downlink and uplink transmission are symmetric with identical performance. Besides, the high capacity OLT-less inter-ONUs communications are feasible based on this prototype, providing ultra-low-latency and ultra-reliability compared with over OLT communication. In 5G era, some sets of ONUs will directly connect to wireless backhaul/fronthaul network base stations to provide mobile services within given wireless coverage area [76]. Our prototype also supports this use case as

justified in Figure 3.4. Here, 3 bands of upconverted 10-Gb/s 16-QAM signals are generated in the OLT and transmitted through 100-km SSMF. Figure 3.4 demonstrates the BER over RoP for each band after a similar signal detection and recovery processes as described above. The sensitivity for band-1, band-2 and band-3 are -28.9-dBm, -30.3-dBm and -31.8-dBm, respectively. Besides, at the ONU side, the lite coherent receiver output can be further amplified and directly emitted into air through the multi-bandpass sectors antenna to provide carrier aggregated mobile services.

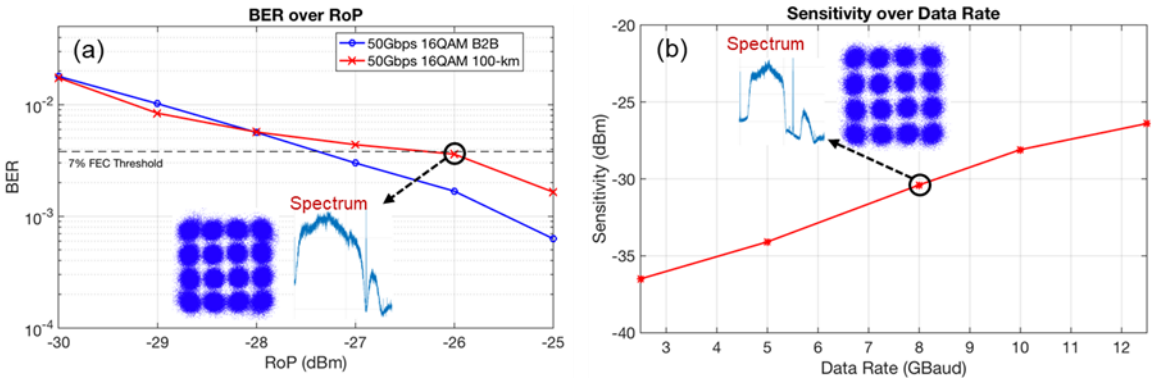


Figure 3.3 (a) BER over RoP at 50-Gb/s over 100-km SSMF, (b) sensitivity over data rates.

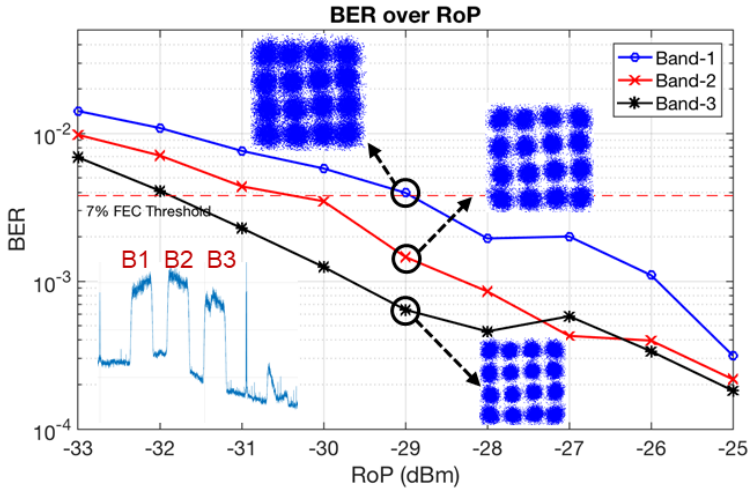


Figure 3.4 BER over RoP for multi-band transmissions over 100-km SSMF.

3.1.3 Summary

We demonstrate a pace-setting symmetric 50 Gb/s/λ 16-QAM transmission over 100-km SSMF with 40.1-dB link budget based on 10G-class electrical/optical components to support the implementation of NG-PON2. The prototype system utilizes a low-cost lite coherent receiver to achieve high receiver sensitivity, flexible wavelength channel selection, and digital dispersion/link impairment compensation. The symmetrical architecture for OLTs and ONUs enables identical downlink/uplink transmission performance. Besides, the prototype lite coherent system is capable of supporting high-capacity OLT-less inter-ONU communications. Carrier aggregation for 5G wireless services can also be realized through the prototype. Our scheme provides a promising low-cost, high-capacity solution to tackle access networking challenges of NG-PON2 and beyond.

3.2 Intelligent Dynamic Bandwidth Allocation

In this section, we propose a novel method for intelligent bandwidth allocation (IBA) in PON by using reinforcement learning (RL) for latency management. The reinforcement learning scheme consisting of three core factors, including State, Action, and Reward, is implemented to proactively update the bandwidth allocation parameters. Based on the network traffic information as the input State, the core Agent updates the maximum allocated bandwidths of the target user for latency management. We verify the capability of the proposed scheme under both fixed and dynamic traffic loads scenarios to achieve <1ms average latency. The RL agent demonstrates an efficient intelligent

mechanism to manage the latency, which provides a promising IBA solution for the next-generation access network. Our related work has been published in CLEO [77].

3.2.1 Principles of IBA

Figure 3.5 shows the principle and implementation of reinforcement learning (RL) for IBA in PON. Particularly, Figure 3.5(a) shows the detailed processing flow, and Figure 3.5(b) shows the implementation in PON environment. RL has demonstrated prominent performance in strategy selection and optimization tasks. The RL agent can obtain positive/negative reward on its executed action under a certain state through interaction with the environment. The feedbacks on state-action pairs can be saved and updated using Q-table or deep neural network, such that the agent is able to make the decision with most positive expected reward. For our demonstration, the input *State* S_t is the traffic information as the average load of traffic in ONU over the time, while the Action is the optimal maximum bandwidth $W_{max}(i)$ allocated to the i_{th} ONU, which is a series of discrete values $W(n) \in [W_1, \ldots, W_N]$. The Q-table updating, and action selection are based on State–action–reward–state–action (SARSA) algorithm, which is an on-policy temporal difference value-based RL algorithm with more conservative action to ensure reliable operations in NG-EPON. Here the *Reward* is the traffic latency of specific ONU to the target latency after action.

A two-layer implementation of proposed IBA in PON with different time scales is shown in Figure 3.5(b). In the lower layer, fast DBA is implemented in μsec to msec scale for real-time bandwidth assignments according to the actual bandwidth request from ONUs. On the higher layer, we have the IBA agent to update the DBA policy in the

scale of hundreds of msec to seconds range based on the inputs from the network. The IBA actions are implemented in a non-real-time manner, which is based on the RL algorithm described above.

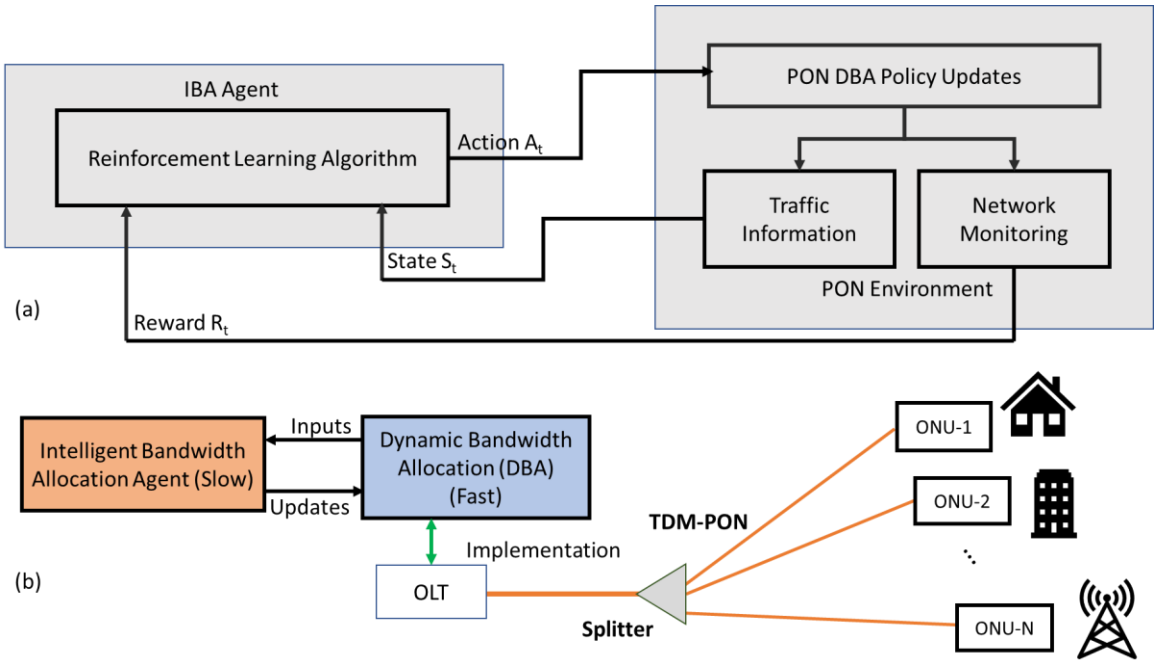


Figure 3.5 The Principle and implementation of RL for IBA in PON: (a) the principle; (b) the implementation in PON environment.

3.2.2 Physical Layer Parameters Test

To find out the key parameter settings of IBA simulation, we conduct physical layer verifications including laser on-off time, burst-clock-data-recovery (BCDR) time and link sensitivity test for 25G EPON. As such, the total effective guard interval time can be confirmed. The experimental setup is shown in Figure 3.6. We first randomly generate a 100k length bit sequence and perform pre-equalization based on estimated channel coefficients derived by least-mean-square (LMS) algorithm. The pre-equalized sequence

is then downloaded to an arbitrary waveform generator (AWG) to produce an NRZ-OOK signal running at 25 Gb/s to comply with NG-EPON standard. The signal is amplified by a modulator driver and is modulated onto a directly modulated laser (DML) for fiber transmission. The optical signal is detected by a PIN receiver with optical preamplification and is digitized by an oscilloscope for the following off-line signal decoding. A variable optical attenuator (VOA) is implemented to control the received optical power (RoP). Figure 3.7(a) shows the measured laser on-off behavior, the laser rise time is 27 ns while the laser fall time is 34 ns. Besides, to achieve a line rate of 25 Gb/s in NG-EPON, a fast synchronization and signal recovery using BCDR for the burst-mode upstream is necessary. We implement a first-order digital CDR scheme based on 2-times oversampling and nonlinear bang-bang phase detector [78] to test the BCDR speed. Follow the BER threshold in IEEE 802.3ca specification, Figure 3.7(b) demonstrates the clock phase vs time result at BER = 1e-2. The clock phase is converged at around 0.23 symbol time after 800 symbols of training. At a sampling rate of 50 GSa/s, the BCDR convergence time is derived to be 16 ns. Figure 3.7(c) is the BER over RoP measurement to calculate the link budget of the NG-EPON. The receiver sensitivity is -26 dBm at BER = 1e-2. By taking the difference between the DML output power of 5.5 dBm and the receiver sensitivity, the link budget of the system reaches 31.5 dB. Since there is no commercially available burst-mode transimpedance amplifier (BM-TIA) for burst gain setting time measurement, we calculate the total effective guard interval time based on the gain setting time of 48 ns from the reference work of 25-GBaud BM-TIA [79]. We confirm that considering the laser on-off time, BCDR time as well as the BM-TIA setting time, a total effective guard interval time of 1 μs should be safe enough for future

commercial systems. The effective guard time setting of 1 μs will be used in our following simulations.

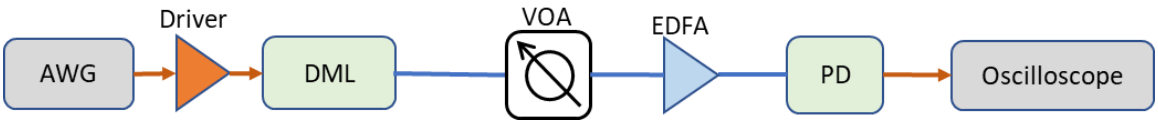


Figure 3.6 Experimental setup for guard time evaluation.

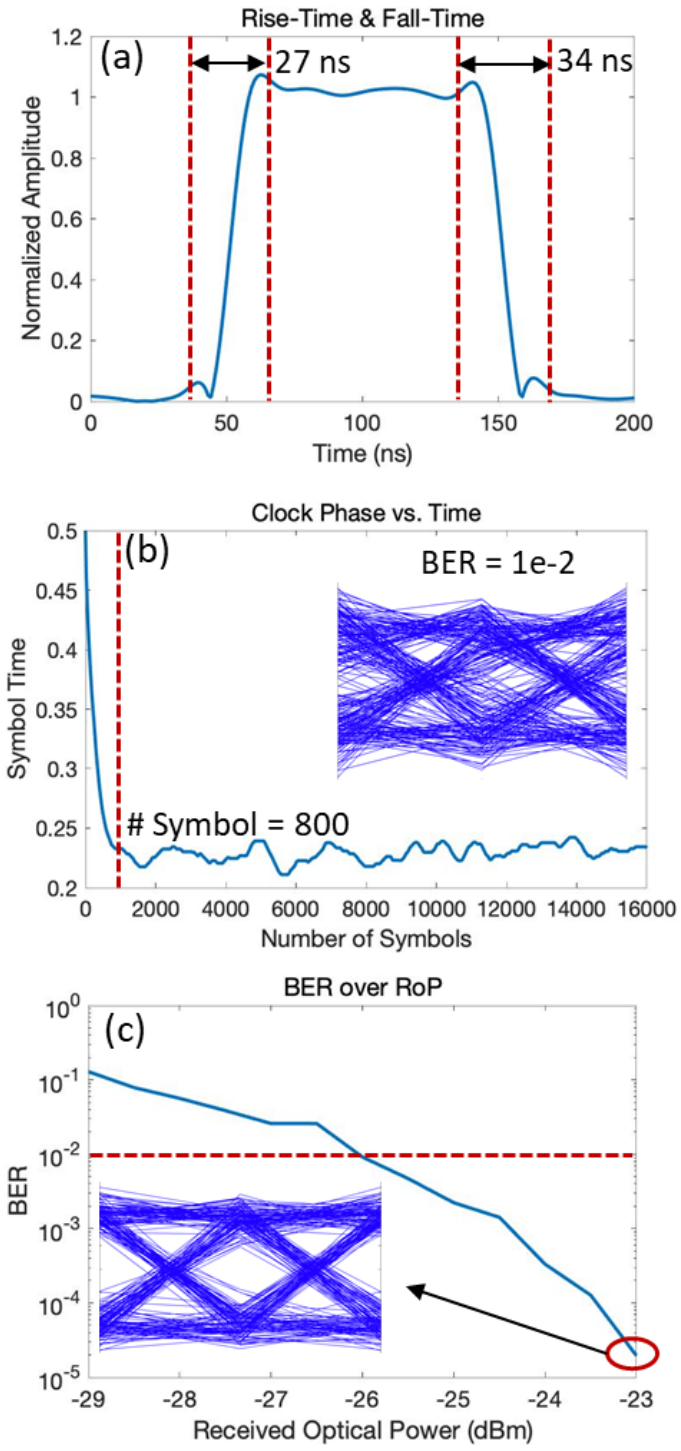
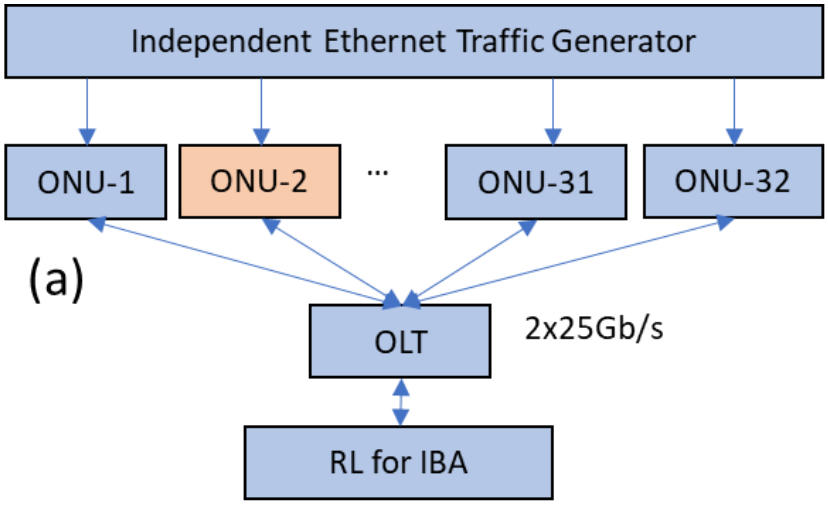


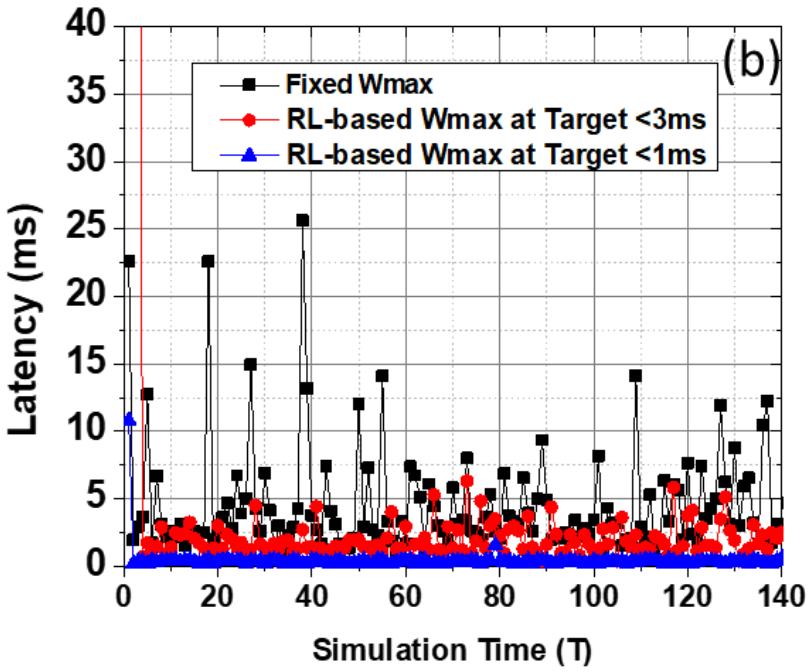
Figure 3.7 (a) Laser on-off time measurement; (b) BCDR performance at BER = 1e-2; (c) BER over received optical power (RoP) measurement.

