Optical Performance Monitoring in Digital Coherent Communications: Intelligent Error Vector Magnitude Estimation

YUCHUAN FAN

Stockholm, Sweden 2022
Optical Performance Monitoring in Digital Coherent Communications: Intelligent Error Vector Magnitude Estimation

YUCHUAN FAN

Academic Dissertation which, with due permission of the KTH Royal Institute of Technology, is submitted for public defence for the Degree of Doctor of Philosophy on Monday the 5th September 2022, at 1:00 p.m. in FB42, Roslagstullsbacken 21, Stockholm.
Abstract

The rapid development of data-driven techniques brings us new applications, such as fifth-generation new radio (5G NR), high-definition video, Internet of things (IoT), etc., which has greatly facilitated our daily lives. Optical networks as one fundamental infrastructure are evolving to simultaneously support these high-dimensional data services, with a feature of flexible, dynamic, and heterogeneous. Optical performance monitoring (OPM) is a key enabler to guarantee reliable network management and maintenance, which improving network controllability and resource efficiency. Accurately telemetry key performance indicators (KPIs) such as bit error rate (BER) can extend monitoring functionality and secure network management. However, retrieving the BER level metric is time-consuming and inconvenient for OPM. Low-complexity OPM strategies are highly desired for ubiquitous departments at optical network nodes.

This thesis investigates machine learning (ML) based intelligent error vector magnitude (EVM) estimation schemes in digital coherent communications, where EVM is widely used as an alternative BER metric for multilevel modulated signals. We propose a prototype of EVM estimation, which enables monitoring signal quality from a short observation period. Three alternative ML algorithms are explored to facilitate the implementation of this prototype, namely convolutional neural networks (CNNs), feedforward neural networks (FFNNs), and linear regression (LR). We show that CNN conjunction with graphical signal representations, i.e., constellation diagrams and amplitude histograms (AHs), can achieve decent EVM estimation accuracy for signals before and after carrier phase recovery (CPR), which outperforms the conventional EVM calculation. Moreover, we show that an FFNN-based scheme can reduce potential energy and keep the estimation accuracy by directly operating with AH vectors. Furthermore, the estimation capability is thoroughly studied when the system has different impairments. Lastly, we demonstrate that a simple LR-designed model can perform as well as FFNN when the information on modulation formats is known. Such LR-based can be easily implemented with modulation formats identification module in OPM, providing accurate signal quality information.

Keywords
Optical performance monitoring, optical fiber communication, machine learning, coherent optical communications, error vector magnitude
Sammanfattning

Den snabba utvecklingen av informationsteknik ger oss nya applikationer, såsom femte generationens radiosystem (5G NR), högupplöst video, Internet of things (IoT) etc., vilket i hög grad underlättar vårt dagliga liv. För att kunna tillhandahålla dessa högdimensionella datatjänster måste den grundläggande infrastrukturen, de optiska nätverken, utvecklas till att bli mer flexibla, dynamiska och heterogena. Optisk prestandaövervakning (OPM) behövs för tillförlitlig styrning, kontroll och underhåll av det optiska nätverket för att därigenom garantera ett effektivt resursutnyttjande. Telemetri av exakta prestandaindikatorer (KPI) som bitfelsfrekvens (BER) kan utöka övervakningssfunktionaliteten och leda till säkrare nätverkshantering. Att mäta BER direkt är dock för tids- och resurskrävande och OPM-strategier med låg komplexitet behövs för en mer allmän användning i optiska nätverksnoder.


Nyckelord
Optisk prestandaövervakning, optisk fiberkommunikation, maskininlärning, sammanhängande optisk kommunikation, felvectormagnituden
Acknowledgments

I would like to take this opportunity to express my sincere gratitude to all people whose support during my Ph.D. studies, without whom this thesis would have been impossible. I start with my supervisor team. Docent Oskars Ozolins, my main supervisor, thank you for taking me in as a Ph.D. student, and for all the careful guidance, nice discussions, and encouragement I need on the way forward. Prof. Sergei Popov, for hosting me to study at KTH and also for all the support and discussions. Dr. Aleksejs Udalcovs, I was very lucky to have you as a mentor during the early stages of my scientific journey. Thank you for your brilliant ideas, patience, and willingness to talk with me in my tough moment. Docent Xiaodan Pang, I deeply appreciate your support and technical expertise, which are essential to solving difficult obstacles during this journey. Docent Richard Schatz, thank you for your kind help and suggestions during our meetings. Thank all my supervisors for helping me to improve my writing skills and for the interesting discussions we had during lunch or coffee breaks.

I am really grateful to Assistant Prof. Marija Furdek and Dr. Carlos Natalino Da Silva for all the fruitful discussions and the most productive collaborations that we had.

I continue thanking Dr. Anders Berntson for his support from RISE research institutes of Sweden. Dr. Stephane Junique, for his kind help and organizing after-work joyful sports.

I am also grateful for all my friends’ help and support that made my fabulous life in Stockholm. Thanks, Xiao Tan, help me pick up the room key when the first day I arrived in Stockholm. Thanks to Qichen Xu and Yang Song’s group work collaborations to conquer data science courses. Thanks to Yucheng Deng, Yan Xiang, Changle Li, Xiaolei Bian, Hongdi Wang, Qigui Yang, and Jingjian Zhou for cycling, partying, gaming, and all the great moments we have experienced together.

Financial support from the China Scholarship Council (CSC), RISE Research Institutes of Sweden, and AI-NET-PROTECT are highly acknowledged.
List of Publications

Publications included in this thesis:

Paper A


Paper B


Paper C


Paper D

Paper E


Paper F


Paper G


Paper H


Publications not included in this thesis:

Paper A

Paper B

Paper C

Paper D

Paper E

Paper F
Paper G

List of acronyms

ADC  Analog to digital converter
AH   Amplitude histogram
AlGaAs Aluminium GaAs
AR   Augmented reality
AWGN Additive white Gaussian noise
BER  Bit error rate
CD   Chromatic dispersion
CNN  Convolutional neural network
CPR  Carrier phase recovery
CU   Centralized unit
DAC  Digital to analog converter
DC   Data center
DCI  Data center interconnection
DSP  Digital signal processing
DU   Distributed unit
ECL  External Cavity Laser
EDFA Erbium-doped fiber amplifier
EVM  Error vector magnitude
FC   Fully connected
FEC  Forward error correction
FFNN Feedforward neural network
5G NR Fifth-generation new radio
GaAs Gallium Arsenide
IQ   In-phase and quadrature
IM/DD Intensity modulation/direct detection
IQI  In-phase and quadrature imbalance
ITU  International telecommunication union
LO   Local oscillator
LP   Launch power
LR   Linear regression
LSTM Long short-term memory
MAE  Mean absolute error
ML   Machine learning
MSLE Mean squared logarithmic error
mQAM m-ary quadrature amplitude modulation
OOK  On-off keying
OPM  Optical performance monitoring
OSNR Optical signal to noise ratio
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAN</td>
<td>Radio access networks</td>
<td></td>
</tr>
<tr>
<td>RC</td>
<td>Reservoir computing</td>
<td></td>
</tr>
<tr>
<td>ReLU</td>
<td>Rectified linear unit</td>
<td></td>
</tr>
<tr>
<td>ROADM</td>
<td>Reconfigurable optical add-drop multiplexing</td>
<td></td>
</tr>
<tr>
<td>RRU</td>
<td>Remote radio unit</td>
<td></td>
</tr>
<tr>
<td>Rx</td>
<td>Receiver</td>
<td></td>
</tr>
<tr>
<td>SSMF</td>
<td>Standard single-mode fiber</td>
<td></td>
</tr>
<tr>
<td>Tx</td>
<td>Transmitter</td>
<td></td>
</tr>
<tr>
<td>VR</td>
<td>Virtual reality</td>
<td></td>
</tr>
<tr>
<td>WDM</td>
<td>Wavelength division multiplexing</td>
<td></td>
</tr>
</tbody>
</table>
# Contents