3.2.3 Intelligent Bandwidth Allocation Simulations and Results

Figure 3.8 shows the simulation setup and results obtained. We simulated 32 ONUs in the NG-EPON system with two wavelengths each carrying 25-Gb/s data, as shown in Figure 3.8(a). We assume 1 μs for guard interval time according to our experimental verification above. As such, a total of 50-Gb/s capacity is shared by the 32 ONUs using first-fit scheme for DBA on the two wavelengths. All 32 ONUs have random RTTs within the range of 100 to 200 μs. ONU2 is the target ONU that is enabled with latency management based on RL method. All traffic is generated by an Ethernet traffic generator model that is described in [80], where self-similar traffic is generated based on the aggregation of multiple streams, each consisting of an alternating Pareto-distributed ON/OFF period [80]. The Ethernet traffics are with the packet size of 64 to 1518 bytes, and maximum traffic load for each ONU is 2 Gb/s. The default W_{max} for simulation is set at 30000 bytes. The Q-table update interval and W_{max} adjustment interval are all set as 0.8 s. Figure 3.8(b) shows the latency management results at the fixed load rate of 1.0 at 2 Gb/s. To verify the latency management capability, we set two target latency values at 3 ms and 1 ms. It is seen that in result of Figure 3.8(b) the target ONU2 follows the latency targets < 3 ms and < 1 ms with our latency management. As a comparison, we plot the latency performance under a fixed W_{max} setting to present that the variance of latency is significantly reduced by employing latency management. Figure 3.8(c) shows the Q value distribution of the Q-table after training with 1-ms target latency and different traffic loads, we can see that the peak data rate of upstream burst traffic can be as high as 5.5 Gb/s. Finally, the latency management performance with dynamic traffic loads is shown in Figure 3.8(d), with the simulation the traffic load changes based on the

trend obtained from a real user traffic behavior during a day. As a comparison, the latency performance of fixed W_{max} is also presented. The latency of fixed W_{max} (30000 bytes) can increase beyond 20-ms at high traffic load. By using RL method, with a target latency of 1 ms, the determinism and reliability of latency management are demonstrated by the simulation result that the average latency of ONU2 is below 1 ms, with a peak latency about 2 ms due to the bursty traffic.





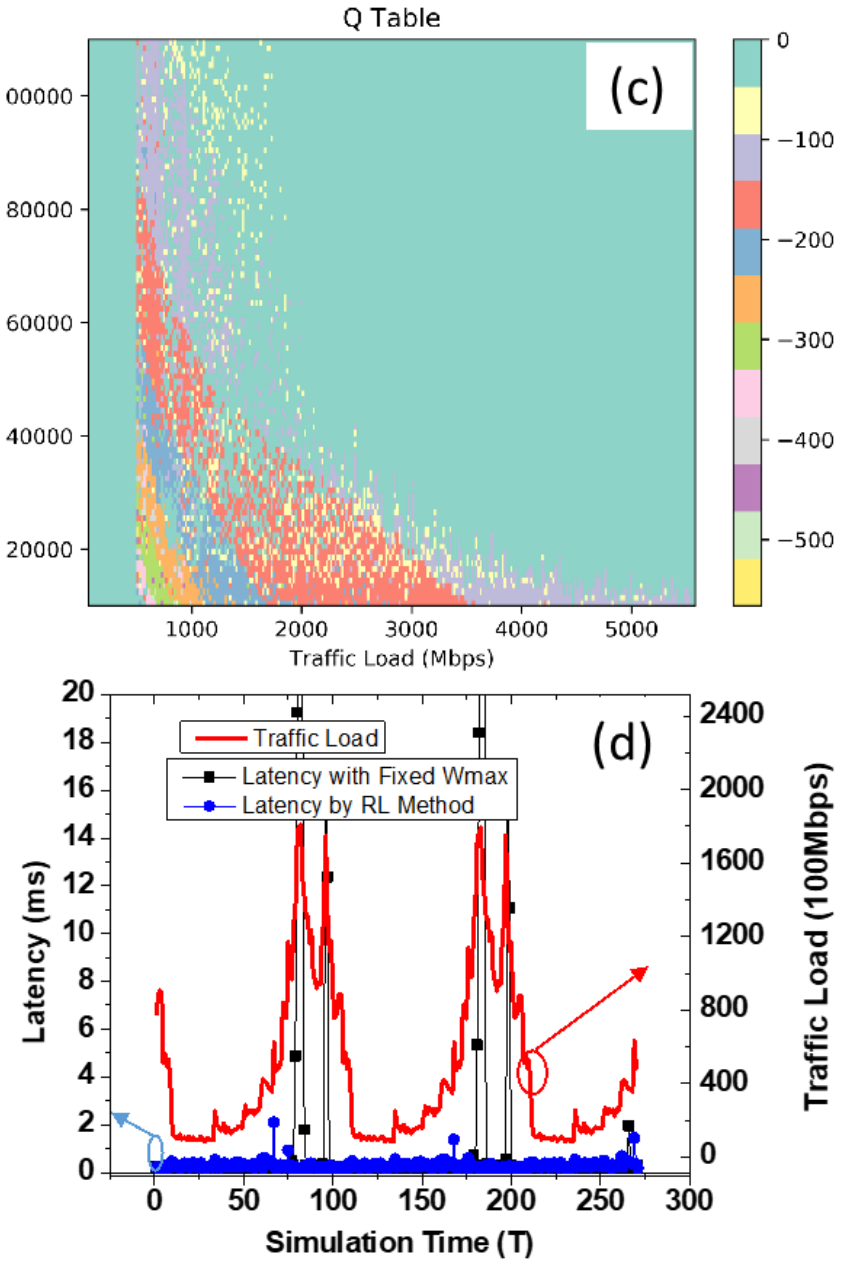


Figure 3.8 The simulation setup and results: (a) the simulation setup; (b) simulation of RL to target 3-ms latency; (c) the Q-table value obtained; (d) the latency performance with dynamic traffic load.

3.2.4 Summary

In this section, we propose a novel method for intelligent bandwidth allocation in PON by using SARSA reinforcement learning for latency management. The proposed scheme's capability to achieve < 1 ms average latency under both fixed and dynamic traffic loads scenarios is verified. An efficient intelligent mechanism to manage latency is demonstrated by the agent, offering a future-proof IBA solution for NG-EPON.

CHAPTER 4. INTERFERENCE AVOIDANCE AND CANCELLATION

In this chapter, the works on proactive interference avoidance and self-interference cancellation are demonstrated. The proof-of-concept experiments are conducted on mmWave-RoF platforms. In Section 4.1, the reinforcement learning based interference avoidance scheme provides an envision for self-organized interference management in the RAN. In Section 4.2, a simultaneous self-interference cancellation and signal-of-interest recovery scheme is realized by a dual-inputs DNN, outperforming conventional subtraction-based interference cancellers. In Section 4.3, we design a novel parallel interference cancellation scheme to facilitate NOMA decoding.

4.1 Proactive Interference Avoidance

In this section, we firstly propose and experimentally demonstrate a proactive real-time interference avoidance scheme in a mmWave-RoF system using (SARSA) RL. The scheme can proactively avoid the multi-band interference and ensure the transmission quality under both static and dynamic conditions, along with the self-evolution via agent-environment interactions. Our related work has been published in [81].

4.1.1 Operating Principles

The mobile fronthaul network consists of the CU, DU, and RRU. A DU typically connects to the plural of RRUs, providing a tactic for coordinated multi-point (CoMP) transmission and deep multi-cell resource utilization [82]. Here we focus on the transmission from the DU to RRU and ultimately to UE. Analog radio-over-fiber stands

out as the technology to connect the DU and RRU due to its higher bandwidth efficiency, lower latency, and cost. In 5G, due to the small cell architecture and the dense deployment of RRU in the urban area, there are multiple sources of unwanted interference, including the other carriers, macro cells to small cells, adjacent small cells, and intentional jammers. These interferences may collide with the desired signal. Since it is complex to acquire the full information of the external interference, the RRU is very unlikely to be able to cancel the interference. In addition, the interference cannot be filtered out using an RF filter, since it is inside the signal band. To eliminate the impairment of the interferences, the RRU can alternatively shift the signal frequency in their allowable range based on the RL algorithm to find the idle interference-free band. In RL, there are three core factors: the state, the action, and the reward. The state design needs to be informative to ensure the reinforcement learner's performance. Here the state consists of the vector of the signal's center frequency and sub-(EVMs) and is discretized when implemented. We divide the (OFDM) signal into three zones to calculate the left, center, and right sub-EVMs. To be more specific, the left sub-EVM is calculated from the first 1/3 subcarriers EVM, while the center sub-EVM and the right sub-EVM are calculated similarly, except for different zones of subcarriers. The signal center frequency is divided into 25 sections and denoted by integers 1–25, while each sub-EVM has five intervals and is denoted by integers 1–5. In other words, the value of the state represents the index of the discretization intervals. We take the discretized signal's center frequency, the left, center, and right sub-EVMs as the sub-states. All the sub-states are then combined as an overall state based on the formula below:

$$s = s0 \times 125 + s1 \times 25 + s2 \times 5 + s3$$

(4.1)

Here the s is the overall state; $s0$ is the sub-state of the center frequency; and $s1$, $s2$, and $s3$ are sub-states of the left, center, and right sub-EVM, respectively. The overall state covers all the sub-states and simplifies the RL algorithm. We define shifting the frequency as the action. There are five associated actions which are shift −20, −10, 0, 10, and 20 MHz. In this case, the agent has options for adjusting the desired signal frequency allocations, depending on different levels of aggression. Finally, the reward is defined via taking the log difference of the (BER) between the current and last state as

$$r = \log BER_{prev} - \log BER_{curr} \tag{4.2}$$

The r is the immediate reward BER_{curr} is the current BER, while BER_{prev} is the BER of the previous state. If the current BER is lower than the previous state, the reward is positive, and vice versa. Therefore, the agent can get a positive experience when the executed action reduces the BER. In addition to those core factors, we set the initial exploration rate to 1 with an exponential decay rate at 0.95 per interaction. The initial exploration rate is high, because the agent needs to get familiar with and explore more in the environment to accumulate positive and negative experience. With abundant experience acquired, the exploration rate becomes lower such that the agent can exploit its experience more often through taking beneficial actions. For a similar reason, we set the initial learning rate to 0.5, and its value is decayed by a factor of 0.95. Besides, to let the agent consider the future reward, we define a discount factor with a value of 0.8, such

that the future reward can still affect the action at the current state though alleviated. Therefore, the element value of the Q table is updated as below:

$$Q(s_n, a_n) = \alpha \cdot r + \gamma \cdot Q(s_{n+1}, a_{n+1}) + (1 - \alpha) \cdot Q(s_n, a_n) \tag{4.3}$$

where s_n, a_n represent the current state and action, respectively. α is the learning rate; γ is the discount factor; s_{n+1}, a_{n+1} are the next state and action, respectively. In this experiment, we implement SARSA learning, an on-policy temporal difference learning algorithm for value-based RL, to realize the action selection and Q table updating. SARSA learning is more suitable than the off-policy Q learning, as the latter is cliff walking with too much aggression for a real-time system [83]. Algorithm 1 for the implemented SARSA learning is designed as below:

Algorithm 1. SARSA RL Algorithm

Initialize Q table, $\epsilon = 1$, $\alpha = 0.5$, $\gamma = 0.8$
Measure current state s_0, current BER BER_0
Set current action a_0 randomly from action space
for < TP in number of TPs > **do**
    f = f + a_0
    measure and calculate next state s_1, and BER_1
    r = log BER_0 – log BER_1
    if < rand( ) < ϵ > then
        Select a_1 randomly from action space
    else
        $a_1 = argmax_a Q(s_1, a)$
        $\epsilon = \epsilon \cdot 0.95$
        $Q(s_0, a_0) = Q(s_0, a_0) + \alpha \cdot (r + \gamma \cdot Q(s_1, a_1) - Q(s_0, a_0))$
        $\alpha = \alpha \cdot 0.95$
        $a_0 = a_1$, $s_0 = s_1$, $BER_0 = BER_1$

4.1.2 Experimental Setup

Figure 4.1 shows the experimental setup to verify the proposed scheme. An external cavity laser (ECL, PPCL100) centered at 1549.8 nm is launched to the Mach–Zehnder modulator 1 (MZM1) with a 13-dBm input power, the spectrum of the ECL's output is shown as the inset i of Figure 4.2. We bias the MZM1 at null point to suppress the optical carrier. A 30-GHz RF single-tone signal is created by an RF signal generator (Anritsu 68369B) and modulates onto the MZM1. The optical spectrum after the MZM1 is shown in inset ii. To further suppress the central optical carrier and boost the 60 GHz spacing sideband carriers, we utilize an optical interleaver and an EDFA (AEDFA13-B-FC). Next, A 16-QAM OFDM signal with 52 subcarriers and 100 MHz bandwidth are generated using an arbitrary waveform generator (AWG) (M8195A) and boosted by a modulator driver (Picosecond 5865). We modulate the amplified OFDM signal onto the MZM2 biasing at the quadrature point. The optical signal then transmits through a 20 km standard single-mode fiber and is detected by a 60 GHz photodetector (PD, XPDV2020R). Due to the square law detection of the two optical carriers in the PD, a mmWave signal with 60-GHz center frequency is generated and radiated out through a horn antenna (ARH-1525-02). We use an independent channel of the AWG to create a multi-band interference with a 32-QAM OFDM waveform. Besides, another RF signal generator (E8247C) is employed to create a 15-GHz RF singletone signal which is then quadrupled to 60 GHz band using a frequency quadrupler (DBS-4060X410). Through mixing with the 60 GHz local oscillator in an RF mixer (SFB-15), the interference is upconverted to a 60 GHz band and is emitted by another horn antenna (ARH-1520). The receiving antenna (ARH-1520) located 1.5 m away from the transmitting antenna captures both the signal and interference, which are downconverted to the baseband

through an envelope detector (DET-15-RPFW0). After digitization by an oscilloscope (DSOZ254A), we use an OFDM decoder to recover the transmitted information.

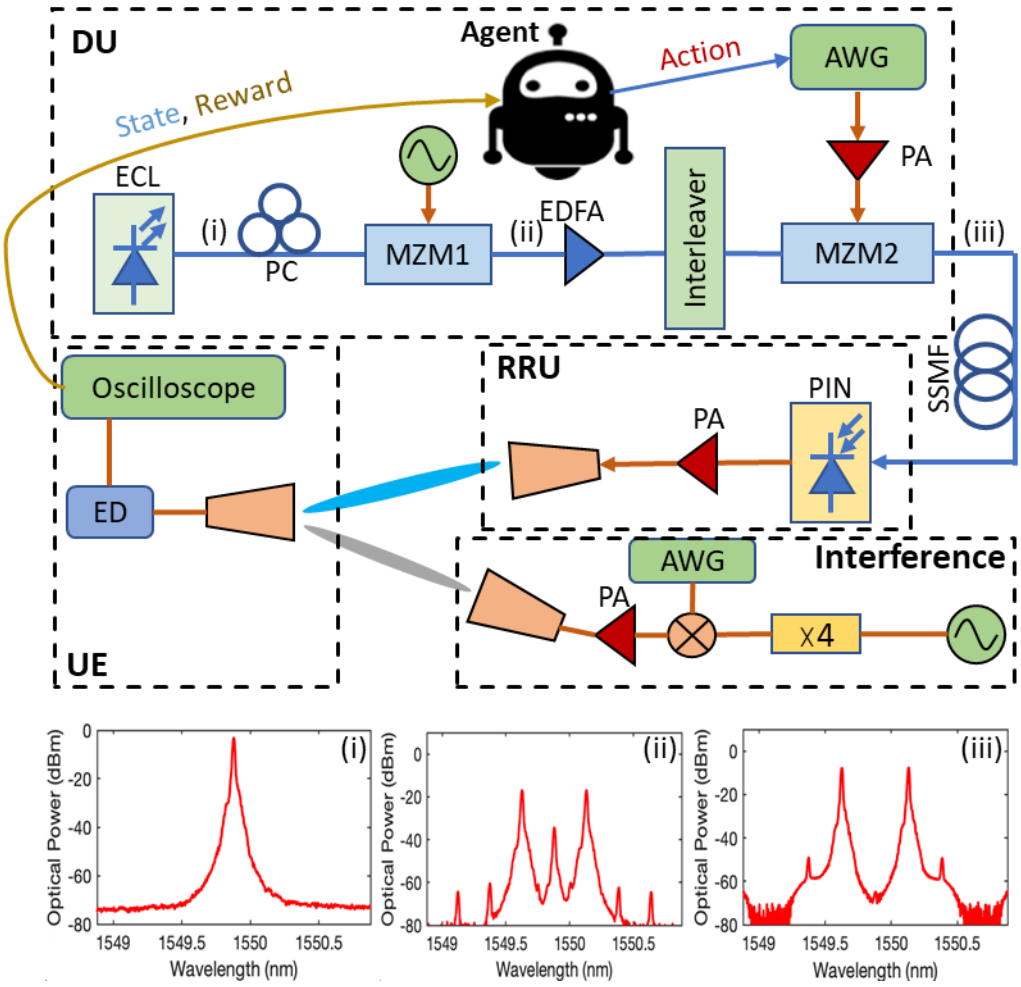


Figure 4.1 Experimental setup for proactive interference avoidance.