1 Introduction 1
   1.1 Background and motivation .......................... 1
   1.2 Overview of the thesis contribution .................. 4
   1.3 Thesis organization .................................. 5

2 Background 7
   2.1 Related techniques in OPM ............................ 8
      2.1.1 Machine learning tools ......................... 10
   2.2 Coherent optical communications ...................... 13
      2.2.1 Modulation formats ............................. 14
      2.2.2 Rx digital signal processing .................... 14
      2.2.3 Metrics for evaluating system performance ...... 16

3 Intelligent EVM estimation schemes 19
   3.1 Dataset collection ................................. 20
   3.2 CNN-enabled EVM estimator ......................... 22
   3.3 FFNN-enabled EVM estimator ........................ 24
   3.4 Linear Regression enabled EVM estimator ............ 29

4 Summary of appended papers 33

5 Conclusions and Future Work 37
   5.1 Conclusions ......................................... 37
   5.2 Future work .......................................... 38

6 Appended papers 47
Chapter 1

Introduction

With the development of communication technology, optical networks are becoming more complex and dynamic to support ever-increase transmission capacity. Advancements in monitoring capability can ensure reliable network operation. This thesis focuses on designing machine learning-based fast signal quality monitoring schemes. In this chapter, the research background and motivation are first introduced in Section 1.1. In Section 1.2, we highlight the contribution of this thesis. Finally, Section 1.3 gives the outline of this thesis.

1.1 Background and motivation

History has witnessed the vigorous development of communication technology from early wired telegraph to modern optical fiber communication applications and integrated services digital networks. In 1880, Alexander Graham Bell used sunlight as a carrier to transmit voice and invented “the optical telephone”, which marked the beginning of optical communication. Two fundamental problems need to be addressed for optical communication: the stable and low-loss transmission media; the other is finding a high-intensity and reliable light source [1]. In the following several decades, optical communication transmission media exploration was fruitless due to attenuation and cost. In 1951, the invention of glass fiber for medical applications was a stepping stone for optical fiber communication, and even the attenuation was about 1000 dB/km. C. K. Kao first theoretically demonstrated the possibility of using glass fiber as the transmission medium in 1966 [2]. Then scientists in Corning Glass Co., USA, successfully manufactured low-loss fiber with attenuation of less than 20 dB/km in 1970 [3]. Simultaneously, intensive research was carried out on light sources. The first semiconductor laser made of Gallium Arsenide (GaAs) was reported in 1962 [4]. Later, the development of Aluminium GaAs (AlGaAs) lasers made the lifetime of the light source and optical detector reaches a practical level of $10^5$ hours.
CHAPTER 1. INTRODUCTION

After several generations of technology evolution, optical fiber communication plays a vital role as the foundation of digital infrastructure. According to the transmission distance, the applications of optical communication systems and networks can be classified as short-reach links (<100 km), metro links (100 ∼ 1000 km), long-haul and transoceanic links (>1000 km). Figure 1.1 illustrates modern optical fiber communication application scenarios. The cost-sensitive short-reach links are large-scale deployed to support access networks, 5G radio access networks (RAN), data center (DC) intra- and inter-links, edge cloud computing networks, etc. Short-reach links are connected to metro links. The metro links are connected across a city or metropolitan area. Techniques such as optical amplifiers, dispersion compensators, wavelength-division multiplexing (WDM), intensity modulation / direct detection (IM/DD), coherent transmission, etc., are used to guarantee reliable and high-capacity transmission. The other end of the metro core connects the long-haul and transoceanic links, which are the core of global networks.

Figure 1.1: The applications of optical communication. DC: data center; CO: central office; RN: remote node; PON: passive optical network; RAN: radio access network; CU: centralized unit; DU: distributed unit; RRU: remote radio unit.

Cisco envisions that global Internet users, devices, and connections will con-
1.1. BACKGROUND AND MOTIVATION

siderably increase in 2023 [5]. There will be 5.3 billion global Internet users and 29.3 billion networked devices in 2023, up from 3.9 billion and 18.4 billion in 2018, respectively. The rapid development of services, such as cloud computing, 5G new radio, 4K/8K high-definition video, virtual reality (VR) / augmented reality (AR), puts forward higher and higher bandwidth demand. This leverages the development of high-capability and high-speed optical transmission technology. Coherent optical technology and advanced modulation formats have become the mainstream techniques in deploying over 100 G optical links. Simultaneously, optical networks evolve more complex, heterogeneous, and dynamic to meet the demands of various services. Service providers need to monitor the link performance to maintain and manage their ecosystems. The currently deployed optical performance monitoring (OPM) modules based on simple coherent front-ends and tunable local oscillators are imperfect, which mainly work in the optical domain and report the optical signal to noise ratio (OSNR). However, OSNR cannot truly reflect the signal quality when the system is impacted by other linear/nonlinear impairments rather than additive white Gaussian noise (AWGN). Advanced accurate physical layer monitoring of the transmission performance, i.e., bit error rate (BER) monitoring, is demanded by customers [6]. The BER can intuitively provide evaluation for a service or light path quality and a predefined threshold.

Error vector magnitude (EVM) is a viable alternative BER performance metric and is widely used in wireless communication systems to qualify the received signal’s quality [7]–[11]. In optical fiber communication systems, the relation between EVM and other figure of merits, such as OSNR and BER, has been investigated in [12], [13] under the AWGN channel. Figure 1.2 shows an error vector measurement between a received signal symbol and its associate ideal symbol on the in-phase (I) and quadrature (Q) planes. The EVM is calculated as the root-mean-square (RMS) of a collection of error vectors over a large number of received symbols [7], [12], [13]. This process needs to accumulate millions of received symbols to retrieve the EVM over a monitoring period, which is time-consuming and improperly tracking the fast network dynamics. Recognizing this need, developing an EVM estimation scheme cost-effectively is the research focus to promote the monitoring functionality of EVM in OPM modules.

As one branch of artificial intelligence, machine learning (ML), like its name, enables machines to simulate or realize human intelligence ability by learning historical data. ML-based applications have been launched in the optical communication field [14]–[16]. Research activities focus on improving the efficiency of network operation and management, such as automatically network performance monitoring and fault diagnosis [17], [18], resource allocation [19], quality of transmission (QoT) estimation [20], [21], routing and wavelength assignment (RWA) [22], etc. These applications can be separated according to three clas-
Figure 1.2: The error vector between the ideal (reference) and received signal symbols.

Basic ML algorithms: (i) supervised learning, mapping input to output through learning statistical features of label-driven feedback, e.g., OPM and QoT estimation; (ii) unsupervised learning, representing and mining potential features of the unlabeled dataset, e.g., signal dimension reduction, traffic clustering; and (iii) reinforcement learning, building and updating a model by obtaining the maximum cumulative rewards when agents take a series of behaviors in the environment, e.g., network planning, reconfiguration [23]. Therefore, we are naturally looking for supervised learning-powered intelligent EVM estimation schemes. How to explore valuable features from a small amount of signal sequence in EVM estimation and light DSP routine are the research interest.

1.2 Overview of the thesis contribution

The contribution of this thesis is dedicated to exploring the intelligent EVM estimation scheme in coherent optical systems. In this regard, the research content of this thesis can be characterized into three parts, which follow designing a fast signal quality monitoring scheme and evolving it into a more flexible one.