4.1.3 Experimental Results and Evaluation

To verify the functionality of the proposed scheme, we conduct a proof-of-concept experiment. In the first scenario, multiple interferences are involved in the receiving-end with their center frequency and bandwidth fixed. The first interference is allocated at 30 MHz with a 20-MHz bandwidth, the second interference locates at 80 MHz with a 20-

MHz bandwidth, and the third interference has a 40 MHz bandwidth and centered at 270 MHz. The transmitted signal is a 100-MHz wide 16-QAM modulated OFDM signal with 52 subcarriers. We set the initial frequency of the signal to 250 MHz, such that the third interference has strong spectrum overlapping with the signal. To mimic the realistic scenario with comparable signal and interference power, the power of the received signal at the UE is set to −12.3 dBm, while the total power of the received multi- band interference is −14 dBm. Therefore, the signal is severely degraded by the interference initially. Figure 4.2(a) demonstrates the actions taken by the agents over a time period (TP, the time for one interaction between the agent and the environment) to avoid the multiband interference. The red area represents the interfered spectrum over the TP, while the blue area corresponds to the frequency range of the desired signal. Here the center frequency of the signal is reset to 250 MHz every 100 TPs. As shown, at the first 100 TPs, the convergence time is much longer than the following episodes because of the high exploration rate at the beginning. In this condition, the agent is more intended to explore the environment. This intention helps the agent capture a broader scope in the environment to avoid the local optima. Besides, the agent accumulates a lot of related experience, such that it takes more decisive and beneficial actions in the subsequent episodes leading to a faster convergence. Figure 4.2(b) is the corresponding BER over the TP. The agent does not take any action when the BER falls below the 3.8×10^{-3} HD-FEC threshold. The change in BER clearly demonstrates the action's effect on the transmission quality. The contour map of the Q table is shown in Figure 4.3(a). The warmer color indicates a higher reward at a certain state/ action combination, and vice versa. We update the element values of the Q table according to the formula (4.3) over

the TP. The value updating amplitude of the Q table is decaying as well, since the agent becomes more familiar with the environment based on the past interactions reflected in the Q table. According to the previous miscellaneous experience, the agent acts more intelligently and makes a more beneficial decision. We plot the BER over the TP at the eighth episode, as shown in Figure 4.3(b). The agent takes decisive and requitable actions which quickly avoid the multi-band interference and recover the signal transmission quality, the actual convergence time is within a second. The corresponding signal constellations are plotted as insets in Figure 4.3(b).

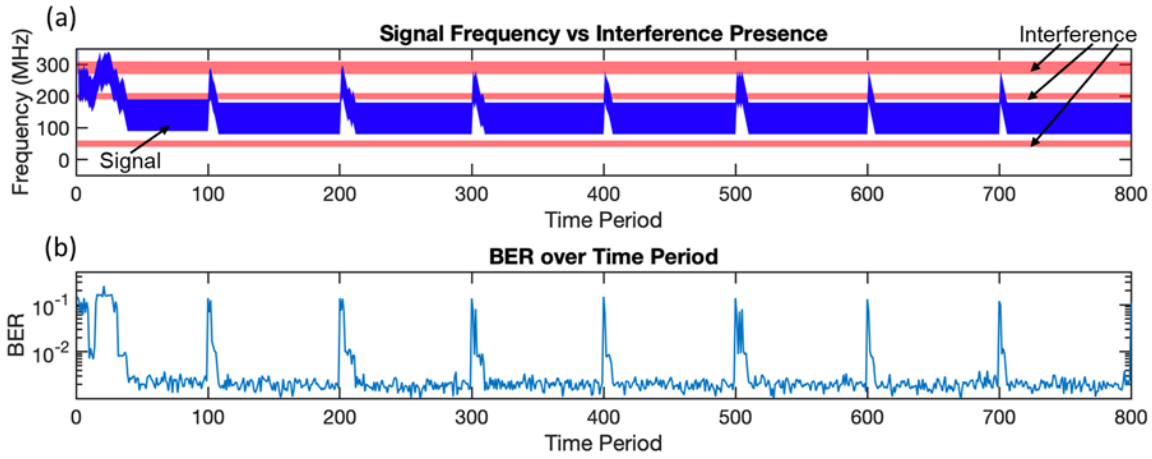


Figure 4.2 Static Interference: (a) action taken over TP; (b) BER over TP.

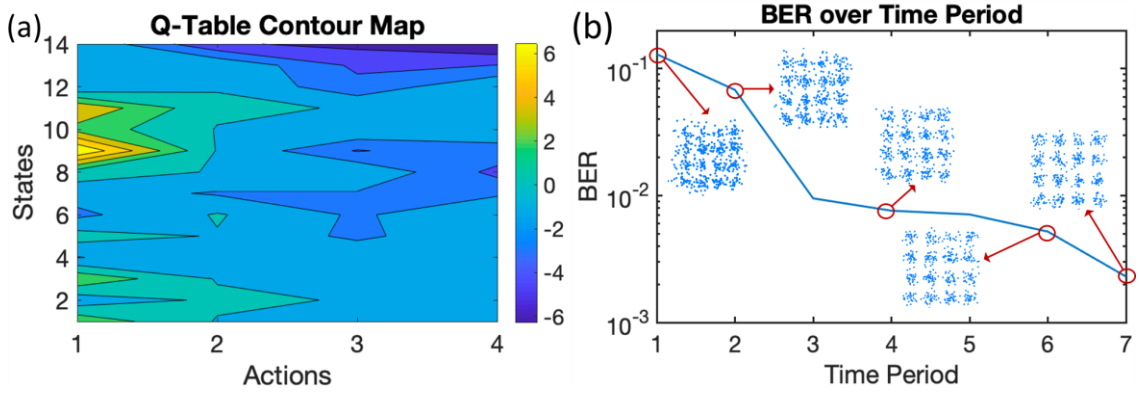


Figure 4.3 (a) Decision map shown by Q-table; (b) BER over TP at eighth episode.

In realistic scenario, the interference characteristics are not fixed. Our scheme also works in a dynamic environment. Here, to better characterize the scheme's performance, we create four types of multi-band interference, namely, Type 1, Type 2, Type 3, and Type 4. Those interference features are summarized in Table 4-1. The chronological order of the interference appearance is 1-3-2-1-2-4-3-2-1-4; each type of interference lasts 100 TPs. In this experiment, we also set the initial signal frequency to 250 MHz as the starting condition. Besides, we create an empty Q table to learn a new strategy and observe the actions on frequency shifting from TP 1 to TP 1000, as shown in Figure 4.4(a). The agent performs excellently within the first 100 TPs, as the type 1 interference is easier to avoid. Besides, though the interference is dynamic, the agent keeps interacting with the environment and learns in real time. Once the HD-FEC threshold of BER is reached, the agent stops further action, as the interference impairment is successfully eliminated. However, the states and the BER are still actively monitored; when there is an abrupt change in the interference or transmission quality, the agent takes new actions to avoid the interference until the BER reaches below the HD-FEC threshold again. Figure 4.4(b) demonstrates the corresponding BER over the TP. The result matches well with the actions taken, illustrated in Figure 4.4(a). The proposed scheme can potentially be extended to handle a more complicated realistic wireless environment efficiently. For example, its feasibility can be further improved through more diverse actions based on refined interference feature identification, while the online convergence speed would be accelerated by profiling the interference characteristics and via proper off-line training. The proposed scheme requires active communication within the control plane between the controller of the DU to take actions and the UE to calculate state and reward. This

interaction is viable, since it just requires some simplified channel state information from the UE, which can be fed back to the DU over a highly reliable scheme [84], without noticeable traffic load increasing and high latency. Meanwhile, it saves a lot of time in finding a solution to deal with the interference autonomously with the help of RL.

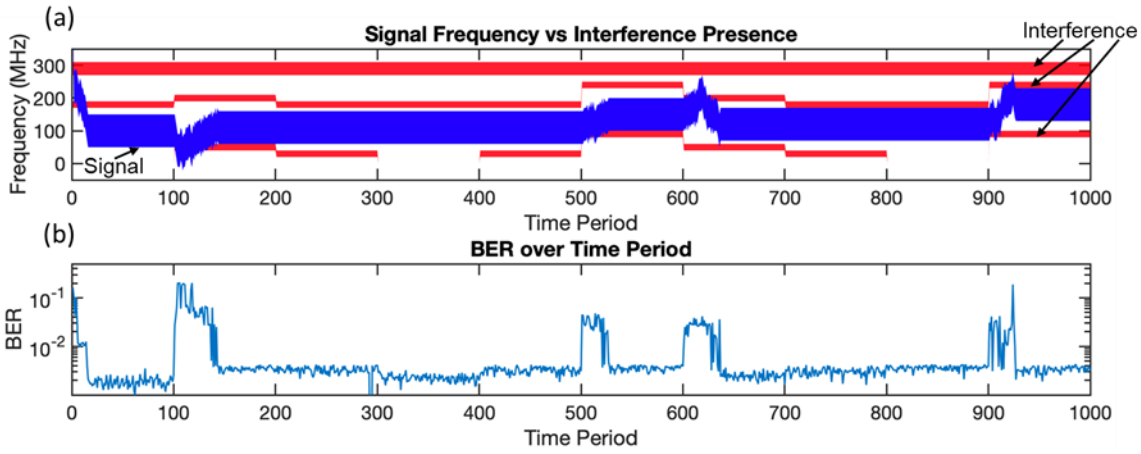


Figure 4.4 Dynamic Interference: (a) action taken over TP; (b) BER over TP.

Table 4-1 Summary of interference center frequency and bandwidth.

(MHz)	Type 1	Type 2	Type 3	Type 4
$FREQ_{B1}$	N.A.	30	50	90
$FREQ_{B2}$	180	180	200	240
$FREQ_{B3}$	290	290	290	290
BW_{B1}	N.A.	20	20	20
BW_{B2}	20	20	20	20
BW_{B3}	40	40	40	40

4.1.4 Summary

In this section, we propose and experimentally verify a proactive real-time interference avoidance scheme in a mmWave-RoF based RAN using SARSA RL. The agent can interact with the environment in real time and learn from the trial and error. The past positive and negative experiences are reflected on the element values of the Q-

table. According to the Q-table, the agent can take actions with a maximum potential reward. The proposed scheme works well for both fixed and dynamic interference scenarios. The agent becomes more intelligent and takes decisive and requitable actions when the same state appears again. In addition, the agent monitors the BER and states to detect the severity of the dynamic interference. Once there is a change degrading the signal transmission quality, the agent takes new actions on shifting the frequency to avoid the dynamic interference.

4.2 Simultaneous SI Cancellation and SOI Recovery

In this section, we propose a DI-DNN based interference canceller for simultaneous SI cancellation and SOI recovery, taking both the received signal and the known SI as dual inputs to the neural networks. We conduct a proof-of-concept experiment based on an mmWave-RoF platform to verify the performance of the proposed interference canceller systematically. To the best of our knowledge, this is the first attempt for simultaneous SI cancellation and SOI recovery considering receiver nonlinearity, along with pilot feasibility evaluation on the mmWave-band full duplex. Moreover, we expand the applicability of the DI-DNN into the full-duplex DOCSIS (Data Over Cable Service Interface Specification) and validate its effectiveness for echo cancellations. Our related works have been published in [85-87].

4.2.1 Operating Principles

This section reviews and explains the principles of neural networks based cancellers. Figure 4.5 shows a simplified wireless full-duplex system diagram to help explain cancellers' models and principles. The digital baseband signal at time instant n is denoted

as *x(n)*. The signal is converted by a digital-to-analog converter (DAC) into the analog domain, then is upconverted and amplified before being transmitted by the base station (BS) antenna. Here, we denote the signal's baseband equivalent as *x(t)*. Due to the proximity of the transmit and receive antenna, the transmitted *x(t)* causes severe SI at the local receiver. At the user equipment (UE) side, *s(t)* is generated and sent to the BS as the SOI. At the BS receiver, after amplification and downconversion, *r(t)* is obtained with received SOI *s'(t)* superimposed with the strong received SI *x'(t)*. The *r(n)* represented the digital baseband of *r(t)=f(x'(t)+s'(t))* after the analog-to-digital converter (ADC), where *f* represents a nonlinear transfer function at the BS receiver. Besides, the digital SI cancellation and DSP part do not need to be located at the BS based on different fronthaul function splits.

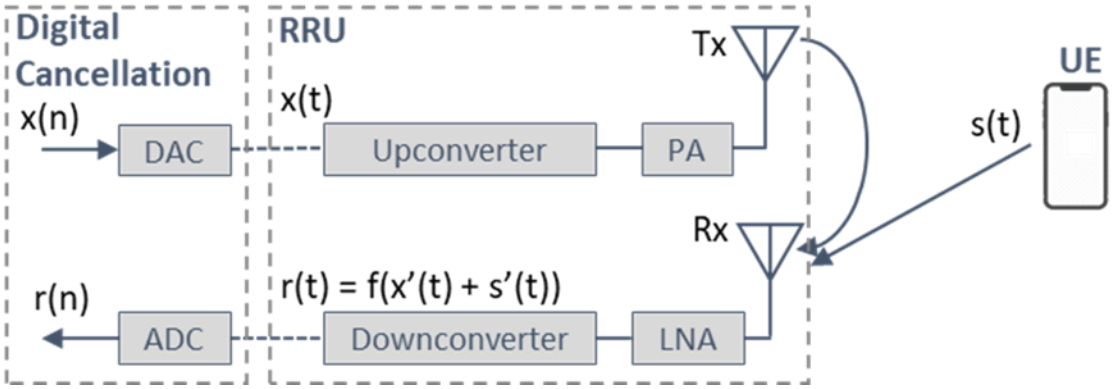


Figure 4.5 Simplified full-duplex wireless communication diagram.

To recover the SOI s(t) from the interfered received signal *r(t)*, the received SI *x'(t)* needs to be cancelled. Assume h_{SI} represents the impulse response of the SI channel, the received digital SI signal *x'(n)* can be modeled as

$$x'(n) = \sum_{l=-L}^{L} h_{SI}(l)x(n-l)$$

(4.4)

where the memory length of the channel is 2L+1. In polynomial interference cancellers, a SI cancellation signal $\hat{x}(n)$ is constructed based on polynomial estimated h_{SI} and can be derived by the Volterra series of *x(n)* as follows:

$$\hat{x}(n) = h_0 + \sum_{k=1}^{K} \sum_{l_1=-L_1}^{L_1} \dots \sum_{l_K=-L_K}^{L_K} h_k(l_1, \dots, l_K) \prod_{j=1}^{k} x(n-l_j)$$

(4.5)

Here, K is the highest nonlinearity order, L_k is the one-sided memory length, while $2L_k+1$ represents the total memory length of the k-th order nonlinearity, and h_k is the coefficient for each Volterra series element. Based on the formula (4.5), the $\hat{x}(n)$ can be estimated using the MMSE optimizer. In this case, the recovered SOI $\hat{s}(n) = r(n) - \hat{x}(n)$. The polynomial canceller assumes no nonlinear crosstalk between the SI and SOI, such that the SI can be directly subtracted from the received signal.

The polynomial canceller is normally associated with high computation complexity, especially when K is large (eg. $K = 5$). This computation complexity comes from a massive number of basis functions and the calculations of basis functions. As the multiplications dominate the hardware resources consumption, we evaluate the cancellers' complexity based on how many multiplications are required at the inference step. The complexity of the polynomial cancellers is:

$$N_{poly} = \sum_{k=1}^{K} C_k^{2L_k+k} + \sum_{k=1}^{K} (k-1) C_k^{2L_k+k} = \sum_{k=1}^{K} k C_k^{2L_k+k}$$

(4.6)

Here, the first item calculates the total number of the basis of a polynomial canceller, while the second item is the multiplications required for the basis function calculation. The linear canceller is a special case of the polynomial cancellers with $K=1$.

The neural networks have demonstrated superior performance over the Volterra series in nonlinear equalizations, such that an intuitive idea to use neural networks for interference cancellation is to estimate the h_{SI} replacing the polynomial formula. Figure 4.6 demonstrates a conventional DNN-based canceller structure. To make it simpler, we use I_N-H_1-H_2-H_3-O_N to represent the structure of the shown DNN. The processing flow is similar to polynomial cancellers but using a DNN for SI channel estimation. The total memory length for DNN input is $2L + 1$. In this case, the $\hat{x}(n) = f_{DNN}(\sum_{l=-L}^{L} x(n-l))$. The DNN canceller has a lower computation complexity than the compared polynomial cancellers based on pruned Volterra series at the inference step due to its narrow and deep architecture. The total complexity based on multiplications is:

$$N_{DNN} = I_N \times H_1 + H_D \times O_N + \sum_{i=1}^{i=D-1} H_i \times H_{i+1}$$

(4.7)

Here, I_N is the number of neurons at the DNN input layer, H_i is the neurons' number at each hidden layer, O_N is the output neuron number, and D is the depth of the DNN (the total number of the hidden layers).

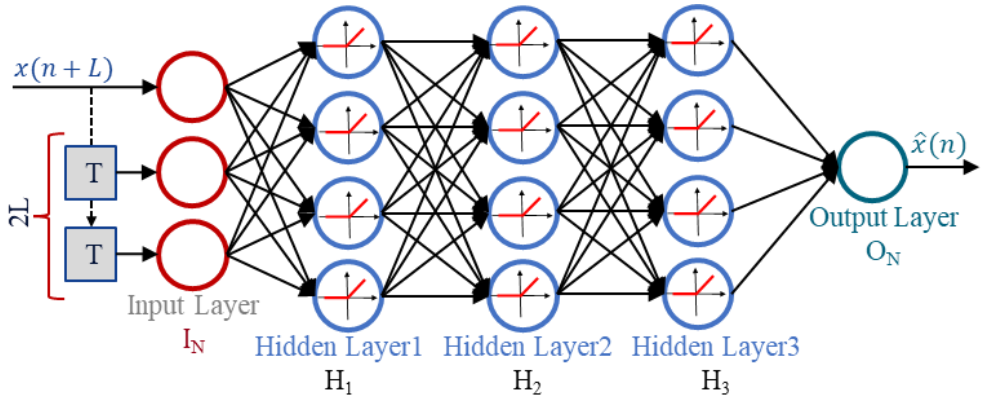


Figure 4.6 Structure and parameters of a conventional DNN canceller.