- (1) In this part, the EVM estimation is first formulated as a regression task. We attempt to estimate signal quality from a signal sequence as short as possible. The estimation scheme here relies on convolutional neural networks (CNNs) in conjunction with image data sets. The image data sets are a collection of signal’s constellation diagrams generated before and after the CPR module in the DSP routine. With this method,
1.3. THESIS ORGANIZATION

the signal quality can be estimated from a short observation period. The first part concludes Paper A and Paper B.

- (2) Then, we proposed a time- and energy-efficient EVM estimation scheme, which utilizes feedforward neural networks (FFNNs) to estimate EVM from signal amplitude histogram (AH) vectors. This scheme can accurately estimate the signal quality before CPR with a short response time. Besides, the average GPU power is reduced by half compared with the CNN-powered scheme. Furthermore, the tolerance of the proposed scheme against various impairments in coherent optical systems is evaluated. The second part concludes Paper C, Paper D, and Paper E.

- (3) Finally, we evaluate the capability of a conventional machine learning algorithm, i.e., linear regression (LR), for the EVM estimation task. This scheme provides good accuracy when we train a model for each modulation format separately. In particular, a modulation format identification module in OPM can help to select pre-trained EVM estimation models. Thus, this scheme could be an idea to design future intelligent OPM schemes with parsimony. The third part concludes Paper F, Paper G, and Paper H.

1.3 Thesis organization

The rest of the thesis is organized as follows. Chapter 2 presents the background and optical performance monitoring state-of-art. Besides, coherent optical communication systems' basics are introduced, e.g., digital signal processing and performance metrics. Chapter 3 demonstrates three machine learning models for EVM estimation. The first section outlines the methodology for dataset collection in coherent optical communication systems. The second section describes the proposed CNN-based EVM estimation scheme and the implementation performance before and after carrier phase recovery is studied. The third section proposes a low-complexity EVM estimation scheme based on FFNN. The last section discusses a simple linear regression-powered EVM estimation scheme. Chapter 4 gives a summary of appended papers included in this thesis. Chapter 5 summarizes the overall conclusions and proposes future work, followed by a collection of publications in Chapter 6.
Chapter 2

Background

In the context of the information age, as one of the essential technologies in telecommunications, optical communication technology provides energy for the development of information technology. Optical transmission links have evolved to have the feature of high spectrum efficiency, narrow channel spacing, long-distance transmission and high data rate, which brings challenges to network management. Optical performance monitoring modules are indispensable in optical networks to guarantee the quality of service of high-capacity coherent transmission links [24]–[26]. Figure 2.1 illustrates the evolution of OPM technology with the development of optical communication networks. The number and requirements of monitoring performance parameters are increasing. The early deployed low-speed optical fiber communication systems have a simple network structure without multiplexing techniques; the system typically operates with IM/DD and simple modulation formats, e.g., on-off keying (OOK). Therefore, few performance parameters are considered in the basic OPM, such as optical channel optical power, wavelength, in-band OSNR, etc. The implementation of advanced technologies, such as WDM, erbium-doped fiber amplifier (EDFA), high-order modulation, reconfigurable optical add-drop multiplexing (ROADM), etc., diversifies transmitted signal modulation formats and data rate, which in turn makes the network structure dynamically changeable. Accordingly, higher requirements are put forward for optical performance parameters. For example, the nonlinear distortion and crosstalk make the traditional monitoring of OSNR, chromatic dispersion (CD), etc., more complex. The performance monitoring appears to have new requirements. It is necessary not only to realize accurate performance parameter monitoring in a complex environment but also to realize dynamic and real-time simultaneous processing of multiple parameters to maintain the network state changing. This chapter revisits related schemes in OPM and machine learning tools and briefly describes modules in coherent optical fiber communication systems.
CHAPTER 2. BACKGROUND

OPM technology
Simple Advanced
IM/DD CD, PMD, non-linear noise, OSNR Data rate, MFI, multiple parameters monitoring, BER, EVM

In-band OSNR, optical power

Digital coherent QAM, OFDM
Self-coherent Analog coherent
...

Transmission schemes
IM/DD
OOK, PAM, DMT

P2P Single channel EDFA WDM ROADM

Optical network schemes

Optical network properties

Static Heterogeneous

Figure 2.1: The OPM technology evolution map. P2P: point-to-point; EDFA: erbium-doped fiber amplifier; WDM: wavelength division multiplexing; ROADM: reconfigurable optical add-drop multiplexing; IM/DD: intensity modulation/direct detection.

2.1 Related techniques in OPM

OPM is a set of measurements to acquire physical layer performance indicators, which basically has two common types: OPM at intermediate nodes and OPM at optical transceivers [26]. Besides, OPM is tightly related to QoT modeling, which can provide channel quality-related parameters as feedback inputs for QoT estimation [36]. Common parameters include OSNR, CD, Q-factor, polarization-dependent loss (PDL), fiber nonlinearity, BER, etc. Figure 2.2 depicts an optical link installed with OPM modules at different monitoring locations. Conventional techniques for OPM can be divided into two types: time-domain monitoring technology and frequency-domain monitoring technology. The parametric information of the transmitted signal can be obtained by effectively representing the features from the time-domain waveform and frequency-domain spectrum. The time-domain monitoring techniques are mainly based on asynchronous sampling [37], polarization-related processing [38–40], Mach-Zehnder interferometer [41], [42], and digital coherent receivers [43]. Frequency-domain monitoring schemes include analyzing optical spectrum [44], radio frequency (RF) tone [45], [46], and DSP-based pilot tone scheme [47]. However, it is hard to build accurate, low-complexity modeling and monitoring tools with a conventional analytical method in modern heterogeneous
2.1. RELATED TECHNIQUES IN OPM

Table 2.1: Optical performance monitoring based on machine learning tools.

<table>
<thead>
<tr>
<th>ML Tools</th>
<th>Feature Representation</th>
<th>Monitored Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNN [27]</td>
<td>Constellation Diagram</td>
<td>OSNR</td>
</tr>
<tr>
<td>CNN [28]</td>
<td>Asynchronous sampled data</td>
<td>OSNR</td>
</tr>
<tr>
<td>CNN [29]</td>
<td>Phase Portrait</td>
<td>OSNR, CD, DGD Modulation Format, Bit Rate</td>
</tr>
<tr>
<td>FFNN [30]</td>
<td>Amplitude Histogram</td>
<td>OSNR, Modulation Format</td>
</tr>
<tr>
<td>FFNN [31]</td>
<td>Asynchronous Complex Histogram</td>
<td>OSNR, Modulation Format, Nonlinear Noise Power</td>
</tr>
<tr>
<td>FFNN [32]</td>
<td>Asynchronous Amplitude Histogram</td>
<td>OSNR, CD, PMD</td>
</tr>
<tr>
<td>LSTM [33]</td>
<td>Frequency domain transformation</td>
<td>OSNR, Nonlinear Noise Power</td>
</tr>
<tr>
<td>LSTM [34]</td>
<td>Four-tributary digital output</td>
<td>OSNR</td>
</tr>
</tbody>
</table>

Figure 2.2: Optical fiber transmission links with enabled OPM modules.
optical networks [36]. Besides, some estimation schemes are only suitable for experienced engineers, and the potential manual operation errors are inevitable. Many of the proposed schemes need extra devices, such as optical spectrum analyzers, interferometers, polarimeters, etc., which will increase implementation complexity. Moreover, some schemes need to modify the transmitter by inserting pilot tones, which will limit scenarios of practical implementation. On the other hand, data-driven machine learning has become popular again in the optical communication community. The development of intelligent OPM schemes by neural networks in coherent optical systems has greatly increased the accuracy and efficiency of automatic monitoring optical networks. Such as feedforward neural networks, convolutional neural networks, long short-term memory (LSTM) neural networks, photonic reservoir computing (RC) enabled applications in modulation formats identification (MFI) and OSNR, CD, non-linear noise power, differential group delay (DGD), PMD, etc., monitoring [27]–[35], [48]–[51]. Table 2.1 briefly summarizes these techniques. In this thesis, we investigated EVM estimation schemes in coherent optical transmission with different enabling tools, such as CNN, FFNN, and LR. Our EVM estimation prototype utilizes CNN and constellation diagram, which enables automatically reading the EVM value from a small amount of symbols constellation diagram. Then, we studied EVM estimation without part of DSP, CNN- and FFNN-powered schemes are developed. Finally, we designed a simple but with limited functionality solution, i.e., an LR-based EVM estimator.