Unfortunately, the intuitive DNN canceller is still subtraction-based, which cannot mitigate the nonlinear crosstalk from the SI to the SOI. For the purpose of solving this limitation, we design a novel DI-DNN canceller. As shown in Figure 4.7, the proposed DI-DNN has two separate inputs, one is from the known current SI *x(n)* and the other input is the received signal *r(n)* with its L precursor and L postcursor samples. In this case, the SOI can be directly recovered as:

$$\hat{s}(n) = f_{DI-DNN}(x(n), \sum_{l=-L}^{l=L} r(n-l)) \tag{4.8}$$

Because of the nonlinear Relu activation function at each hidden layer, the DI-DNN can capture and mitigate the nonlinear SI to SOI crosstalk, which significantly improves the

SOI recovery performance. The DI-DNN only has a negligible complexity increase than DNN canceller due to the extra $x(n)$ as the input:

$$N_{DI-DNN} = (I_N + 1) \times H_1 + H_D \times O_N + \sum_{i=1}^{i=D-1} H_i \times H_{i+1}$$
$$= N_{DNN} + H_1 \tag{4.9}$$

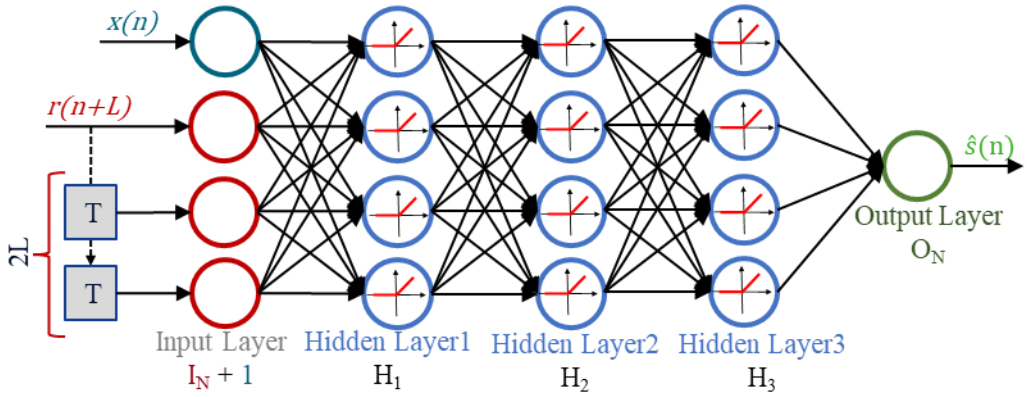


Figure 4.7 Structure and parameters of the proposed DI-DNN canceller.

The weights of the DI-DNN are optimized via supervised learning using training samples containing known inputs and corresponding outputs. The loss function is defined as the MSE and the loss with respect to the weights can be computed using backpropagation. The weight optimization is based on the Adam gradient descent algorithm, which implements momentum and adaptive learning rate during each weight update. The gradient update rule at step t is:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2$$

$$w_t = w_{t-1} - \eta \frac{m_t}{\sqrt{v_t} + \in}$$

(4.10)

The m_t is the exponential average gradients, the v_t is the exponential average of gradient squares, the g_t is the current gradient, η is the initial learning rate; β_1, β_2, and $\in$ are hyperparameters to fine-tune the Adam optimizer. This optimizer promises a fast and optimized convergence.

4.2.2 Experimental Setup

Figure 4.8 depicts the experimental setup and processing architecture. In the DU side, a C-band external cavity laser (ECL, PPCL100) is employed to generate an optical carrier to the Mach-Zehnder Modulator 1 (MZM1) biased at the null point. At the same time, a 30-GHz RF single-tone signal is modulating onto the MZM1 to obtain 60-GHz spacing optical sidebands based on optical carrier suppression (OCS). To compensate for the modulation loss from the MZM1, an EDFA (AEDFA13-B-FC) is used to boost the output optical signal. On the other hand, we generate a 16-QAM OFDM signal serving as the SI. The subcarrier spacing is $2^7\times$ 15kHz=1.92 MHz compliant with 5G conventions. The SI signal has 104 active carriers which imply a signal bandwidth of around 200 MHz, while the intermediate frequency is 300 MHz. The SI signal is digital-to-analog converted through a 16-GSa/s arbitrary waveform generator (AWG, M8195A) and boosted by a modulator driver (Picosecond 5865) to modulate the optical signal at the MZM2, which is biased at the quadrature point. The RRU is connected to the DU with a 15-km standard single-mode fiber (SSMF). After the detection by a V-band photodetector (XPDV2020R), an upconverted SI signal is generated with a 60-GHz carrier frequency.

The SI signal is amplified by a 25-dB gain power amplifier (PA1) before entering a 15-dBi horn antenna. The transmit antenna is intentionally partially positioned toward the receive antenna to mimic strong self-interference. In the UE side, a random bit-sequence is generated and modulated into QPSK symbols, following by conventional OFDM DSP to serve as the SOI. The SOI has 52 active subcarriers with 1.92-MHz subcarrier spacing. The intermediate frequency is the same as the SI signal to intensify the in-band self-interference. The SOI is upconverted to 60-GHz carrier frequency via mixing with a quadrupled 15-GHz RF single tone signal. The UE antenna is placed 2-meters away from the RRU receive antenna and the power of the SOI is controlled to verify the digital cancellers performance under different SOI to SI power ratio. At the RRU receive antenna, both the SOI and SI are captured and amplified by a low-noise amplifier (LNA). The received signal is down-converted by an envelope detector (ED, DET-15-RPFW0). Since the RRU lacks processing capability, the received signal is sent back to the DU for the interference cancellation and DSP using a directly modulated laser (DML) based IMDD link. A real-time oscilloscope (DSOZ254A) is used at the DU to sample the received signal to obtain data for experimental verification.

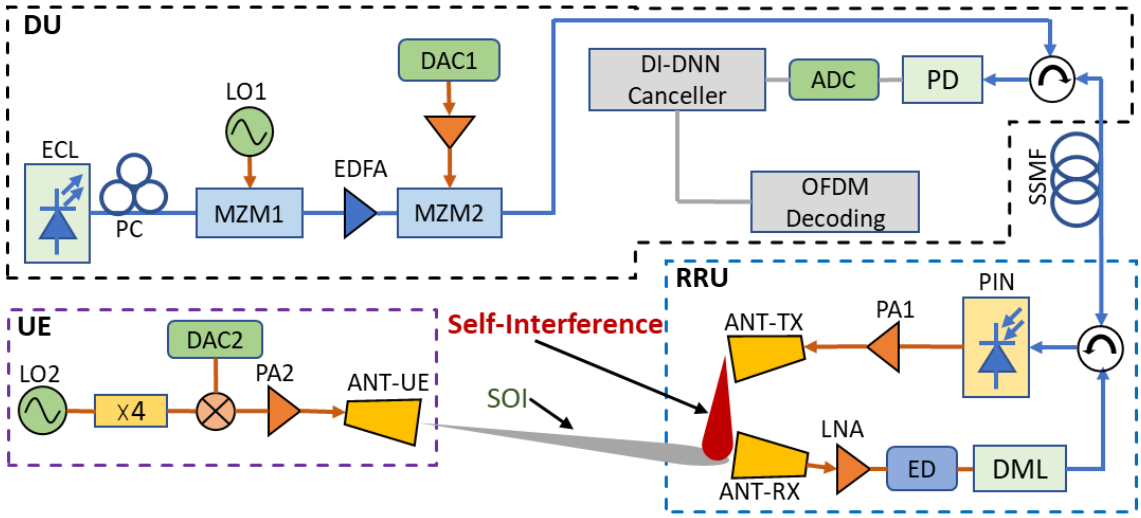


Figure 4.8 Experimental setup of full-duplex mmWave over fiber access system.

4.2.3 Experimental Results and Evaluation

Firstly, we evaluate the performance of a linear canceller. Figure 4.9(a) shows the SI cancellation coefficients of a 128-tap digital canceller derived from the correlation between pilot Tx signal and the SI. As observed, besides the dominant peak of SI, there are multiple notable SI components originated from microreflections due to impedance mismatch. Figure 4.9(b) demonstrates the spectra of the baseband OFDM SI signal before and after digital cancellation. The orange curve is the spectrum of a 1-GHz bandwidth SI signal with 520 efficient carriers over 2048 FFT size, while the blue curve is the 2-GHz bandwidth SI signal with 1040 efficient carriers. After the digital cancellation, both of the SI signals are cancelled completely down to the noise floor as shown by the purple and yellow curves, respectively. The cancellation performance is derived via calculating the difference in dB between the spectra before and after the cancellation. We experimentally demonstrate a 24.1-dB digital cancellation over 1-GHz bandwidth and a 19.5-dB cancellation over 2-GHz bandwidth in a mm-wave over fiber system, as shown in Figure

4.9(c). This prominent performance enables full-duplex operations over a wide bandwidth.

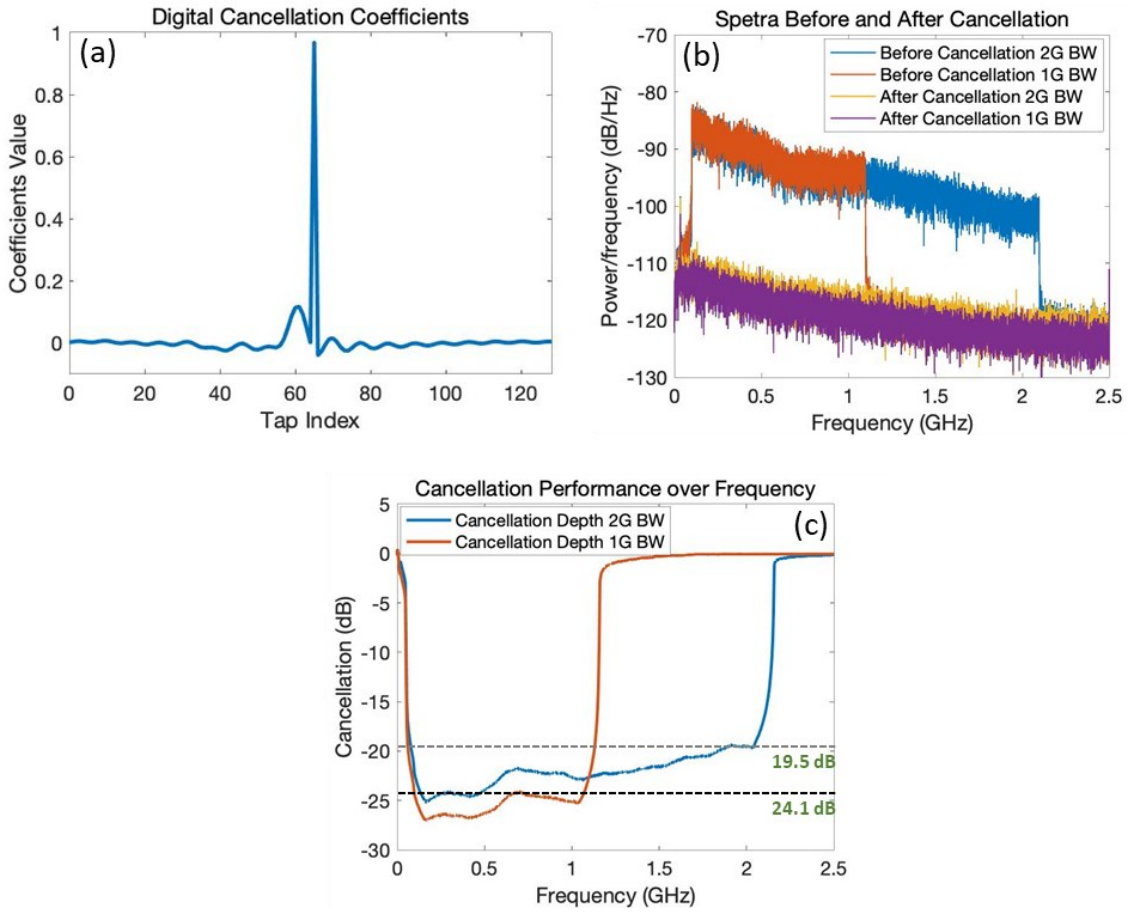


Figure 4.9 (a) Cancellation coefficients over tap index; (b) self-interference signal spectra before and after SI cancellation; (c) cancellation performance over frequency range.

Besides, we evaluate the cancellation performance of nonlinear cancellers based on Volterra series and DNNs. The proposed DI-DNN is implemented using the Tensorflow Keras framework. The DI-DNN consists of 1 input layer, 1 output layer and 3 hidden layers. The structure is (21+1)-8-4-4-1. We construct a feature dataset for the DI-DNN training and testing using the received signal samples and synchronized known SI

samples with a dimension of 195820 by 22. Each feature dataset element consists of two separate inputs. The first input is formed by the currently received signal sample with its 10 precursor and 10 postcursor samples, while the second input is obtained from the current SI sample. The target dataset for the DI-DNN is built from the synchronized SOI samples with each element be the current SOI sample. The dimension of the target dataset is 195820 by 1. The first 2/3 of the dataset is used for DI-DNN training while the remaining 1/3 is used for testing. The initial step size of the Adam optimizer is set to 0.02, the mini-batch size is 8000, and the total training epoch is 100. Figure 4.10(a) demonstrates the monitored training loss and testing loss over 100 epochs when the SOI to SI power ratio is -6 dB. The training loss and testing loss come very close to each other over the whole training epochs, the testing loss is even slightly smaller than the training loss because the training loss is measured during each epoch while testing loss is measured after each epoch. The close match of the training and testing loss verifies there is no overfitting. To connect the testing loss to SOI recovery performance, we plot the recovered SOI EVM in dB over the first 20 epochs, as the DI-DNN model already converges to a relatively low MSE loss (below 0.014) after epoch 20. As shown in Figure 4.10(b), the recovered SOI EVM follows the trend of the descent training/testing loss. The insets show how the SOI recovers from a corrupted constellation (at epoch 3) to a clear constellation (at epoch 19) along with the model's evolution.

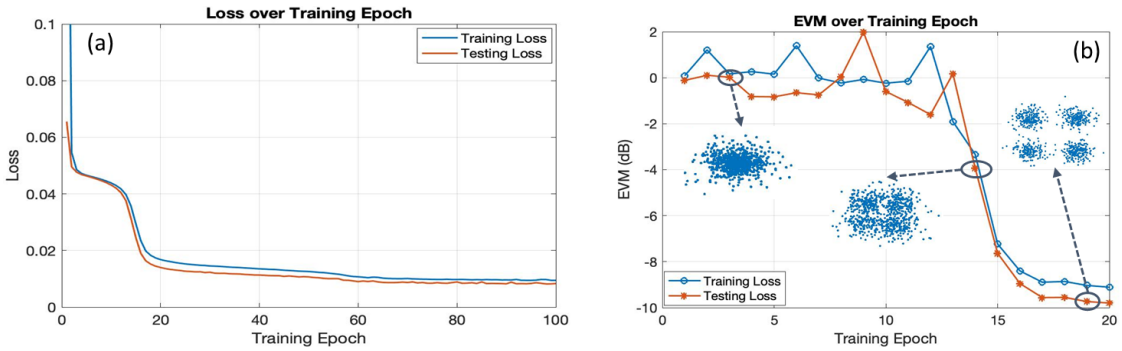


Figure 4.10 (a) Training and testing loss comparison; (b) EVM in dB and constellation of recovered SOI over the first 20 training epochs for both training and test set.

At the last training epoch (epoch 100), the MSE loss converges to 0.008, such that the DI-DNN model is well-trained for performance evaluation. Figure 4.11 demonstrates the constellation and spectrum recovery of the SOI from the received signal. After applying the DI-DNN canceller, the EVM in dB of the SOI is improved from -0.2 dB to -15.0 dB with a 14.8-dB gain achieved. The recovered constellation is clear and the SOI spectrum is fully recovered from a completely overwhelmed received spectrum. To visualize the subtraction of SI and the recovery of SOI in conventional cancellers, the received signal and the recovered signals are illustrated in the frequency domain. The hyper-parameter setting of the DNN canceller is similar to the DI-DNN for a fair comparison. Figure 4.12 shows the spectrum of the received signal, linear, nonlinear, and DNN recovered SOI, respectively. Unlike many previous papers that evaluate the cancellation performance with only the presence of SI, we measure the metrics more realistically having the SI and SOI coexist in the received signal. As shown, the linear canceller has a 17.4-dB cancellation depth, resulting in a 9.6-dB SNR of the recovered SOI. The nonlinear canceller and the DNN canceller perform similarly, both show 21.3-dB cancellation and 12.1 dB recovered SNR. The extra 3.9-dB cancellation comes from a

better estimation of the SI cancellation signal considering the transmitter-side nonlinearity.

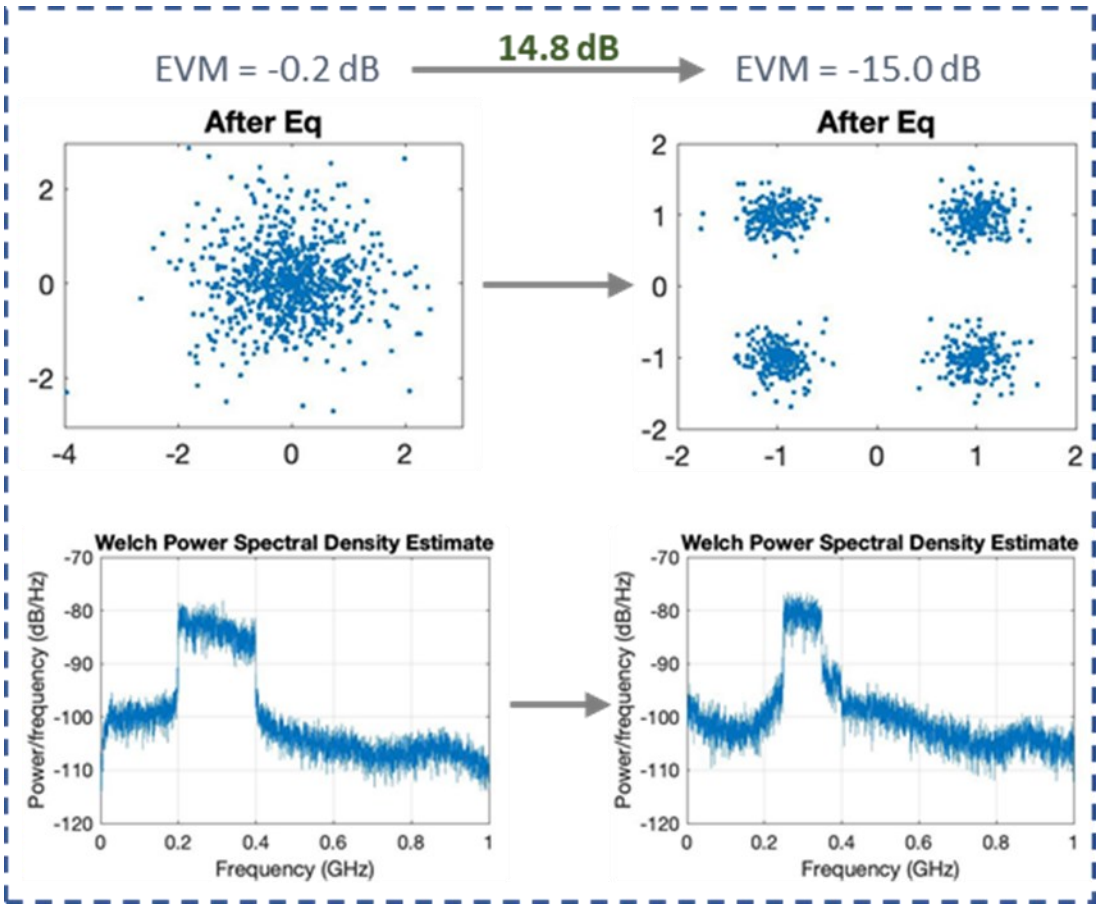


Figure 4.11 Constellation and spectrum recovery of SOI based on the DI-DNN canceller.

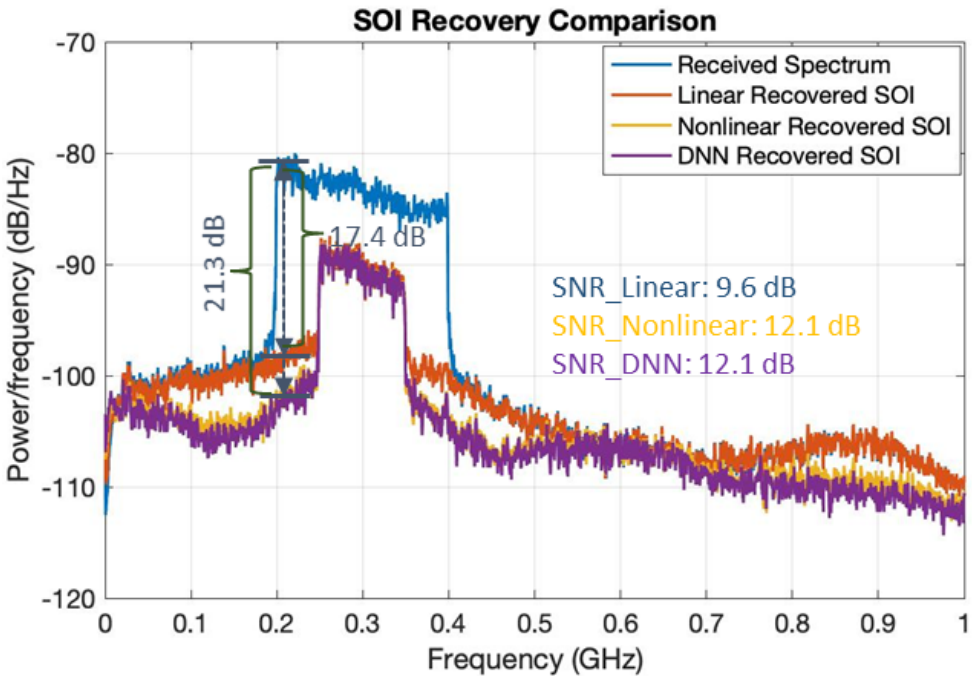


Figure 4.12 Spectra comparison between before and after cancellation of linear canceller, nonlinear canceller and DNN canceller.

The SOI to SI power ratio is an important factor to evaluate the cancellers' performance, as its value significantly affects the nonlinear crosstalk from the SI to the SOI. Besides, with a higher SOI to SI ratio, the nonlinear channel estimation on SI will be less accurate, which remarkably degrades the conventional subtraction-based cancellers performance. Due to nonlinearity at the receiver, the received SI and SOI will compete for the received power. For example, -6-dB input SOI to SI ratio corresponds to -9.3-dB received SOI to SI ratio. In the whole context, we refer the SOI to SI ratio as input SOI to SI ratio.

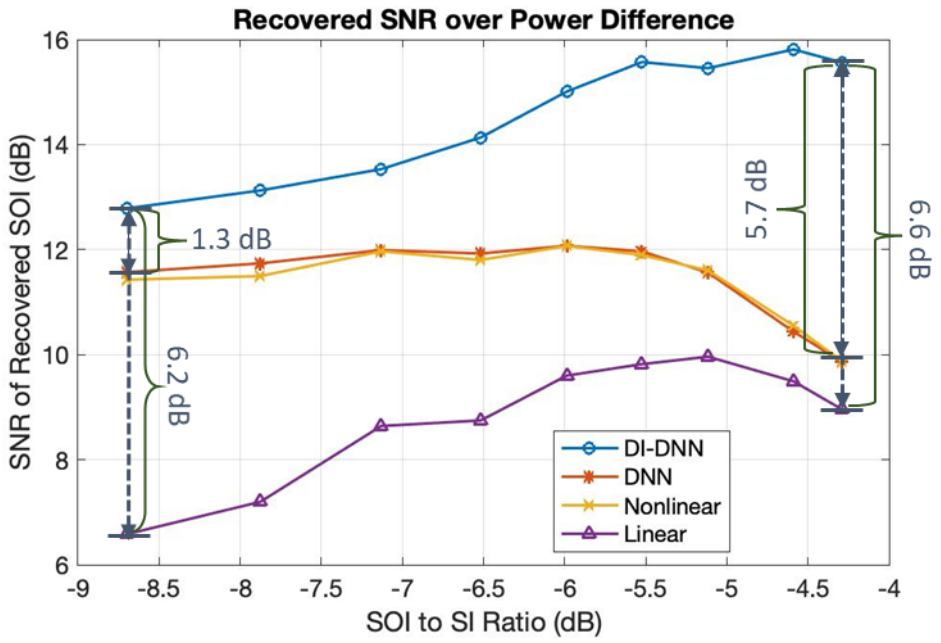


Figure 4.13 SNR of recovered SOI comparison among linear, nonlinear, DNN and DI-DNN cancellers over input SOI to SI ratio.

Figure 4.13 demonstrates the direct comparison of SOI recovery performance under different SOI to SI ratio. From the shown curves, the SNR of DI-DNN recovered SOI has 1.3-dB gain over the DNN and the nonlinear cancellers at -8.7-dB SOI to SI ratio. The gain is relatively small because the SI is dominant in the received signal, such that the nonlinear channel estimation on the SI is more accurate. Conversely, at -4.2-dB SOI to SI ratio, the DI-DNN canceller outperforms the DNN and the nonlinear cancellers by 5.7 dB. The huge margin results from a more accurate estimation of the SI cancellation signal, which is affected by the presence of the SOI, especially when the SOI is close to the SI in terms of power. This conclusion can also be justified by the declining trend of the red and yellow curve when SOI to SI ratio surpasses -6 dB, which indicates stronger nonlinear crosstalk between SOI and SI. The linear canceller performs poorly as shown by the

purple curve. The DI-DNN demonstrates more than 6-dB gain over the whole SOI to SI ratio range.

Time complexity is another essential metric to evaluate potential real-time performance. Since the multiplication operations consume most of the hardware resources, here we take the total number of multiplications as the time complexity. Due to a similar neural network structure, the DI-DNN and the DNN cancellers' complexities are close to each other. Based on formula (4.7) and (4.9), the total multiplications required at the inference step are $(21+1)\times 8+8\times 4+4\times 4+4\times 1=228$ and $21\times 8+8\times 4+4\times 4+4\times 1=220$, respectively. The Volterra nonlinear canceller, on the other hand, has a much higher complexity. Its highest nonlinear order is 3, the linear memory length is 21, both the second and the third order nonlinear memory lengths are 11. Based on formula (4.6), the total multiplications required is $21+66+286+66+286\times 2=1011$. The DI-DNN saves 77.5% required multiplications compared with the Volterra nonlinear canceller. However, if compared with simpler polynomial models like parallel Hammerstein, the complexity reduction can be smaller. Obviously, the linear canceller has the least complexity with only 21 multiplications but at the cost of insufficient performance. The complexity comparisons are summarized in Table 4-2.

Table 4-2 Complexity comparisons among digital cancellers.

Methods	Computational Complexity (Multiplications Required)
DI-DNN	228
Conventional DNN	220
Volterra Canceller	1011
Linear Canceller	21

4.2.4 Application in Echo Cancellation

Full-duplex DOCSIS has been encountering a continuous uphill battle to double the channel capacity through resource sharing of uplink and downlink channels [88]. As the downstream and upstream are delivered via the same spectrum at the same time, echoes will arise from the internal couplings or micro-reflections in a cable system. To ensure the proper operations of full-duplex DOCSIS, echo cancellation is necessary, which shares a similar process and technology as the self-interference cancellation. In this subsection, we employed the proposed DI-DNN for simultaneous echo cancellation and upstream signal recovery with experimental validation.

The experimental setup of the OFDM based full-duplex transmission system is illustrated in Figure 4.14. A 16-GSa/s arbitrary waveform generator (AWG) is employed to generate the downlink echo and the desired uplink signal. To evaluate the impairment due to the interference power from the reflected echo, a 10-dB attenuator cascaded by a 35-dB gain amplifier are applied to boost the downlink signal. Echo signal and the desired signal are combined via a power combiner before entering a 10 GSa/s real-time scope (RTS). Both echo and uplink signal are offline encoded and decoded via Matlab. The typical OFDM processing is employed. The FFT size is 2048 and the subcarrier spacing is set as 1.92 MHz. The bandwidth of echo is fixed as 990MHz and the desired uplink signal bandwidth under tested are 90, 450, and 990 MHz, respectively. The downstream transmitted power is ranging from -10 to -18 dBm before the power amplifier.

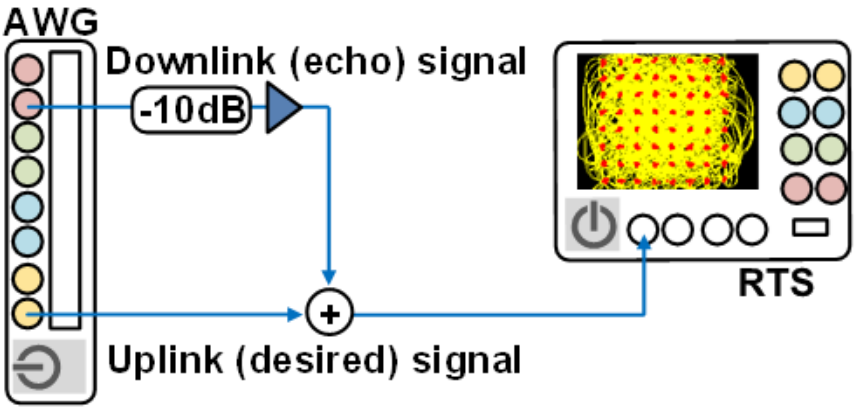


Figure 4.14 Experimental setup of the OFDM-based full-duplex cable system.

As one can note that in Figure 4.15, the received BER performance is getting worse as the bandwidth of uplink increasing. This can be understood since the output Vpp of the desired signal, i.e., uplink, is fixed by the AWG as 1 V. Thus, as the bandwidth increasing, the per subcarrier SNR is decreased. Meanwhile, the BER is getting worse when the DS power increasing because the echo impairment is also increasing for the uplink signals in this case. On the other hand, the BER performance with nonlinear canceller, i.e., Volterra equalizer, always outperforms the linear canceller, i.e., minimum mean square error (MMSE), at the cost of a much higher DSP complexity. In the case of 90-MHz US, the BER of nonlinear canceller performs irrelevant to the power of echo interference.

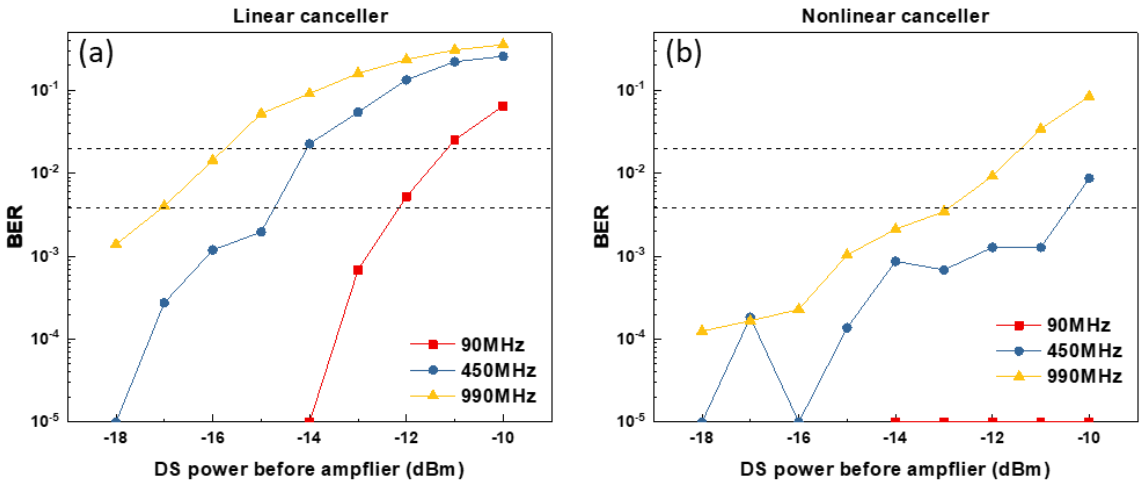


Figure 4.15 BER performance over DS power using (a) linear (b) nonlinear cancellers.

As shown in Figure 4.16, the DNN decoding performance of 90 MHz US is close to the nonlinear canceller. In this case, the DS information is not needed and thus all the RE can be fully utilized for the full-duplex operation. However, the performance of DNN drops quickly as the bandwidth increasing to 450MHz. DNN only outperforms the linear canceler in the strong interference range of DS power over -15 dBm, in such range, both DNN and linear canceler cannot achieve the FEC threshold requirement even with 20% overhand. This performance degradation may be caused by the insufficient vertical resolution of the received signal, since the peak to average power ratio (PAPR) increase proportional to the active subcarrier number. To circumvent this restriction, a straightforward method is reducing the active subcarrier number at the cost of reducing the bandwidth.

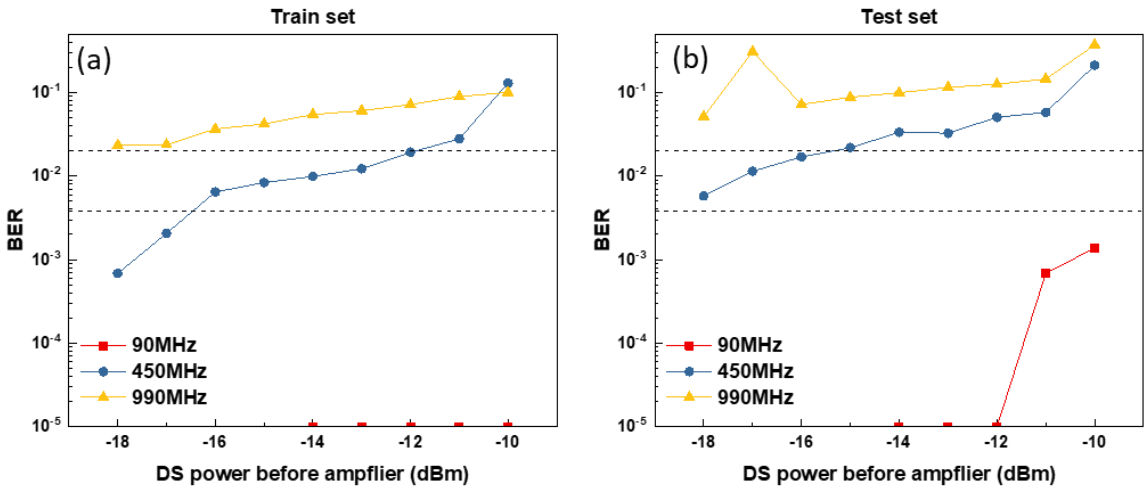


Figure 4.16 BER performance versus interference with the DNN canceller: (a) training set; (b) test set.

To get the better DNN performance and keep the transmission bandwidth, we can reduce the fast Fourier transform (FFT) size and active subcarrier number as well as increase the subcarrier spacing. Bandwidth under tested are 1GHz and 1.8GHz to comply with the DOCSIS 4.0 requirement. The active subcarrier number are 32 and 58, respectively. With the new settings, the US and echo have the same bandwidth for the FDX operation. the DNN canceller outperforms the conventional linear and non-linear DSPs in both scenarios as shown in Figure 4.17. The test set data performance is close to the training set, which implies no overfitting issue. The received performance is independent of the bandwidth increment.