2.1.1 Machine learning tools

In general, most of the existing ML-based OPM schemes are supervised learning, which means that we train a model with labels. A model is actually a function approximator, mapping an input variable to an output variable. There are basically two types of pattern recognition tasks: classification and regression. The classification task predicts a label of a discrete category. The mapping function predicts a category label for a given observation sample, e.g., recognizing dog and cat. The regression task predicts a label of continuous numbers. The mapping function predicts a continuous output variable for given information, e.g., predicting house price. Figure 2.3 illustrates the classic applications. To deal with these tasks, ML algorithms, such as LR, support vector machine, random forest, Bayesian learning, neural networks, etc., can be implemented [52]. Since EVM values are continuous numbers, we develop estimation schemes based on the regression algorithm. In the following, we mainly introduce LR, FFNN, and CNN algorithms, which are used for developing EVM estimation schemes in this thesis work.

Linear regression is simple regression analysis, mapping the independent
2.1. RELATED TECHNIQUES IN OPM

Figure 2.3: Pattern recognition tasks (a) classification, (b) regression.

variable \( x \) and the dependent variable \( y \) by the linear function,

\[
\hat{y} = wx + b
\]  

(2.1)

where \( w \) and \( b \) are weight and bias. The \( w \) and \( b \) are updated during the learning process by minimizing residual error between true and estimated values (also known as loss function), e.g., in the mean squared error (MSE) form

\[
\mathcal{L} = \arg \min_w \sum_{i=1}^{n} (y_i - \hat{y}_i).
\]  

(2.2)

A typical feedforward neural network contains input, hidden, and output layers. Figure 2.4 depicts a 2-hidden-layer neural network, and the input \( x \) arrives output layer is unidirectional propagation. In addition to the nonlinear activation function, each neuron in the hidden layer has a similar weight and bias operation as linear regression. The neurons in the two adjacent layers are fully connected. The propagation can be expressed as

\[
\begin{align*}
\mathbf{s}_1 &= \mathbf{w}_1 \mathbf{x} + \mathbf{b}_1, \\
\mathbf{s}_2 &= \mathbf{w}_2 \mathcal{F}_1(\mathbf{s}_1) + \mathbf{b}_2, \\
\hat{\mathbf{y}} &= \mathbf{w}_3 \mathcal{F}_2(\mathbf{s}_2) + \mathbf{b}_3.
\end{align*}
\]  

(2.3) (2.4) (2.5)

The \( \mathbf{w}_i, \mathbf{b}_i \) are weight matrix and bias matrix for layer \( i - 1 \) to layer \( i \). The \( \mathbf{s}_i, \mathcal{F}_i(\cdot) \) are \( i \) layer output matrix and activation function, respectively. In modern neural networks, the commonly option for activation function is the rectified linear unit (ReLU), which is applied independently to each neuron in the hidden layer. After the forward propagation process, we compute gradients of error (a.k.a loss or cost) matrix between the true label matrix and estimated value matrix by the loss function. Then, the \( \mathbf{w}_i, \mathbf{b}_i \) can be iteratively updated by the gradient descent algorithm during the backward propagation of gradients.
In FFNN, we see that neurons of one layer are fully connected to the previous and next layer, which brings about the potential problem of parameters expansion when we process images. Thanks to convolutional layers, CNNs are popular for many computer vision tasks, e.g., image recognition, object detection, etc. The well-known deep CNN architectures, e.g., AlexNet, ResNet, VGGNet, etc., have high accuracy on pattern recognition tasks. Figure 2.5 shows the general structure for CNN. In the convolution layer, the neurons of each layer are connected through the filter kernel as a medium. Convolution operation can mine local structure; this process can be understood as that we use a filter kernel to filter each small region of the image to obtain the features of these small regions. The same filter kernel is shared in the image, and the output feature maps still retain their original position relationship after the convolution operation. The \( i^{th} \) feature map of convolutional layer \( l \) output can be expressed as

\[
    x^l_i = \mathcal{F}\left(\sum_{j=1}^{m} x^{l-1}_{j} \ast k^l_{i,j} + b^l_i\right) \tag{2.6}
\]

where \( x^{l-1}_{j} \) is the \( j^{th} \) feature map in layer \( l - 1 \). The \( k^l_{i,j} \) denotes the filter kernel connecting feature maps \( x^{l-1}_{j} \) and \( x^l_{i} \), and \( b^l_i \) is the bias. After the convolution operation, we normally have a max-pooling layer to reduce the dimensions of the feature map passed to the next layer. The fully connected (FC) layers are followed by a series of convolutional layers to recognize the output.
2.2. COHERENT OPTICAL COMMUNICATIONS

Coherent optical communication technology is often used to deploy 100G and beyond transmission scenarios. Figure 2.6 shows a typical single-polarization coherent optical communication system. The transmitter (Tx) is composed of Tx DSP, digital to analog converter (DAC), in-phase and quadrature (I/Q) modulator, and external cavity laser (ECL). The Tx DSP module maps the incoming bits sequence to the selected modulation format, then filtered by the Nyquist pulse filter to generate the band-limited signal. In addition, the DAC generates the required electrical signal. The Mach-Zehnder IQ modulator converts electrical signal to optical signal by means of modulating the electrical signal onto a continuous-wave light source. The optical signal reaches the coherent receiver front-end after several spans of the optical filter, e.g., standard single-mode fiber (SSMF), and is amplified by an erbium-doped fiber amplifier (EDFA). On the receiver side (Rx), the balanced photodetectors detect coherent beating product between transmitted signals and the local oscillator (LO) laser. Finally, the analog to digital converter (ADC) generates data traces, which are then processed by the offline Rx DSP module to recover the transmitter signal.

Figure 2.6: Basic structure of coherent optical communication system. DAC: digital to analog converter; ECL: external cavity laser; EA: electrical amplifier; EDFA: erbium-doped fiber amplifier; SSMF: standard single-mode fiber; LO: local oscillator; ADC: analog to digital converter.
CHAPTER 2. BACKGROUND

2.2.1 Modulation formats

![Constellation examples of QAM signals.](image)

The Shannon-Hartley theorem is derived from C. E. Shannon in 1948 [61], which mathematically gives the limit of information that can be error-free transmitted over a channel, using ideal error correction and coding schemes. The capacity of one additive white Gaussian noise (AWGN) channel is defined as maximum mutual information (MI) between the channel input \(X\) and output \(Y\) (discrete format):

\[
C = \max_{P_X(x)} I(X, Y) = \frac{1}{2} \log_2(1 + \text{SNR}) \quad \text{[bits/symbol]},
\]

where \(P_X(x)\) is the probability of \(x\), and SNR is input power divided by noise power. Nowadays, the transceivers’ bandwidth is becoming a limiting factor. Thus, coherent optical systems normally transmit modulated multilevel symbols from selected constellations (a.k.a. known as symbol alphabets) to increase capacity. Quadrature amplitude modulation (QAM) is a widely used family of single-carrier modulation schemes, which encodes information bits onto both discrete amplitude and phase combinations in a complex plane. Figure 2.7 shows the constellation examples for different modulator order QAM with unit power. High order modulation formats can increase the spectral efficiency, which in turn improves the channel capacity, as shown in Fig. 2.8.