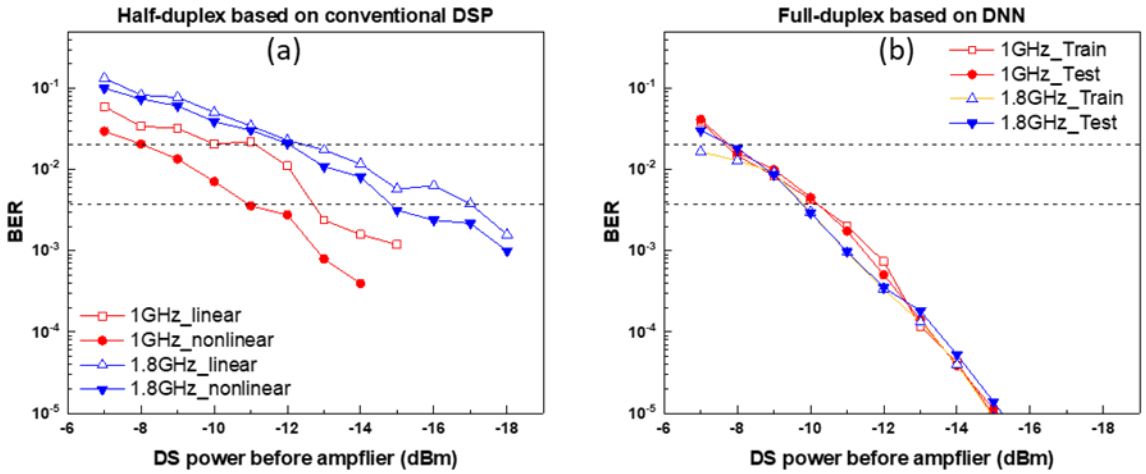


Figure 4.17 Wideband echo cancellation BER performance based on (a) conventional DSPs; (b) DNN.

4.2.5 Summary

We propose a novel DI-DNN canceller for simultaneous nonlinear SI cancellation and the SOI recovery. The DI-DNN canceller demonstrates significant SNR gain over the nonlinear and the DNN cancellers especially when the SOI to SI ratio is small. A remarkable 5.7-dB gain is realized at a ratio equals -4.3 dB. The proposed DI-DNN canceller, together with the mmWave-RoF implementation serves as a promising candidate for the 5G and beyond full-duplex wireless communication networks. Besides, we expand the applicability of the DI-DNN for echo cancellation in full-duplex DOCSIS. Better BER performance is observed at any measured DS transmission power for both 1 GHz and 1.8 GHz bandwidth cases. The results show the DNN canceller is robust for different configurations and working conditions of the coaxial cable network, offering a future-proof solution for cable operators to satisfy the exponentially growing bandwidth demand without dramatic infrastructure change.

4.3 Parallel Interference Cancellation

In this section, for the first time, we propose a Convolutional Neural Network based PIC (CNN-PIC) for decoding the NOMA signals in mmWave-RAN. The CNN-PIC effectively avoids the error propagation while mitigating the nonlinear impairments from the transmission link simultaneously. Experimental results show significant gain over the SIC in NOMA signal recovery and decoding performance. Moreover, the CNN-PIC is capable to recover the UE signals even under small power difference among the UEs, where the SIC fails to suffice. Our related work has been published in OECC [89].

4.3.1 Operating Principles

Figure 4.18(a) shows a schematic diagram of NOMA with two UEs. Here, we denote $s_1(t)$ and $s_2(t)$ as the transmitted signals from UE1 and UE2, respectively. The $s_1(t)$ and $s_2(t)$ are superimposed at the RRU, where the power ratio are determined by their transmitted power and the corresponding channel conditions. Assume that the UE1 is the user with a stronger transmitted power P_1, while the UE2 has a weaker transmitted power P_2. Then the NOMA signal received at the RRU can be represented as

$$r(t) = f\left(\sum_{i=1}^{2} h_i\sqrt{P_i}\cdot s_i(t)\right) \tag{4.11}$$

where h_i is the channel attenuation coefficient while the $f(*)$ represents a nonlinear channel transfer function. The $r(t)$ is then downconverted and digitized by an analog to digital converter (ADC) for the NOMA decoding. To avoid additional DSP and extra

delays from ranging and time synchronization in UEs, the uplink NOMA signal is asynchronous which requires waveform-level SIC in conventional solutions [90]. In the SIC scheme, the UE1 signal is decoded first using a standard OFDM decoding process by treating the UE2 signal as noise. The decoded UE1 symbol $z_1(n)$ is used to recontruct an estimated UE1 waveform $\hat{s}_1(n)$, which is then subtracted in time-domain from the digitized received signal $r(n)$ for UE2 signal recovery. In the case of low signal-to-interference-plus-noise ratio (SINR), decoding the UE1 signal could cause an inaccurate

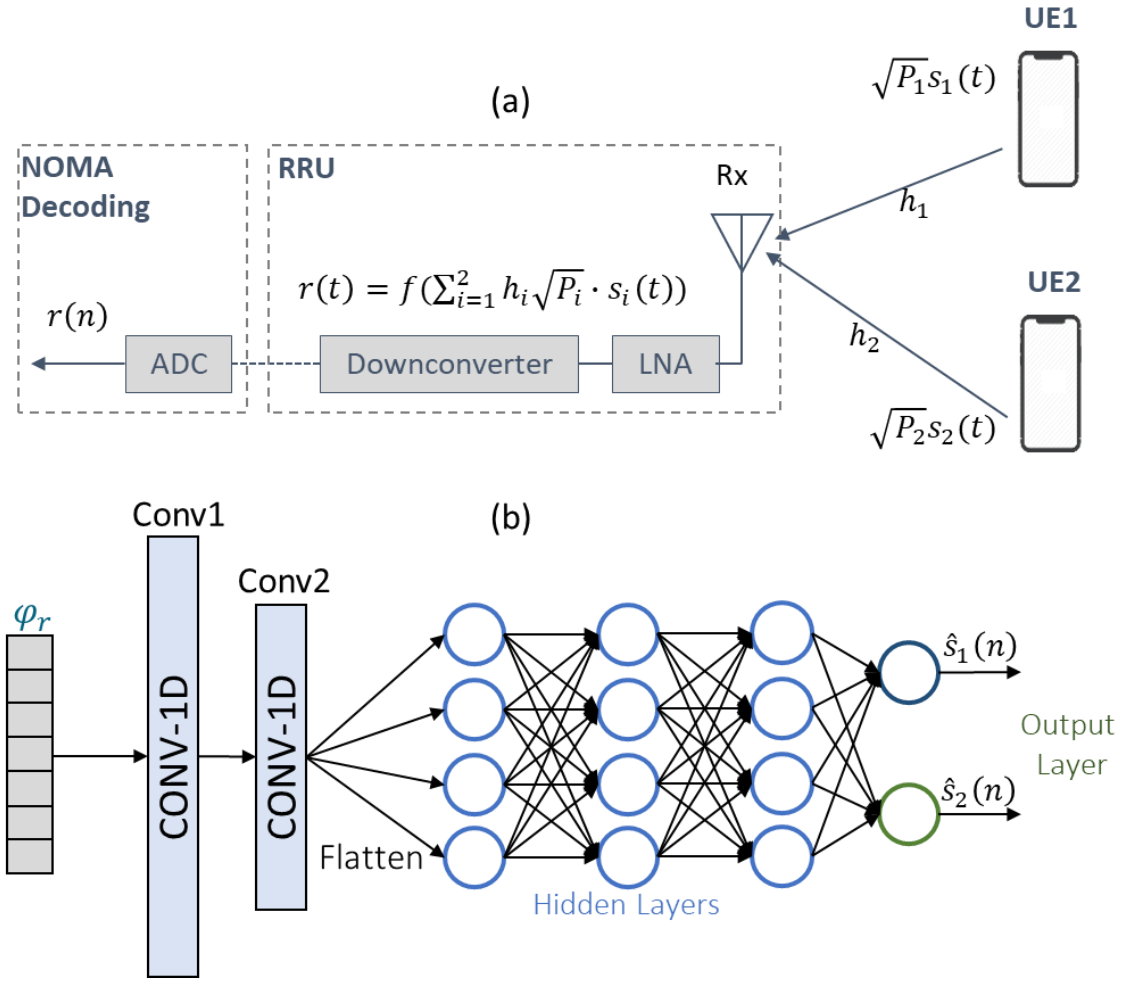


Figure 4.18 (a) NOMA illustration diagram. (b) CNN-PIC architecture.

channel estimation and high BER, which further lead to severe error propagation. Besides, the simple subtraction process of SIC neglects the receiver nonlinearity at the RRU, resulting in a strong residual interference. However, in our proposed CNN-PIC, UE1 and UE2 signals can be jointly recovered for NOMA decoding without the aforementioned drawbacks. As shown in Figure 4.18(b) of the CNN-PIC architecture, there are two neurons at the output layer, corresponding to the recovered UE1 signal $\hat{s}_1(n)$ and UE2 signal $\hat{s}_2(n)$. The φ_r is a 1-D tensor formed by the current received sample $r(n)$ and its precedent/subsequent samples. The φ_r goes through two layers of convolution process with 10 different kernel filters at each layer. Batch normalization is inserted before the Relu activation function to prevent overfitting, and the maxpooling is followed for downsampling. The output of the second convolution layer is flattened as the input to the feedforward neural network with two hidden layers. Each hidden layer has 32 neurons with Relu as the activation function. At the output layer, a joint MSE loss function is defined:

$$Loss = \alpha \cdot (\hat{s}_1(n) - s_2(n))^2 + (1 - \alpha) \cdot (\hat{s}_2(n) - s_1(n))^2 \tag{4.12}$$

The α is a weighting factor to determine the contribution of each MSE loss to the total loss for optimization.

4.3.2 Experimental Setup and Results

The proof-of-concept experimental setup is depicted in Figure 4.19. For each UE, a uniformly distributed binary sequence is generated in MATLAB, which is then 4QAM

modulated and followed by the conventional OFDM DSP. The subcarrier spacing is set to 1.92 MHz, complying to the 5G conventions. To differentiate the signals from the two UEs, 104 subcarriers are data-loaded for the UE1 signal indicating a 200-MHz bandwidth, whereas UE2 signal has a 100-MHz bandwidth with 52 active subcarriers. Each signal is digitally upconverted to a 300-MHz intermediate frequency. The OFDM waveform is generated by a 16-GSa/s arbitrary waveform generator (AWG, M8195A), while the transmitted power is tuned by the output voltage (Vp) of the corresponded AWG channel. At each UE, a mixer is used to upconvert the analog signal to 52.92-GHz carrier generated by quadrupling a 13.23-GHz RF source. The mmWave signal is further boosted by a 25-dB gain power amplifier (PA) before entering the 15-dBi transmit horn antenna. The Tx and Rx antennas are separated by 4 feets. At the RRU side, another 15-dBi receive horn antenna captures both signals from the two UEs, forming a NOMA signal. The NOMA signal is downconverted by an envelope detector (ED, DET-15-RPFW0) and converted to optical signal through a directly modulated laser (DML). After propagating through a 10-km standard single mode fiber (SSMF), the signal is detected by a photodetector (PD) at the DU, whose output connects to a 10-GSa/s oscilloscope (DSOZ254A) for digitization. The digitized samples are used in the following waveform-level SIC or CNN-PIC to compare the decoding performance.

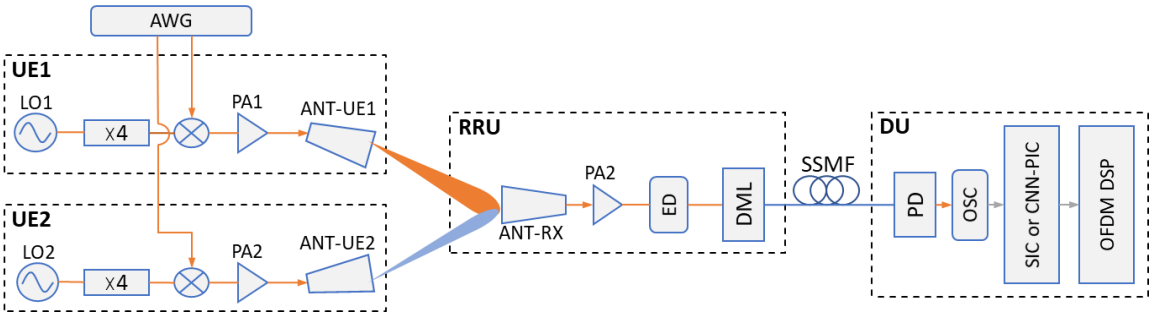


Figure 4.19 Experimental setup for NOMA PIC.

Figure 4.20(a) shows the training and testing loss of the CNN-PIC versus the training epoch. Among a total 195840 received digitized samples, the first 70% percent is separated as the training set and the remaining 30% is used as the test set. The testing loss follows a similar descending curve as the training loss without significant deviation, showing no overfitting issue. Figure 4.20(b) illustrates the waveform-level SIC in the spectral domain, where the UE2 signal is clearly recovered from the received spectrum. However, the corrupted UE2 constellation indicates a severe error propagation issue with a 92.31% EVM, which arises from the decoding error of UE1 signal with a 51.74% EVM, causing the failure of SIC. On the contrary, the proposed CNN-PIC recovers both the UE1 signal and UE2 signal in parallel, while mitigates the nonlinear channel impairments simultaneously as demonstrated by the spectra in Figure 4.20(c). Besides, clear constellations are obtained for both UE signals with a 23.02% EVM for UE1 and a 39.56% EVM for UE2, respectively.

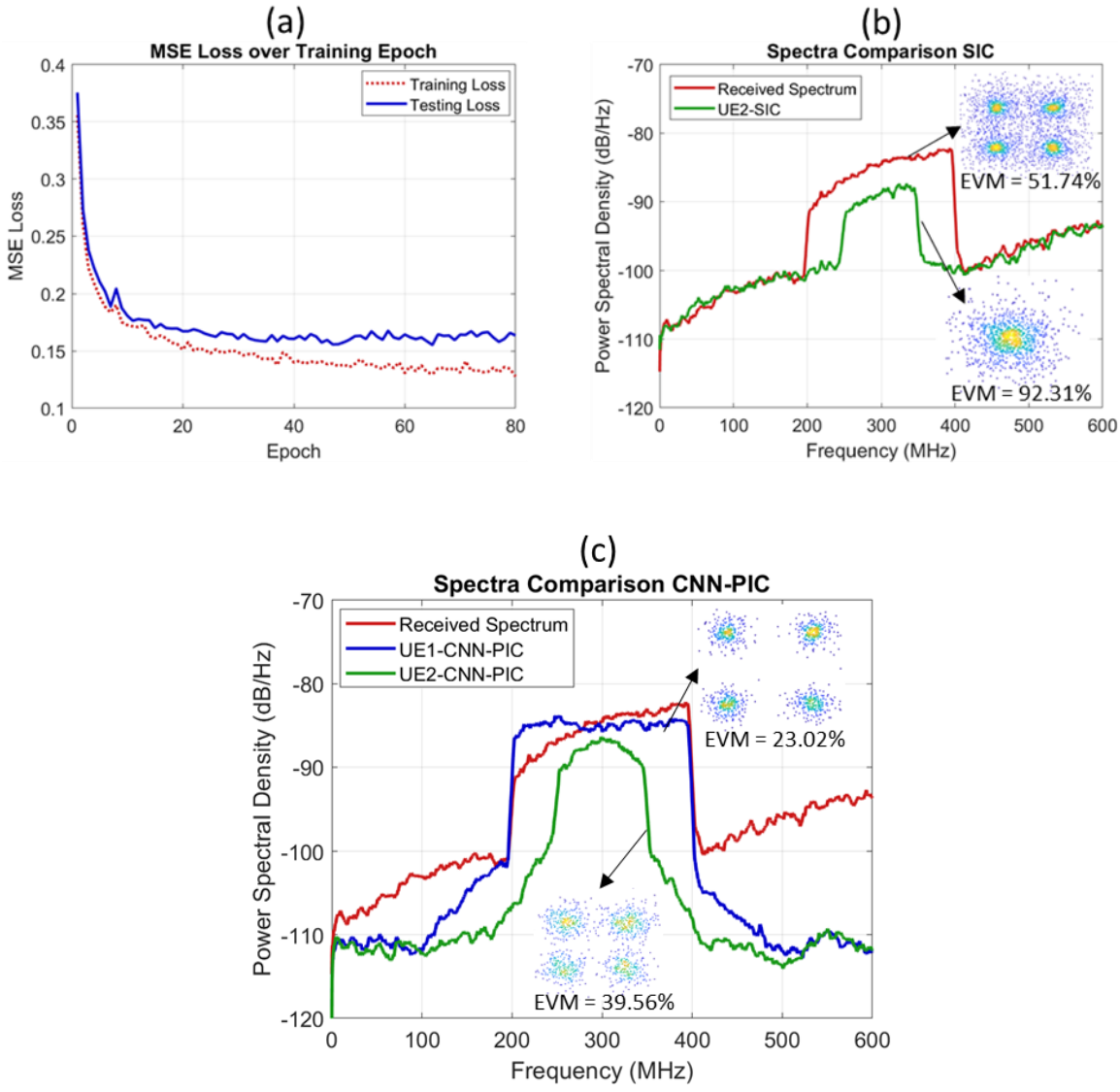


Figure 4.20 (a) Training and testing MSE loss versus training epoch. (b) Spectra comparison before and after applying SIC. (c) Spectra comparison before and after applying CNN-PIC.