2.2.2 Rx digital signal processing

Offline digital signal processing is needed at the receiver to recover the received I and Q samples. Figure 2.9 illustrates the necessary modules in the DSP routine and how the constellation changes from 'chaotic' to clear. In general, the RX DSP includes chromatic dispersion compensation, timing recovery, adaptive equalization, carrier phase recovery, and BER counting. Due to the inherent characteristics of optical fiber, the optical signal pulses will be broadened after transmission. Therefore, it is hard to discern correct symbols without compensation; as a result, the received samples are chaotically distributed on the constellation diagram. The compensation process can be implemented in the
2.2. COHERENT OPTICAL COMMUNICATIONS

Figure 2.8: The mutual information (MI) versus signal to noise ratio for QAM signal.

Figure 2.9: Rx DSP modules in coherent optical systems.
CHAPTER 2. BACKGROUND

time domain or frequency domain by different algorithms [62]–[64]. Then, the
timing recovery module is responsible for synchronizing the clock between the
transmitter and receiver ADC sampling rate to produce samples with optimal
SNR [65]. An adaptive equalizer then equalizes the samples to eliminate time-
varying impairments of the transmitted channel. Algorithms like the constant
modulus algorithm (CMA) [66], and the multi-modulus algorithm [67] can be
used for single and multilevel amplitude constellations, respectively. The am-
plitude levels are corrected in this process, and clear circles appear on the
constellation diagram. At this stage, the constellation circles are mainly im-
acted by phase noise. A non-ideal transmitter and LO laser induce the phase
noise, and a higher-order modulation format is more susceptible. Besides, the
existing CPR schemes for high order modulation formats are based on Viterbi
& Viterbi algorithm [68], [69], and the blind phase search (BPS) algorithm [70].
Finally, symbols detection and forward error correlation (FEC) is performed in
the BER counting module to retrieve the information bits and compared them
with transmitted bits to calculate the BER.

2.2.3 Metrics for evaluating system performance

In the design and evaluation of a communication system, we should establish a
set of indicators that can reflect the performance of all aspects of the system.
There are many performance evaluation metrics for optical communication sys-
tems, the common ones are as follows: eye diagram opening penalty, Q-factor,
optical signal to noise ratio, bit error rate, error vector magnitude, chromatic
dispersion tolerance, etc.

- BER: BER is the ultimate performance indicator describing the reliability
  of the optical link. BER is defined as the proportion of the error re-
  ceived bits in the total number of bits transmitted, that is, BER is the
  probability of the error transmitted bits in the transmission system. A
  lower BER value indicates a more reliable transmission system. However,
  demodulation and decoding take a long time to retrieve transmitted bits.

- Eye diagram: Eye diagram is an accumulation of 1-bit and 0-bit current
  periods on an oscilloscope, which provides a quick and qualitative measure
  of a received signal. The smaller the ‘eye opening’ of the eye diagram
  indicates the greater inter-symbol interference (ISI). On the contrary, the
  larger the opening of the eye diagram, the smaller of ISI. When there
  is noise, the edge line of the eye diagram becomes fuzzy and broad.
  The larger noise will result in the thicker edge lines, the more fuzzy eye
  diagram, and the smaller ‘eyes opening’.

- Q-factor: The quality factor Q (Q-factor), like its name, can describe
  the signal quality and can be used to represent the bit error rate of the
2.2. COHERENT OPTICAL COMMUNICATIONS

The calculation formula of the Q value is as follows:

\[ Q = \frac{I_1 - I_0}{\sigma_1 + \sigma_0} \tag{2.8} \]

where the \( I_1, I_0 \) are values of 1-bit and 0-bit current, the \( \sigma_1, \sigma_0 \) are values of the standard deviation of 1-bit and 0-bit current. A higher Q value is associated with a smaller BER, and vice versa.

- **OSNR**: OSNR plays an important role in the estimation and measurement of the optical transmission system that includes optical amplifiers since BER is related to OSNR. The OSNR is defined as the ratio of optical signal power to noise power within a bandwidth of 0.1 nm. The power of the optical signal generally takes the peak value, while the power of noise generally measures amplified spontaneous emission noise within 0.1 nm optical channel gaps. In general, a higher OSNR value leads to a lower BER, however, this does not hold when systems operate with fiber nonlinearity.

- **EVM**: EVM suggests a degree of impairment for complex modulation formats at the symbol level. It counts the magnitude difference of each DSP compensated symbol to its ideal (or reference) constellation point. The value can directly transfer to BER and OSNR when the system noise is dominated by AWGN [12], [13].
Chapter 3

Intelligent EVM estimation schemes

To efficiently monitor the signal, we propose estimating signal quality from a short observation period, as illustrated in Fig. 3.1. We collect \( L = 100 \) signal sequences, and one signal sequence contains \( N \) number of symbols per constellation cluster. The length of the signal sequence is equal to \( N \) multiplied by the modulation order \( M \). The \( N \times M \times L \) is related to a training period. We can use the data from the training period to train an ML-based model to mine signal features from limited symbols. After the training period, the trained model is available to estimate EVM from a short observation period. We study how a short observation period is enough to provide a good estimation accuracy. Conventional mathematical EVM calculation is implemented in the fully DSP recovered symbols. Our study has designed EVM estimation schemes to skip some DSP modules when the symbols have clear amplitude circles. In

![Figure 3.1: The proposed EVM monitoring mechanism. N is the number of symbols, M is the modulation order, L is the number of collected signal sequence.](image-url)
this chapter, we first describe dataset collection, and then we discuss proposed ML-based intelligent EVM schemes for the limited length of signal symbols.

### 3.1 Dataset collection

![Figure 3.2: The schematic diagram of 32 GBaud optical coherent transmission link.](image)

In our study, we collect both simulation and experimental datasets. The simulation datasets are using software VPItransmissionMaker™ and co-simulate with MATLAB. Initially, we built a 32 GBaud coherent optical back-to-back (OB2B) transmission system to investigate EVM estimation schemes. Figure 3.2 shows a schematic diagram of the setup. The coherent transmitter includes an I/Q driver, a dual-parallel Mach-Zehnder modulator (MZM), and a continuous wave laser. The I/Q driver is used for bits to symbols mapping and pulse shaping. The considered modulation formats are QPSK and mQAM, which are widely used in coherent optical systems. Then the MZM modulates a pulse-shaped signal onto a CW lightwave to generate optical signals. An ONSR adjustable module controls the quality of the signal. We use different OSNR regions for different modulation formats to ensure the system BER is under a practical interest threshold. Hard-decision forward error correction (HD-FEC, 3.8e-3) threshold for QPSK, soft-decision forward error correction (SD-FEC, 1e-2) for 16QAM and 64QAM. Thus, we set corresponding 10 OSNR values with a 2 dB step size in 12~30 dB, 20~38 dB, 26~44 dB for QPSK, 16QAM, and 64QAM, respectively. After coherent detection, we save signals before and after the carrier phase recovery module in the DSP routine for further dataset generation.

The EVM label of a received signal is an averaged value for symbols. Here, we use $2^{19}$ symbols. The mathematical calculation is defined as

$$EVM_{rms} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} |E_r(i) - E_t(i)|^2}$$

(3.1)
where \( E_r(i) \) and \( E_t(i) \) denote the received (measured) signal vector and ideal signal vector, respectively. The normalization term \( |E_{t,0}| \) is either the maximum magnitude or average magnitude of the used modulation. Normally the value of EVM is expressed in percentage or decibel. In this thesis, We use maximum magnitude as \( |E_{t,0}| \) and percentage expression of EVM. The ideal signal vector can be obtained by transmitting pilot symbols, we use centroids for each constellation cluster determined by the k-means clustering algorithm as \( E_t(i) \), which can achieve a high accuracy [71], [72]. Figure 3.3 shows the QPSK constellation diagram and centroids for constellation cluster obtained k-means algorithm.