To examine the performance of the CNN-PIC over different power ratio, we scan the Vp of the UE2 from 0.08 V to 0.36 V. The power density of the UE2 signal even surpasses the UE1 when UE Vp $\geq$ 0.28V. As shown in Figure 4.21(a), significant improvement of NOMA signal recovery is achieved by the CNN-PIC compared to the

SIC especially under smaller power ratio. The recovered SNR gain of UE2 over the SIC becomes smaller at a lower UE2 Vp. This phenomenon is owing to the fact that the CNN-PIC recovers the NOMA signal by find the correlations while the SIC is susceptible to severe error propagation under limited SINR. At a lower UE2 Vp, the decoding error of UE1 signal is smaller due to a higher SINR. Conversely, the correlation between the UE2 signal and the received signal becomes weaker with a lower UE2 signal power. Figure 4.21(b) demonstrates the sensitivity test in the NOMA mmWave-RoF system at 0.18-V UE2 Vp. The UE1 and UE2 using the CNN-PIC reach -5.7-dBm and -1.5-dBm sensitivity at 7% overhead HD-FEC threshold, respectively. On the other hand, the waveform-level SIC based NOMA decoding fails to reach any threshold over the received optical power (RoP) because of the dominant residual interference causing by nonlinear degradations and error propagations.

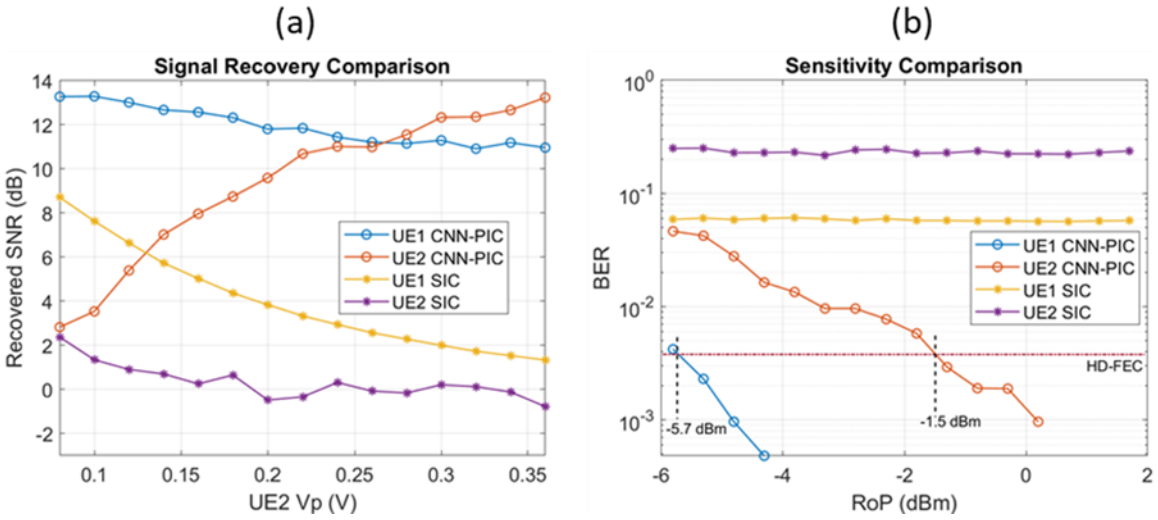


Figure 4.21 (a) NOMA signal recovery comparison under different UE2 power. (b) Sensitivity test of the link over 10-km SSMF.

4.3.3 Summary

We propose a novel CNN-PIC scheme for NOMA signal decoding and recovery in mmWave-RAN. The CNN-PIC recovers all the UE signals in parallel while mitigating the channel nonlinear impairments simultaneously. Significant improvements are observed in the recovered signal SNR due to the elimination of error propagation, which is otherwise suffered by the conventional waveform-level SIC. Moreover, -5.7dBm and -1.5dBm sensitivity are achieved in the link transmission test for UE1 and UE2, respectively. The demonstrated results justify the effectiveness of the CNN-PIC scheme, which provide a viable solution for NOMA implementation in 5G/6G mmWave-RAN.

CHAPTER 5. CONCLUSIONS

5.1 Technical Contributions

This dissertation studies the methodologies for the transmission performance optimization in fiber-wireless access network. We investigate the underlying infrastructure and improve the system design to support higher capacity and bandwidth utilization. The core innovations are centered on the implementation of machine learning techniques. We design neural networks based equalizer and pre-distorter to mitigate the nonlinear impairments arising from the overwhelmed small dynamic range of modern RANs, which are even more severe with advanced modulation formats. Considering the unprecedently complex interference in 5G and beyond era, we design a proactive interference avoidance scheme using RL and an efficient SI cancellation method using DI-DNN. Besides, a CNN based PIC solution is demonstrated to facilitate NOMA decoding. The technical contributions of the research thrusts are summarized below.

5.1.1 Advanced Signal Recovery and Modulation

The fronthaul link in a modern RAN is associated with non-negligible nonlinear impairments due to the limited dynamic range of the deployed low-cost optics/electronics components. The nonlinear degradations will be more severe when advanced modulation formats are implemented to improve the spectral efficiency. In this study, we design a DNN decoder which can effectively mitigate the nonlinear distortions. Both PAM8 and PS-PAM8 transmission are experimentally validated. Besides, to utilize the centralized

DSP resources of RANs, we invent a CNN enhanced DPD scheme that resolves the nonlinear degradations from the whole link at the transmitter side.

To enhance the multi-level signal recovery, a deep neural network (DNN) decoder which combines the function of both equalization and decoding is proposed and experimentally demonstrated for mobile fronthaul (MFH) transmission. This DNN consists of one input layer, one output layer, and several hidden layers. Adamax algorithm is implemented for finding the global minima while dropout mechanism and early stopping are utilized to avoid overfitting. The DNN accepts the received samples as the input and generates the decoded samples directly to recover the transmitted samples. Using this DNN decoder, a pace-setting data-rate transmission-distance product at 1800-Gbps·km based on the directly modulated laser (DML) with intensity-modulation direct-detection (IMDD) is obtained. Besides, the PAM8 modulation with 3-bps/Hz spectral efficiency is implemented. Due to the smaller source spectrum bandwidth compared with traditionally widely used PAM4 and OOK at a certain data-rate, the power fading limited transmission distance is extended by 1.5 and 3 times, respectively.

To mitigate the capacity crunch in RANs, we integrate the PS-PAM8 and the DNN decoder. The PS-PAM8 is power efficient and offers a higher flexibility to adapt the channel frequency response. However, the higher PAPR introduced by PS will make the nonlinear impairments even more severe. We therefore implement the DNN decoder for nonlinear compensation and improve the receiver sensitivity by 3.2-dB. Adding up the 4.1-dB gain from PS-PAM8, an 80-Gb/s over 20-km SSMF transmission performance is realized with a beyond 7.3-dB gross gain over uniform PAM modulations with linear post-equalization.

To better utilize the centralized computing resources in RANs and simplify the DSP design in the receiver. We combine the DMT and PS to exploit the channel capacity and integrate CNN-DPD to mitigate the impairments from the amplifier nonlinearity and inherent nonlinearity of electrical to optical (E/O) and optical to electrical (O/E) conversion. A 68.2-Gb/s net data-rate probabilistic shaped discrete multi-tone transmission in passive optical network with 11G-class devices. The convolutional neural network strengthens the digital predistortion performance with 1.1-dB improvement of system sensitivity over linear pre-equalization.

5.1.2 Enhanced PON supporting D-RoF Transmission

The PON infrastructure allows the sharing of optical fibers and transmission equipment to support low-cost configuration of RANs. However, the bandwidth of conventional PONs is becoming deficient to support the dramatic traffic in D-RoF transmission. Besides, the bandwidth utilization and latency are not optimized to support 5G mobile services.

To improve the capacity of PON, we utilize a low-cost lite coherent receiver to achieve high receiver sensitivity, flexible wavelength channel selection, and digital dispersion/link impairment compensation. A pace-setting symmetric 50 Gb/s/λ 16-QAM transmission over 100-km SSMF with 40.1-dB link budget based on 10G-class electrical/optical components is demonstrated. The symmetrical architecture in the OLT and ONU enables equivalent transmission rate between the downlink and uplink. Besides, the prototype lite coherent system can also support high-capacity OLT-less inter-ONU communications.

To improve the utilization of the uplink bandwidth and reduce the network latency, a novel intelligent bandwidth allocation scheme in NG-EPON using reinforcement learning is proposed and demonstrated for latency management. The proposed scheme can realize a future-proof IBA in NG-EPON, with < 1 ms average latency under both fixed and dynamic traffic loads scenarios.

5.1.3 Interference Avoidance and Cancellation

The wireless interferences are unprecedently complex in 5G and beyond RANs. Conventional static planning and human-centric interference hunting schemes are becoming inadequate. Besides, self-interference is a rising issue as full-duplex communication is introduced to magnify the channel capacity.

To avoid the interference, we propose and experimentally verify a real-time proactive interference avoidance scheme in a mmWave-RoF based RAN using reinforcement learning. The RL agent have several essential factors, including state, action, and reward. The state is defined as a discretized value from the center frequency, the left, right, and center sub-EVM of the signal. The action space has 5 options, which are −20, −10, 0, 10, and 20 MHz. The RL agent will choose the most proper action to avoid the dynamic interference. To learn from the experience, the reward is defined as the log value of BER difference between the past and the present state. The RL-based approach is an online learning algorithm, which learns in real time based on environmental feedbacks. Besides, the agent learns from past trails/errors and updates Q table accordingly, which enables it to respond intelligently when facing a similar

situation again. We verify the capability of the proposed scheme under both fixed and dynamic interference scenarios.

For the self-interference cancellation, an efficient method for simultaneous nonlinear self-interference cancellation and signal-of-interest recovery is proposed and experimentally verified in a mmWave over fiber testbed, based on a specially designed dual-input deep neural network (DI-DNN). The mmWave band is more realistic to implement full-duplex communication compared with the conventional sub-6GHz band, as it operates as a highly directional beam and shorter transmission distance resulting in a higher SOI to SI ratio. We mitigate the nonlinearity arising from the transmitters of both the SI and SOI, as well as the nonlinear crosstalk from the SI to the SOI after the detection by a nonlinear receiver. Significant SNR gain is achieved over the nonlinear and the DNN cancellers, especially when the SOI to SI ratio is small. A remarkable 5.7-dB gain is realized at a ratio equals -4.3 dB. Moreover, the expanded application of DI-DNN on echo cancellation for full-duplex DOCSIS is also successful. Higher signal quality is obtained at various DS transmission power for both 1 GHz and 1.8 GHz bandwidth cases.

To mitigate the inter-user interference in NOMA decoding and alleviate the error propagation in SIC, we propose a novel CNN-PIC scheme for NOMA signal decoding and recovery in mmWave-RAN. The CNN-PIC recovers all the UE signals in parallel while mitigates the channel nonlinear impairments simultaneously. Prominent improvements are observed in the recovered signal SNR because of the alleviation on error propagation, which is otherwise suffered by the conventional waveform-level SIC. Moreover, -5.7dBm and -1.5dBm sensitivity levels are achieved in the link transmission

test for UE1 and UE2, respectively. The demonstrated results prove the effectiveness of the CNN-PIC scheme, which provides a viable solution for NOMA implementation in 5G/6G mmWave-RAN.

5.2 Future Research Topics

5.2.1 Joint Equalization and Digital Predistortion

An important part of this dissertation is designing and implementing neural networks based digital predistortion and post-equalization to mitigate the nonlinear impairments in the fiber-wireless network. However, the dedicated DSP modules in the transmitter and receiver are not optimized for the end-to-end transmission, resulting in a deficient nonlinear compensation. Besides, they may also enhance the noise power and introduce excess power loss. Consequently, the implementation of a fiber-wireless link as an end-to-end neural network would help derive the optimal joint equalization and digital predistortion functions. This method allows the optimization for a specific metric to be done in a single deep learning process for the whole link, which deserves further investigations in the future.

5.2.2 Proactive Interference Mitigation using Deep Q Network

We have demonstrated a proactive interference avoidance scheme through shifting the signal carrier frequency. In a more complex scenario, limiting the actions to frequency shifting only might not be sufficient. Other straightforward interference mitigation actions can be power controlling, beam steering, waveform adapting, etc. Besides, the decision-making of a reinforcement learning agent needs more precise state

information to analyze the characteristics of interferences. Under those circumstances, Q-table based learning process becomes ineffective. On the other hand, the Deep Q Network (DQN) is able to accept continuous states value such that a significantly larger state-action space can be supported. The DQN only saves the DNN weights instead of a massive Q-table, which remarkably relaxes the memory requirements [91]. Therefore, the DQN based RL could be an impactful candidate to solve more complex interferences.

REFERENCES

[1] Union, I. “IMT traffic estimates for the years 2020 to 2030.” Report ITU M.2370-0 (2015).

[2] S. Liu et al., “A Multilevel Artificial Neural Network Nonlinear Equalizer for Millimeter-Wave Mobile Fronthaul Systems,” Journal of Lightwave Technology, vol. 35, no. 20, pp. 4406–4417, 2017.

[3] S. Zhou, X. Liu, F. Effenberger and J. Chao, "Mobile-PON: A high-efficiency low-latency mobile fronthaul based on functional split and TDM-PON with a unified scheduler", Proc. Opt. Fiber Commun. Conf. Exhibit. (OFC), pp. 1-3, Mar. 2017.

[4] “Transport requirement for CU&DU functional splits options.” R3-161813 (2016).

[5] M. Nazir et al, "Learning based mechanisms for interference mitigation in self-organized femtocell networks", Signals, Systems and Computers (ASILOMAR), 2010 Conference Record of the Forty Fourth Asilomar Conference on, 7-10 Nov. 2010.

[6] Ali, Samad, et al. “6G white paper on machine learning in wireless communication networks.” arXiv preprint arXiv:2004.13875 (2020).

[7] Liu X, Zeng H, Chand N, Effenberger F. Efficient mobile fronthaul via DSP-based channel aggregation. Journal of Lightwave Technology, 2016, 34(6): 1556–1564.

[8] M. Xu, F. Lu, J. Wang, L. Cheng, D. Guidotti, and G.-K. Chang, “Key technologies for next-generation digital rof mobile fronthaul with statistical data compression and multiband modulation,” Journal of Lightwave Technology, vol. 35, no. 17, pp. 3671–3679, 2017.

[9] T. Jiang, Y. Yang, and Y.-H. Song, “Companding technique for papr reduction in ofdm systems based on an exponential function,” in Global Telecommunications Conference, 2005. GLOBECOM’05. IEEE, IEEE, vol. 5, 2005, 4–pp.

[10] T. Jiang, Y. Yang, and Y.-H. Song, “Exponential companding technique for papr reduction in ofdm systems,” IEEE Transactions on broadcasting, vol. 51, no. 2, pp. 244–248, 2005.

[11] D. F. Paredes-Páliz, G. Royo, F. Aznar, C. Aldea, and S. Celma, “Radio over Fiber: An Alternative Broadband Network Technology for IoT,” *Electronics*, vol. 9, no. 11, p. 1785, Oct. 2020.

[12] K. Tanaka, H. Kao, S. Ishimura, K. Nishimura, R. Inohara, and M. Suzuki, "Bidirectional IFoF Mobile Fronthaul With DSP-based Channel Multiplexer and Demultiplexer," in *26th Optoelectronics and Communications Conference*, P. Alexander Wai, H. Tam, and C. Yu, eds., OSA Technical Digest (Optical Society of America, 2021), paper T1A.2.

[13] A. Ghosh, ‘‘The5GmmWaveradiorevolution,’’*Microw. J.*, vol.59, no.9, pp. 22–36, 2016.

[14] FCC Spectrum Frontiers (mmW) Order on 28/37/39/64-71GHz and Further Notice of Proposed Rulemaking (FNPRM): transition. https://docs.fcc.gov/public/attachments/DOC-339990A1.pdf

[15] *Tg3c channel modeling sub-committee final report*, IEEE, 2007.

[16] S. E. Alavi, M. R. K. Soltanian, I. S. Amiri, M. Khalily, A. S. M. Supa’at, and H. Ahmad, “Towards 5G: A photonic based millimeter wave signal generation for applying in 5G access fronthaul,” *Sci. Rep.*, vol. 6, p. 19891, Jan. 2016.

[17] F. Lu, M. Xu, S. Shen, Y. M. Alfadhli, H. J. Cho and G. Chang, "Demonstration of Inter-Dimensional Adaptive Diversity Combining and Repetition Coding in Converged MMW/FSO Links for 5G and beyond Mobile Fronthaul," in *Optical Fiber Communication Conference*. Optical Society of America, 2018. p. M3K.4.

[18] F. Lu, L. Cheng, M. Xu, J. Wang, S. Shen and G. Chang, "Orthogonal and Sparse Chirp Division Multiplexing for MMW Fiber-Wireless Integrated Systems," in *IEEE Photonics Technology Letters*, vol. 29, no. 16, pp. 1316-1319, 15 Aug.15, 2017.

[19] D. Zibar, R. Sambaraju, A. Caballero, J. Herrera, U. Westergren, A. Wal-ber, J. B. Jensen, J. Marti, and I. T. Monroy. "High-Capacity Wireless Signal Generation and Demodulation in 75- to 110-GHz Band Employing All-Optical OFDM," in *IEEE Photonics Technology Letters*, vol. 23, no. 12, pp. 810-812, Jun. 2011.

[20] J. Cheng, C. Xie, Y. Chen, X. Chen, M. Tang and S. Fu, "Comparison of Coherent and IMDD Transceivers for Intra Datacenter Optical Interconnects," 2019 Optical Fiber Communications Conference and Exhibition (OFC), 2019, pp. 1-3.

[21] J. -. Elbers, N. Eiselt, A. Dochhan, D. Rafique, and H. Grießer, "PAM4 vs Coherent for DCI Applications," in *Advanced Photonics 2017 (IPR, NOMA, Sensors, Networks, SPPCom, PS)*, OSA Technical Digest (online) (Optical Society of America, 2017), paper SpTh2D.1.

[22] S. Gaiarin, X. Pang, O. Ozolins, R. T. Jones, E. P. Da Silva, R. Schatz, U. Westergren, S. Popov, G. Jacobsen, and D. Zibar, "High Speed PAM-8 Optical Interconnects with Digital Equalization based on Neural Network," in Asia Communications and Photonics Conference 2016, OSA Technical Digest (online) (Optical Society of America, 2016), paper AS1C.1.

[23] R. Zhang, N. Kaneda, Y. Lefevre, A. Mahadevan, D. van Veen and V. Houtsma, "Probabilistic and Geometric Shaping for Next-Generation 100G Flexible PON," 2020 European Conference on Optical Communications (ECOC), 2020, pp. 1-4.

[24] G. Böcherer, P. Schulte and F. Steiner, "Probabilistic Shaping and Forward Error Correction for Fiber-Optic Communication Systems," in *Journal of Lightwave Technology*, vol. 37, no. 2, pp. 230-244, 15 Jan.15, 2019.

[25] Z. Qu and I. B. Djordjevic, "On the Probabilistic Shaping and Geometric Shaping in Optical Communication Systems," in *IEEE Access*, vol. 7, pp. 21454-21464, 2019.

[26] M. Schaedler, C. Bluemm, M. Kuschnerov, F. Pittalà, S. Calabrò, and S. Pachnicke, "Deep Neural Network Equalization for Optical Short Reach Communication," Applied Sciences, vol. 9, no. 21, p. 4675, Nov. 2019.

[27] W. Xu, Z. Zhong, Y. Be'ery, X. You and C. Zhang, "Joint Neural Network Equalizer and Decoder," *2018 15th International Symposium on Wireless Communication Systems (ISWCS)*, 2018, pp. 1-5.

[28] K. Burse, R. N. Yadav and S. C. Shrivastava, "Channel Equalization Using Neural Networks: A Review," in *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 40, no. 3, pp. 352-357, May 2010.

[29] S. Cass, "Taking AI to the edge: Google's TPU now comes in a maker-friendly package," in IEEE Spectrum, vol. 56, no. 5, pp. 16-17, May 2019.

[30] J. Choquette and W. Gandhi, "NVIDIA A100 GPU: Performance & Innovation for GPU Computing," in 2020 IEEE Hot Chips 32 Symposium (HCS), Palo Alto, CA, USA, 2020 pp. 1-43.

[31] D. R. Morgan, Z. Ma, J. Kim, M. G. Zierdt and J. Pastalan, "A Generalized Memory Polynomial Model for Digital Predistortion of RF Power Amplifiers," in *IEEE Transactions on Signal Processing*, vol. 54, no. 10, pp. 3852-3860, Oct. 2006.

[32] P. W. Berenguer *et al*., "Nonlinear Digital Pre-distortion of Transmitter Components," in *Journal of Lightwave Technology*, vol. 34, no. 8, pp. 1739-1745, 15 April15, 2016.

[33] S. Wang, M. Roger, J. Sarrazin and C. Lelandais-Perrault, "An Efficient Method to Study the Tradeoff Between Power Amplifier Efficiency and Digital Predistortion Complexity," in *IEEE Microwave and Wireless Components Letters*, vol. 29, no. 11, pp. 741-744, Nov. 2019.

[34] X. Wang, Y. Li, C. Yu, W. Hong and A. Zhu, "Digital Predistortion of 5G Massive MIMO Wireless Transmitters Based on Indirect Identification of Power Amplifier Behavior With OTA Tests," in *IEEE Transactions on Microwave Theory and Techniques*, vol. 68, no. 1, pp. 316-328, Jan. 2020.

[35] C. Eun and E. J. Powers, “A new Volterra predistorter based on the indirect learning architecture,” IEEE Trans. Signal Process., vol. 45, no. 1, pp. 223–227, 1997.

[36] L. Lu, Jin, P. Pang, et al., “Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators,” Nat Mach Intell 3, 218–229 (2021).

[37] A. O. María, M. Yacelga-Pinto and G. V. Arévalo, "Experimental evaluation of a NG-PON2 network," *2018 IEEE Third Ecuador Technical Chapters Meeting (ETCM)*, 2018, pp. 1-4.

[38] “40-gigabit-capable passive optical networks (NG-PON2): general requirements,” ITU-T Recommendation G.989.1, 2015, https://www.itu.int/rec/T-REC-G.989.1/en.

[39] J. Zhang, J. S. Wey, J. Yu, Z. Tu, B. Yang, W. Yang, Y. Guo, X. Huang, and Z. Ma, "Symmetrical 50-Gb/s/λ PAM-4 TDM-PON in O-band with DSP and Semiconductor Optical Amplifier Supporting PR-30 Link Loss Budget," in *Optical Fiber Communication Conference*, OSA Technical Digest (online) (Optical Society of America, 2018), paper M1B.4.

[40] C. Qin, V. Houtsma, D. Van Veen, J. Lee, H. Chow, and P. Vetter, "40 Gbps PON with 23 dB power budget using 10 Gbps optics and DMT," in *Optical Fiber Communication Conference*, OSA Technical Digest (online) (Optical Society of America, 2017), paper M3H.5.

[41] Xiang Liu and Frank Effenberger, "Emerging Optical Access Network Technologies for 5G Wireless [Invited]," J. Opt. Commun. Netw. 8, B70-B79 (2016).

[42] T. Pfeiffer, "Considerations on transport latency in passive optical networks," *45th European Conference on Optical Communication (ECOC 2019)*, 2019, pp. 1-3.

[43] F. J. Effenberger, "PON standardisation status and future prospects," 45th European Conference on Optical Communication (ECOC 2019), 2019, pp. 1-3.

[44] Hiroyuki Uzawa, Kazuaki Honda, Hirotaka Nakamura, Yukio Hirano, Ken-ichi Nakura, Seiji Kozaki, and Jun Terada, "Dynamic bandwidth allocation scheme for network-slicing-based TDM-PON toward the beyond-5G era," J. Opt. Commun. Netw. 12, A135-A143 (2020).

[45] H. Uzawa et al., "Practical Mobile-DBA Scheme Considering Data Arrival Period for 5G Mobile Fronthaul with TDM-PON," 2017 European Conference on Optical Communication (ECOC), 2017, pp. 1-3.

[46] C. Jiang, H. Zhang, Y. Ren, Z. Han, K. -C. Chen and L. Hanzo, "Machine Learning Paradigms for Next-Generation Wireless Networks," in *IEEE Wireless Communications*, vol. 24, no. 2, pp. 98-105, April 2017.

[47] Y. Zhou, L. Liu, H. Du, L. Tian, X. Wang and J. Shi, "An overview on intercell interference management in mobile cellular networks: From 2G to 5G," *2014 IEEE International Conference on Communication Systems*, 2014, pp. 217-221.

[48] M. Lichtman, R. Rao, V. Marojevic, J. Reed and R. P. Jover, "5G NR Jamming, Spoofing, and Sniffing: Threat Assessment and Mitigation," 2018 IEEE International Conference on Communications Workshops (ICC Workshops), 2018, pp. 1-6.

[49] A. Rahmati, X. He, I. Guvenc and H. Dai, "Dynamic Mobility-Aware Interference Avoidance for Aerial Base Stations in Cognitive Radio Networks," *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications*, 2019, pp. 595-603.

[50] S. Hong *et al*., "Applications of self-interference cancellation in 5G and beyond," in *IEEE Communications Magazine*, vol. 52, no. 2, pp. 114-121, February 2014.

[51] W. Nam, D. Bai, J. Lee and I. Kang, "Advanced interference management for 5G cellular networks," in *IEEE Communications Magazine*, vol. 52, no. 5, pp. 52-60, May 2014.

[52] J. X. Chen, "The Evolution of Computing: AlphaGo," in Computing in Science & Engineering, vol. 18, no. 4, pp. 4-7, July-Aug. 2016.

[53] Y. Kong, H. Zang, and X. Ma, “Improving TCP congestion control with machine intelligence,” in Proc. ACM SIGCOMM Workshop Netw. Meets AI&ML (NetAI), Budapest, Hungary, Aug. 2018, pp. 60–66.

[54] S. Hong et al., “Applications of self-interference cancellation in 5G and beyond,” IEEE Communications Magazine, vol. 52, no. 2, pp. 114–121, 2014.

[55] E. Everett, A. Sahai, and A. Sabharwal, “Passive self-interference suppression for full-duplex infrastructure nodes,” IEEE Transactions on Wireless Communications, vol. 13, no. 2, pp. 680–694, 2014.

[56] J.-I. Choi, M. Jain, J. Shalizi, and J. Mehlman, “Systems and methods for frequency independent analog self-interference cancellation,” May 19, 2015.

[57] Q. Zhou, H. Feng, G. Scott, and M. P. Fok, “Wideband co-site interference cancellation based on hybrid electrical and optical techniques,” Optics letters, vol. 39, no. 22, pp. 6537–6540, 2014.

[58] D. Korpi, L. Anttila, V. Syrjälä, and M. Valkama, “Widely linear digital self-interference cancellation in direct-conversion full-duplex transceiver,” IEEE Journal on Selected Areas in Communications, vol. 32, no. 9, pp. 1674–1687, 2014.

[59] Z. Li, Y. Xia, W. Pei, K. Wang, and D. P. Mandic, “An augmented nonlinear LMS for digital self-interference cancellation in full-duplex direct-conversion transceivers,” IEEE Transactions on Signal Processing, vol. 66, no. 15, pp. 4065–4078, 2018.

[60] A. T. Kristensen, A. Burg, and A. Balatsoukas-Stimming, “Advanced machine learning techniques for self-interference cancellation in full-duplex radios,” in 2019

53rd Asilomar Conference on Signals, Systems, and Computers, 2019, pp. 1149–1153.

[61] C. Shi, Y. Hao, Y. Liu, and S. Shao, “Digital Self-Interference Cancellation for Full Duplex Wireless Communication Based on Neural Networks,” in 2019 4th International Conference on Communication and Information Systems (ICCIS), 2019, pp. 53–57.

[62] S. Shen, Y.-W. Chen, Q. Zhou, J. Finkelstein, and G.-K. Chang, “Demonstration of Pattern Division Multiple Access With Message Passing Algorithm for Multi-Channel mmWave Uplinks via RoF Mobile Fronthaul,” J. Lightwave Technol, vol. 38, no. 21, pp. 5908–5915, 2020.

[63] G. Wang et al., “A Joint Interference Cancellation Method For Non-Orthogonal Multiple Access Uplink Signals,” Proc. IEEE ICSP, 2018, pp. 694–697.

[64] X. Li et al., "Transmission of 4×28-Gb/s PAM-4 over 160-km single mode fiber using 10G-class DML and photodiode," 2016 Optical Fiber Communications Conference and Exhibition (OFC), 2016, pp. 1-3.

[65] M. M. Maghrabi, S. Kumar, and M. H. Bakr, “Dispersion compensation of fiber optic communication system with direct detection using artificial neural networks (ANNs),” Optics Communications, vol. 409, pp. 109–116, 2018.

[66] Q. Zhou et al., "Enhanced Multi-Level Signal Recovery in Mobile Fronthaul Network Using DNN Decoder," in IEEE Photonics Technology Letters, vol. 30, no. 17, pp. 1511-1514, 1 Sept.1, 2018.

[67] R. A. Dunne and N. A. Campbell, “On the pairing of the softmax activation and cross-entropy penalty functions and the derivation of the softmax activation

function,” in Proc. 8th Aust. Conf. on the Neural Networks, Melbourne, 1997, vol. 181, p. 185.

[68] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” arXiv preprint arXiv:1412.6980, 2014.

[69] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” nature, vol. 521, no. 7553, p. 436, 2015.

[70] Q. Zhou et al., "Combining Efficient Probabilistic Shaping and Deep Neural Network to Mitigate Capacity Crunch in 5G Fronthaul," 2020 Optical Fiber Communications Conference and Exhibition (OFC), 2020, pp. 1-3.

[71] P. Schulte and G. Böcherer, "Constant Composition Distribution Matching," in IEEE Transactions on Information Theory, vol. 62, no. 1, pp. 430-434, Jan. 2016.

[72] T. Takahara et al., “Discrete multi-tone for 100 Gb/s optical access networks,” Proc. OFC, 2014, pp. M2I-1.

[73] Q. Zhou et al., "Digital Predistortion Enhancement by Convolutional Neural Network for Probabilistic Shaped Discrete Multi-Tone Signal Transmission in Passive Optical Network," 2021 Optical Fiber Communications Conference and Exhibition (OFC), 2021, pp. 1-3.

[74] J. Kani, J. Terada, K. Suzuki and A. Otaka, "Solutions for Future Mobile Fronthaul and Access-Network Convergence," in Journal of Lightwave Technology, vol. 35, no. 3, pp. 527-534, 1 Feb.1, 2017.

[75] Q. Zhou, J. He, S. Shen, R. Zhang, S. Yao, Y. Alfadhli, Y. Chen, and G. Chang, "Symmetric Long-Reach 16-QAM Transmission using Lite Coherent Receiver for Next-Generation Optical Access Network," in *Optical Fiber Communication*

Conference (OFC) 2019, OSA Technical Digest (Optical Society of America, 2019), paper Th2A.29.

[76] H. Beyranvand, M. Lévesque, M. Maier, J. A. Salehi, C. Verikoukis and D. Tipper, "Toward 5G: FiWi Enhanced LTE-A HetNets With Reliable Low-Latency Fiber Backhaul Sharing and WiFi Offloading," in IEEE/ACM Transactions on Networking, vol. 25, no. 2, pp. 690-707, April 2017.

[77] Q. Zhou, J. Zhu, J. Zhang, Z. Jia, B. Huberman and G. -K. Chang, "Intelligent Bandwidth Allocation for Latency Management in NG-EPON using Reinforcement Learning Method," 2020 Conference on Lasers and Electro-Optics (CLEO), 2020, pp. 1-2.

[78] J. L. Sonntag and J. Stonick, "A Digital Clock and Data Recovery Architecture for Multi-Gigabit/s Binary Links," in *IEEE Journal of Solid-State Circuits*, vol. 41, no. 8, pp. 1867-1875, Aug. 2006.

[79] G. Coudyzer, P. Ossieur, L. Breyne, M. Matters, J. Bauwelinck and X. Yin, "A 50 Gbit/s PAM-4 Linear Burst-Mode Transimpedance Amplifier," in IEEE Photonics Technology Letters, vol. 31, no. 12, pp. 951-954, 15 June15, 2019.

[80] G. Kramer, B. Mukherjee and G. Pesavento, "IPACT a dynamic protocol for an Ethernet PON (EPON)," in IEEE Communications Magazine, vol. 40, no. 2, pp. 74-80, Feb. 2002.

[81] Qi Zhou, You-Wei Chen, Shuyi Shen, Yiming Kong, Mu Xu, Junwen Zhang, and Gee-Kung Chang, "Proactive real-time interference avoidance in a 5G millimeter-wave over fiber mobile fronthaul using SARSA reinforcement learning," Opt. Lett. 44, 4347-4350 (2019)

[82] P. Chanclou, L. Anet Neto, K. Grzybowski, Z. Tayq, F. Saliou, and N. Genay, "Mobile Fronthaul Architecture and Technologies: A RAN Equipment Assessment [Invited]," J. Opt. Commun. Netw. 10, A1-A7 (2018).

[83] R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction (MIT, 2018).

[84] D. Lee et al., "Coordinated multipoint transmission and reception in LTE-advanced: deployment scenarios and operational challenges," in IEEE Communications Magazine, vol. 50, no. 2, pp. 148-155, February 2012.

[85] Q. Zhou et al., "Centralized Digital Self-Interference Cancellation Technique to Enable Full-Duplex Operation of Next Generation Millimeter Wave over Fiber Systems," 2020 Optical Fiber Communications Conference and Exhibition (OFC), 2020, pp. 1-3.

[86] Q. Zhou et al., “Simultaneous Echo Cancellation and Upstream Signal Recovery using Deep Learning in Full-duplex DOCSIS Systems,” Cable-Tec Expo-SCTE, 2020.

[87] Q. Zhou, S. Shen, Y. -W. Chen, R. Zhang, J. Finkelstein and G. -K. Chang, "Simultaneous Nonlinear Self-Interference Cancellation and Signal of Interest Recovery Using Dual Input Deep Neural Network in New Radio Access Networks," in Journal of Lightwave Technology, vol. 39, no. 7, pp. 2046-2051, 1 April1, 2021.

[88] B. Berscheid and C. Howlett, "Full Duplex DOCSIS: Opportunities and Challenges," in IEEE Communications Magazine, vol. 57, no. 8, pp. 28-33, August 2019.

[89] Q. Zhou, S. Shen, C. Hsu, Y. Chen, J. Finkelstein, and G. Chang, "Novel Parallel Interference Cancellation Scheme for Non-Orthogonal Multiple Access in Millimeter-Wave RAN Using Convolutional Neural Network," in 26th Optoelectronics and Communications Conference, P. Alexander Wai, H. Tam, and C. Yu, eds., OSA Technical Digest (Optical Society of America, 2021), paper W4A.6.

[90] F. Lu, M. Xu, L. Cheng, J. Wang, and G.-K. Chang, “Power-division non-orthogonal multiple access (NOMA) in flexible optical access with synchronized downlink/asynchronous uplink,” J. Lightwave Technol, vol. 35, no. 19, pp. 4145–4152, 2017.

[91] V. Mnih, K. Kavukcuoglu, D. Silver et al. Human-level control through deep reinforcement learning. Nature 518, 529–533 (2015). https://doi.org/10.1038/nature14236.

VITA

Qi Zhou was born in August 1991 in Changde, Hunan province, China. He received the BE degree in Electronic Information Engineering from University of Science and Technology of China (USTC) in 2013. He obtained the MS degree in Engineering from the University of Georgia in 2015, where his research focus is microwave photonics advised by Dr. Mable P. Fok. Since 2017, Qi has been working toward his PhD degree in Electrical and Computer Engineering at Georgia Institute of Technology under the supervision of Prof. Gee-Kung Chang. His research interests are machine learning assisted DSP and fiber-wireless system design. From 2015-2017, he worked in the telecom industry as an RF engineer.

Qi has authored and co-authored around 50 peer-reviewed journal articles and international conference papers. He also serves as an active reviewer for prestigious journals including IEEE Access, Photonics Research, Journal of Lightwave Technology, Optics Letters, Photonics Technology Letters, Photonics Journal, Optics Express, etc.