![Figure 3.3: The QPSK constellation and cluster centroids of k-means algorithm.](image)

Table 3.1: Specification of the fiber link.

<table>
<thead>
<tr>
<th>Specification</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baud rate</td>
<td>32 GBAud</td>
</tr>
<tr>
<td>Generated symbols</td>
<td>( 2^{19} )</td>
</tr>
<tr>
<td>Chromatic dispersion</td>
<td>( 16e-6 ) s/m²</td>
</tr>
<tr>
<td>Attenuation coefficient</td>
<td>0.2 dB/km</td>
</tr>
<tr>
<td>Nonlinear refractive index</td>
<td>2.6e-20 m²/W</td>
</tr>
<tr>
<td>Each span length of SSMF</td>
<td>100 km</td>
</tr>
<tr>
<td>QPSK measurement distance</td>
<td>400 km (4 spans)</td>
</tr>
<tr>
<td>16QAM measurement distance</td>
<td>300 km (3 spans)</td>
</tr>
<tr>
<td>64QAM measurement distance</td>
<td>200 km (2 spans)</td>
</tr>
</tbody>
</table>

Then, we insert multiple spans of standard single-mode fiber (SSMF) between transceivers to simulate realistic transmission scenarios. Each span of SSMF is 100 km and contains an inline EDFA. The detailed SSMF configuration is shown in Table 3.1. Besides, the launch power is optimized to ensure
the transmission system is working in a linear region. We set 45 dB OSNR at the transmitter and measured OSNR 0.1 nm resolution after a set of distance transmissions by using an optical spectrum analyzer (OSA). The maximum transmission distance for QPSK, 16QAM, and 64QAM is configured at 2000 km, 1500 km, and 1000 km, respectively.

To study the length of signal sequence effects on estimation accuracy, we capture signal with options N number per constellation points from 10 to 1000. We collect 100 short signal sequences for each transmission scenario to make the dataset. Then, the captured signals are represented in the constellation diagram and AH before and after CPR to investigate different estimation schemes further. Figure 3.5 shows the collected dataset examples for 16QAM.

<table>
<thead>
<tr>
<th>OSNR [dB]</th>
<th>22.5</th>
<th>26.9</th>
<th>30.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>After CPR</td>
<td>IQ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before CPR</td>
<td>IQ</td>
<td>16-bin AH</td>
<td></td>
</tr>
</tbody>
</table>

Figure 3.5: The 16QAM dataset examples.

### 3.2 CNN-enabled EVM estimator

Initially, starting from the definition of EVM, we explored the possibility of inferring the EVM value from the constellation diagram. It is natural to formulate it as a task of image recognition. First, we test the performance of the CNN-based EVM estimator on the dataset without fiber effect, i.e., the OB2B
dataset. In the training process, we tune hyperparameters by adjusting different hidden layer parameters, e.g., the number of filter kernels, the number of convolution layers, filter size, etc. Besides, we set the loss function as the mean squared logarithmic error (MSLE) to update the neural network, including small estimation error effects

\[ \mathcal{L} = \frac{1}{k} \sum_{i=1}^{k} (\log(y_i + 1) - \log(\hat{y}_i + 1))^2. \] (3.2)

The optimized CNN architecture for the EVM task has four convolutional layers with 8, 16, 16, 8 filters and a 3-by-3 kernel; two fully connected layers with 500 and 100 neurons; output layer with 1 neuron [73]. Figure 3.6 shows test performance on this architecture, and black lines are references obtained from conventional EVM calculation on 100 symbols per cluster signal. We can see that the 100-symbol/cluster dataset achieves a normalization mean absolute error (MAE) of 3.7% for QPSK, 2.2% for 16QAM, and 1.1% for 64QAM. Paper A [73] provides a detailed discussion about the performance of CNN in conjunction with the constellation diagram EVM estimation scheme.

![Figure 3.6](image)

Figure 3.6: Test performance of the proposed EVM monitoring scheme when datasets containing constellation diagrams with 10 to 500 symbols/cluster. The Ref. curves are baselines obtained using the conventional method applied for the 100-symbol/cluster dataset. (a) QPSK, (b) 16QAM, and (c) 64QAM [73].

Then, we are trying to skip some DSP modules to relax the requirement of EVM inference. Paper [74] reports that the power distribution of main elements for 64QAM coherent metro/DCI operation is 40% for ADC/DAC, 5% for Tx shaping, 13% CD compensation, 7% for polarization/polarization mode dispersion compensation, 14% for CPR, and 21% for FEC. Among them, the Rx DSP elements CPR and FEC take the second most power consumption, lower than the 5% of analog front-end ADC/DAC. Thus, we consider designing the EVM estimation scheme before CPR. As described in Chapter 2.2.2, the signal is mainly affected by phase noise before the CPR processing. The received complex symbols rotate in random degrees on the IQ complex plane, which appear
as ring circles after several accumulations. Alternately, amplitude histograms can show the signal quality and avoid the influence of phase noise. The number of bins when plotting AH can affect signal representation. Thus, we include different AH images with 8, 16, and 64 bins. Figure 3.7 shows the MAE versus the number of symbols per mQAM constellation cluster, where the MAE is calculated over all modulation formats. We can see that the CNN-powered scheme can also provide a good EVM estimation by skipping CPR. One should be noticed that we didn’t realize the power normalization for EVM true labels calculation in Paper A and Paper B. Later after the power normalization is included, the accuracy can also be increased.

![Figure 3.7: Test performance of CNN-based EVM monitoring scheme for different datasets](image)

**Figure 3.7:** Test performance of CNN-based EVM monitoring scheme for different datasets [75].

### 3.3 FFNN-enabled EVM estimator

Bringing the EVM estimation before the CPR, the CNN method first recognizes the amplitude on the AH image by the convolution operations, then feeds feature maps to FC layers for EVM estimation. The convolution operation is computationally demanding, which may outweigh the benefits of operating before CPR and short signal sequence. Alternatively, the AH vectors can directly input to a feedforward neural network model, which is actually the FC layers in CNN architecture. The detailed results for different FFNN architectures, e.g., accuracy, and measure amount of computation in floating-point operations (FLOPs), can be found in Paper C [76]. After tuning hyperparameters, the
3.3. **FFNN-ENABLED EVM ESTIMATOR**

Optimized architecture has four hidden layers with neurons 1000, 500, 500, and 100, respectively. Figure 3.8 illustrates test results MAE versus length of signal sequence for mQAM signals employing different estimation schemes. All schemes achieve an MAE below 0.5% at 100-symbol/cluster unless 8-bin AH is used. Besides, FFNN conjunction with the 64-bin AH vectors scheme has a comparable estimation accuracy performance to CNN operating with IQ after CPR.

![Figure 3.8: Test performance versus N symbols/cluster for various mQAM signal representations](image)

**Next,** we evaluate energy consumption for FFNN and CNN-powered EVM estimators by using NVIDIA System Management Interface (nvidia-smi) tool to record the instantaneous power of GPU at a specific time [77]. This experiment is performed on a device with a 2.4 GHz Intel Xeon E5-2630-v3 CPU, 64 GB of RAM, and a GTX TITAN Black GPU. We use the nvidia-smi tool to record the starting- and ending-time information for training and testing on the 100-symbol/cluster 64-bin AH dataset. The real-time GPU power for CNN and FFNN estimators is shown in Fig. 3.9. Having these measurements, the energy consumption can be calculated as:

$$E = \sum_{i=1}^{T} P_i. \quad (3.3)$$

Table 3.2 summarizes the corresponding energy consumption, where the training dataset size and testing dataset size are 1650 samples and 550 samples,
respectively. The low-complexity FFNN saves over 95% energy compared with the CNN case. Therefore, the FFNN-based scheme not only can ensure decent EVM estimation accuracy but also provide energy savings.

![Graph showing real-time GPU power for CNN and FFNN-powered EVM estimator](image)

Figure 3.9: The real-time GPU power for CNN and FFNN-powered EVM estimator [76].

<table>
<thead>
<tr>
<th>Energy consumption</th>
<th>Training [J]</th>
<th>Testing [J]</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNN</td>
<td>194487.9</td>
<td>290.1</td>
</tr>
<tr>
<td>FFNN</td>
<td>2856.7</td>
<td>10.3</td>
</tr>
</tbody>
</table>

Table 3.2: Energy consumption for a 64-bin AH dataset.

Finally, we evaluate the proposed FFNN scheme tolerance when the system is in the presence of impairments, such as residual IQ imbalance (IQI), fiber nonconformity, and laser phase noise. These impairments might impact the EVM estimation accuracy. We studied the impact of ideally separated degradation sources in simulation and an experimental investigation of the impact of laser phase noise. The considered modulation formats are square 64QAM (Sq-64QAM) and circular 64QAM (C-64QAM). The imperfection of TX and Rx induces IQ imbalance, which will cause a mismatch in signals’ amplitude and phase [78]-[80]. An imperfectly compensated system will result in the signal containing a small amount of residual IQ imbalance. In this study, we narrow down that effect to Rx IQ imbalance since the IQ imbalance of transmitters is normally pre-calibrated and compensated in commercial transceivers. We conduct a simulation for Rx IQ 1% amplitude imbalance and -5°~5 degrees phase...
imbalance in the optical hybrid. Optical fiber nonlinearities can also degrade signal quality, which is non-negligible for long-distance and high launch power (LP) WDM systems [81]. In the optical fiber nonlinearity simulation setup, we modify the single-channel setup (as shown in 3.4) to a 32 Gbaud 5-channel optical fiber transmission system using the 50 GHz international telecommunication union (ITU) grid. The center channel is under test and launch power values range from –4 dBm to +10 dBm per channel with a 1 dBm increment. The cost-efficient semiconductor lasers will introduce phase noise to systems, which may, in turn, impact the proposed EVM estimation scheme. We build simulation and experimental 28 Gbaud optical coherent setups for the laser phase noise study. In the simulation setup, we keep the LO laser linewidth (LW) at 200 kHz and sweep the Tx laser linewidth from 100 kHz to 4.1 MHz. Thus, the emulated laser linewidth scenarios are from 300 kHz to 4.3 MHz. In the experimental setup, we emulate corresponding laser linewidth scenarios by Wiener’s phase noise model in the Tx DSP [70], [82]. The digitally generated phase noise can be expressed as [70]

$$\varphi_{pn}(k) = \varphi_{pn}(k - 1) + \Delta \theta,$$  

(3.4)

where the $\Delta \theta$ is an independent and identical distributed (i.i.d.) random Gaussian variable with the variance [82]

$$\sigma_{\theta}^2 = 2\pi \Delta f T_s.$$  

(3.5)

The $\Delta f$ denotes the linewidth of the signal laser, and $T_s$ is the symbol period. Figure 3.10 illustrates examples of signals before/after CPR constellations, before CPR AHs, BER, and EVM versus OSNR curves, when signals are impacted by different degradation sources. It can be observed that different laser linewidth scenarios have a very similar constellation diagram after CPR and AH before CPR. In contrast, the constellation diagram before CPR presents a different shape. This indicates that the amplitude histogram, in principle, cannot provide phase information.

In the training process, we tested the EVM estimator generalization capability by three training schemes for each impairment type: (i) the case with insignificant impairment (e.g., IQI = 0 deg, LP = 6 dBm, LW = 300 kHz) to use as a benchmark, (ii) all cases: training a single model to incorporate all impairments, and (iii) separate cases: training an independent model for each impairment case. When a single model is used, the FFNN-based EVM estimator achieves a mean estimation error below 0.05% and 0.2% for IQ imbalance with and without 1% amplitude imbalance, respectively. The performance of fiber nonlinear is closer to the residual amplitude imbalance, below 0.3%. More detailed results and discussion about IQ imbalance and fiber nonlinearity can be found in Paper E. As discussed before, the amplitude histogram cannot provide
CHAPTER 3. INTELLIGENT EVM ESTIMATION SCHEMES

Figure 3.10: The examples of before/after CPR constellations, before CPR amplitude histograms, BER versus OSNR curves, and EVM versus OSNR curves for different degradation source: (a) IQ imbalance without amplitude imbalance, (b) IQ imbalance with 1% amplitude imbalance, (c) fiber nonlinearity, (d) laser phase noise [76].
phase information. Thus, when we train all linewidth scenarios together, the FFNN-EVM estimator loses the capability to distinguish the linewidth scenarios (see blue curves in Fig. 3.11 (a)). In this case, the model needs to be trained separately for different linewidth scenarios to achieve desirable performance, as shown in the yellow curves shown in Fig. 3.11 (a). However, the opposite EVM estimation results are observed in blue curves in Fig. 3.11. After further investigation, we noticed that the Lithium Niobate-based IQ modulator in the experimental setup has a significant bias drift, which results in a variation of IQ imbalance in the system. Thus, in addition to the laser phase noise, the degradation source of the experimental dataset also has an IQ imbalance. Finally, after more impairments are included in simulations, the estimation tendency of simulation scenarios can match the experimental setting much better.

![Figure 3.11: The mean estimation error of different training schemes versus linewidth on (a) simulation datasets, and (b) experimental datasets.](image)

### 3.4 Linear Regression enabled EVM estimator

So far, the developed neural-network-based EVM estimation schemes have shown a promising capability for monitoring signal quality. However, under a certain using case, a parsimonious or 'less intelligent' model might be also competent for this work. Linear regression is a classic analytical algorithm, which is the fundamental element of FFNN. We design the LR-based EVM estimator with the AH vector $X_i = [x_{i,1}, x_{i,2}, ..., x_{i,64}]$ as multiple linear regression (a.k.a. multivariable linear regression):

$$\hat{y}_i = X_i \beta + \varepsilon_i$$  \hspace{1cm} (3.6)  

where $\beta$ is the regression coefficient, and $\varepsilon_i$ denotes the i.i.d. estimation error. The regression coefficient can be obtained by using least-squares estimation
\[ \hat{\beta} = (X^T X)^{-1} X^T y. \] (3.7)

We use the long-haul amplitude histogram dataset to evaluate the LR-based EVM estimator performance and include the FFNN-based EVM estimation scheme as a benchmark. The dataset is divided by 75% and 25% for training and testing purposes. Since the LR-based scheme might be affected by nonlinear variables, we explore the generalization capability by training a different number of modulation formats in one LR model: (1) LR1, a single type of modulation format; (2) LR2, 16QAM, and 64QAM; (3) LR3, QPSK, 16QAM, and 64QAM. Figure 3.12 shows the estimation performance of the normalized MAE (NMAE) versus the number of symbols on the testing dataset, where the FFNN reference case is trained on QPSK, 16QAM, and 64QAM three modulation formats. The NMAE allows us to compare the estimation accuracy across different models:

\[ NMAE[\%] = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\%, \] (3.8)

where \( n \) is the number of test samples in a model, \( y_i \) and \( \hat{y}_i \) denote \( i^{th} \) test sample true EVM and estimated EVM. It can be observed that the accuracy of the LR-based estimator degrades along with an increasing number of modulation formats.

Figure 3.12: Normalized mean absolute error (NMAE) of EVM estimation vs. number of symbols for various schemes of training one model. LR1: one modulation format; LR2: two modulation formats; LR3: three modulation formats; FFNN: three modulation formats [85].
When we implement the LR-based model for each single modulation format, which can achieve a performance as well as an FFNN scheme. For a 1000-symbol and a 6400-symbol long signal sequence, the LR1 model achieves NMAE below 3.5% and 1%, respectively. Since the MFI module can provide the modulation format information in OPM, such a low-complexity LR scheme can be easily implemented and provide accurate signal quality information. More detailed results and experimental validation can be found in Paper H.
Chapter 4

Summary of appended papers

Paper A:


In this paper, we designed and proposed a prototype of an EVM estimation scheme in coherent optical communication systems, which facilitate fast and accurate signal quality monitoring from a short observation period. This scheme utilizes a convolutional neural network and constellation diagram. We studied how short/long symbol sequences are required to achieve a decent estimation accuracy.

**Contribution of the author:** Design the prototype model for EVM monitoring, development the machine learning estimator, analysis and characterization of results, preparation of first draft and updated versions.

Paper B:


In this paper, we explored the EVM estimation scheme by reducing the required DSP, and provided insight into the possibility of skipping a carrier phase recovery module. We explored different signal representations, e.g. amplitude histograms, constellation diagrams. We keep using CNN and image representa-
CHAPTER 4. SUMMARY OF APPENDED PAPERS

Contribution of the author: Design the EVM monitoring scheme before the CPR, development the machine learning estimator, analysis and characterization of results, preparation of first draft and updated versions.

Paper C:


This paper studied feedforward neural networks and convolutional neural networks-based EVM estimation schemes. More detailed performance analyses of different types of signal representation results are contained. Besides, it provides experimental validation of the proposed schemes. Furthermore, the energy consumption of the two kinds of neural networks is studied. Results indicate that the FFNN conjunction with AH vectors can be used to perform time-sensitive and accurate EVM estimation for mQAM signal quality monitoring in coherent communication systems.

Contribution of the author: Design and simplicity the EVM monitoring scheme before the CPR, development the machine learning estimator, analysis and characterization of results, preparation of first draft and updated versions.

Paper D:


In this paper, we conducted an experimental investigation on the FFNN-based EVM estimator for a system with different laser phase noise. The results show that an FFNN-based estimator can predict an EVM when the system is impaired by laser phase noise, whereas further investigation needs to be conducted.

Contribution of the author: Conduct the experiment with different laser phase noise, making datasets from the collected signals, analysis and characterization of results, preparation of first draft and updated versions.
Paper E:


This paper is an extension of Paper D, where we studied the FFNN-based EVM estimation scheme for systems with more impairments. In particular, we analyzed the signal quality monitoring capabilities in the presence of residual IQ imbalance, fiber nonlinearity, and laser phase noise. We studied the impact of ideally separating these degradation sources with the amplified spontaneous emission noise on the EVM estimation. The simulation results show that the proposed EVM estimator is robust against IQ imbalance and fiber nonlinearity. In principle, for laser phase noise, the amplitude histogram cannot provide phase information, which will affect the estimation accuracy. We verified it in both simulation and experiment. We found that the experimental investigation of laser phase noise is a hybrid type of impairment, which also has residual IQ imbalance exists in the signal to help reveal phase information on AHs.

**Contribution of the author:** Simulating the coherent communication system in the presence of isolated degradation source, making datasets from the collected signals, analysis and characterization of results, preparation of first draft and updated versions.

Paper F:


This paper designed a simple linear regression-based EVM estimation scheme for coherent optical communication systems. The comparison of estimation capability between the linear regression and FFNN-based model is conducted. The results show that the LR-based scheme can be used for EVM estimation when the modulation format is known a priori.

**Contribution of the author:** Design and simplicity the EVM estimation scheme, development the machine learning model, analysis and characterization of results, preparation of first draft and updated versions.
CHAPTER 4. SUMMARY OF APPENDED PAPERS

Paper G:


In this paper, we experimentally demonstrated a linear regression-based EVM estimation scheme monitoring signal quality before CPR. This scheme is implemented along with a modulation format identification module to extend OPM functionality.

Contribution of the author: Implementing MFI and LR-based EVM estimation scheme, development the machine learning model, analysis and characterization of results, preparation of first draft and updated versions.

Paper H:


This paper is an extension of Paper F and Paper G, where the linear regression-based EVM estimation is thoroughly studied in simulation and experiment. In addition, more detailed performance analysis results are contained.

Contribution of the author: Design and simplicity the EVM estimation scheme, development the machine learning model, collecting both simulation and experiment datasets, analysis and characterization of results, preparation of first draft and updated versions.
Chapter 5

Conclusions and Future Work

5.1 Conclusions

This thesis represents a collection of designing intelligent EVM estimation schemes for monitoring signal quality in digital coherent communications. Specifically, the designed schemes enable us to infer an EVM from a short observation period, providing accurate signal quality information for optical performance monitoring. First, we have proposed a scheme estimating EVM from a constellation diagram, which is a human-friendly graphical representation of the signal. This scheme is designed using a convolutional neural network, and the constellation diagram is generated from a short signal sequence. We have investigated the estimation accuracy considering CNN architecture and the number of symbols contained in the constellation diagram. Both optical back-to-back and long-haul transmission simulations are investigated. The results show that this scheme can accurately monitor the signal quality with a fast response time. Then, we have examined the proposed CNN model EVM estimation capability when operating before the carrier phase recovery block in the DSP routine. More graphical representation types of the signal are investigated in this work, such as before/after CPR constellation diagrams and before CPR amplitude histograms. The results indicate that, before carrier phase recovery, the AHs contain sufficient features for monitoring signal quality.

Further, we investigate a simplified EVM estimation scheme by directly inputting AH vectors into a feedforward neural network instead of convolutional layers recognizing amplitude levels from AH images. In this way, the EVM estimator can be further optimized in terms of complexity and energy consumption. In our case, the FFNN-based scheme achieves a decent performance for EVM inference and saves over 95% energy compared with the CNN-based solution. We noticed that the laser phase noise might impact the EVM estimation accuracy induced by cost-efficient semiconductor lasers when we performed the
experimental validation. Thus, we conduct an experimental investigation on the FFNN-based EVM estimator for a system with different laser phase noise. Although the results show that this scheme can distinguish different laser linewidth scenarios, the amplitude histograms, in principle, exclude phase information. In further investigation, we ideally separate the impairment sources in simulation. The results show that the experimental setup degradation sources are hybrid type, which contains a different degree of IQ imbalance for each linewidth scenario. This can help FFNN recognize the features in linewidth scenarios.

Finally, a linear regression-based EVM estimation model is designed to cooperate with existing modulation formats identification functionality in optical performance monitoring. The performance is validated in both simulation and experiment, where the LR-based model can achieve comparable performance as the FFNN scheme when the modulation format information is known.

5.2 Future work

As discussed, the amplitude histogram can not provide signals’ phase information, which may trigger faults when systems have different phase noises. Therefore, we will investigate how to effectively estimate EVM when the system has laser phase noise in future works. Besides, future work should also explore accurately EVM scheme operating with low-speed ADC since ADC could be occupied another major power consumption part. In this case, the effects of transmission distance and chromatic dispersion should also be evaluated. Last but not least, the accurate mapping of EVM to BER with different system parameters can be studied.
References


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


