UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

Exploiting optical signal analysis for autonomous communications

Diogo Gonçalo Sequeira

Universitat Politècnica de Catalunya

Optical Communications Group

Exploiting Optical Signal Analysis for Autonomous Communications

Diogo Gonçalo Sequeira

diogo.goncalo.sequeira@upc.edu

A thesis presented in partial fulfillment of the requirements for the degree of

Philosophy Doctor

Advisor: Dr. Luis Velasco

Co-advisor: Dr. Antonio Napoli

September 2022

Optical Communications Group (GCO)

Universitat Politècnica de Catalunya (UPC)

C/ Jordi Girona, 1-3

Campus Nord, D4-213

08034 Barcelona, Spain

Acknowledgments

Firstly, I'd like to thank my supervisors, Prof. Luis Velasco and Dr. Antonio Napoli, for their vital advice, continuous support, and patience during these three long years. I'd also like to thank Dr. Marc Ruiz, for all the deep technical discussions and to be always available to clarify me about all my doubts. Also, thank you Dr. João Pedro and Dr. Nelson Costa, for all the valuable meetings and discussions which were important for the success of this PhD thesis.

I'd like to thank all GCO colleagues I met during the last three years, all of you were important during this journey. Thank you, Dr. Alba, Dr. Sima, Dr. Fatemeh, Pai, Morteza, Masab, Hailey, Mariano, Prazunika, Shaoxuan and Pol.

Finally, I'd like to express my deep gratitude to my dear parents, João and Glória, my dear sister, Inês, and all my family and friends. Without your unconditional support and love, it would be harder. Thank you all, my success is your success!

To my brother, André.

Abstract

Optical communications have been extensively investigated and enhanced in the last decades. Nowadays, they are responsible to transport all the data traffic generated around the world, from access to the core network segments. As the data traffic is increasing and changing in both type and patterns, the optical communications networks and systems need to readapt and continuous advances to face the future data traffic demands in an efficient and cost-effective way.

Moreover, with the continuous advances in optical communications systems and networks, they are nowadays more complex infrastructures, in both data and control plane. Consequently, introducing more robust control and management tasks to the optical communications, for example based on machine learning methods, to bring them more intelligence and drive them towards the autonomous operation is a key driver to the future optical communications.

This PhD thesis focuses on investigate and analyze the optical signals in order to extract useful knowledge from them to support the autonomous lightpath operation, as well as to lightpath characterization. In addition, investigations of innovative optical networks architectures and transmission technologies to improve the overall optical performance and cost-effectively route the optical signals are also objectives of this PhD thesis.

The first objective of this PhD thesis is to investigate the optical transmission feasibility of optical signals based on high-order modulation formats (MF) and high symbol rates (SR) in hybrid filterless, filtered and flexible optical networks. It is expected a higher physical layer impairments impact on these kinds of optical signals that can lead to degradation of the quality of transmission. In particular, the impact of the optical filter narrowing arising from the node cascade is evaluated. The obtained simulation results for the required optical-signal-to-noise ratio in a cascade up to 10 optical nodes foresee the applicability of these kinds of optical signals in such scenarios. By using high-order MF and high SR, the number of the optical transponders cab be reduced, as well as the spectral efficiency is enhanced.

The second objective focuses on MF and SR identification at the optical receiver side to support the autonomous lightpath operation. Nowadays, optical transmitters can generate several optical signal configurations in terms of MF and SR. To increase

the autonomous operation of the optical receiver, it is desired it can autonomously recognize the MF and SR of the incoming optical signals. In this PhD thesis, we propose an accurate and low complex MF and SR identification algorithm based on optical signal analysis and minimum Euclidean distance to the expected points when the received signals are decoded with several available MF and SR. The extensive simulation results show remarkable accuracy under several realistic lightpath scenarios, based on different fiber types, including linear and nonlinear noise interference, as well as in single and multicarrier optical systems.

The final objective of this PhD thesis is the deployment of a machine learning-based digital twin for optical constellations analysis and modeling. An optical signal along its lightpath in the optical network is impaired by several effects. These effects can be linear, e.g., the noise coming from the optical amplification, or nonlinear ones, e.g., the Kerr effects from the fiber propagation. The optical constellations are a good source of information regarding these effects, both linear and nonlinear. Thus, by an accurate and deep analysis of the received optical signals, visualized in optical constellations, we can extract useful information from them to better understand the several impacts along the crossed lightpath. Furthermore, by learning the different impacts from different optical network elements on the optical signal, we can accurately model it in order to create a partial digital twin of the optical physical layer. The proposed digital twin shows accurate results in modeled lightpaths including both linear and nonlinear interference noise, in several lightpaths configuration, i.e., based on different kind of optical links, optical powers and optical fiber parameters. In addition, the proposed digital twin can be useful to predict quality of transmission metrics, such as bit error rate, in typical lightpath scenarios, as well as to detect possible misconfigurations in optical network elements by cooperation with the software-defined networking controller and monitoring and data analytics agents.

Resumen

Las comunicaciones ópticas han sido ampliamente investigadas y mejoradas en las últimas décadas. En la actualidad, son las encargadas de transportar la mayoría del tráfico de datos que se genera en todo el mundo, desde el acceso hasta los segmentos de la red troncal. A medida que el tráfico de datos aumenta y cambia tanto en tipo como en patrones, las redes y los sistemas de comunicaciones ópticas necesitan readaptarse y avanzar continuamente para, de una manera eficiente y rentable, hacer frente a las futuras demandas de tráfico de datos.

Además, con los continuos avances en los sistemas y redes de comunicaciones ópticas, hoy en día las infraestructuras son más complejas, tanto en el plano de datos como en el de control. En consecuencia, la introducción de tareas de control y gestión más sólidas para las comunicaciones ópticas, por ejemplo, basadas en métodos de aprendizaje automático, para brindarles más inteligencia y conducirlas hacia la operación autónoma es un factor clave para las comunicaciones ópticas del futuro.

Esta tesis doctoral se centra en investigar y analizar las señales ópticas con el fin de extraer de ellas conocimiento útil para apoyar el funcionamiento autónomo de las conexiones ópticas, así como para su caracterización. Igualmente, la investigación de arquitecturas de redes ópticas innovadoras y tecnologías de transmisión para mejorar el rendimiento óptico y enrutar de manera eficiente las señales ópticas también son objetivos de esta tesis doctoral.

El primer objetivo de esta tesis doctoral es investigar la viabilidad de transmisión de señales ópticas basadas en formatos de modulación de alto orden y altas tasas de símbolos en redes ópticas híbridas con y sin filtros. Se espera un mayor impacto de las degradaciones de la capa física en este tipo de señales ópticas que pueden conducir a la degradación de la calidad de transmisión. En particular, se evalúa el impacto de la reducción del ancho de banda del filtro óptico que surge tras atravesar una cascada de nodos. Los resultados de simulación obtenidos para la relación señal óptica/ruido requerida en una cascada de hasta 10 nodos ópticos prevén la aplicabilidad de este tipo de señales ópticas en tales escenarios. Mediante el uso de modulación de alto orden y altas tasas de símbolos, se reduce el número de transpondedores ópticos y se mejora la eficiencia espectral.

El segundo objetivo se centra en la identificación de formatos de modulación y tasas de símbolos en el lado del receptor óptico para respaldar la operación autónoma de la conexión óptica. Hoy en día, los transmisores ópticos pueden generar varias configuraciones de señales ópticas en términos de formatos de modulación y tasas de símbolos. Para aumentar el funcionamiento autónomo del receptor óptico, se desea que pueda reconocer de forma autónoma la configuración de las señales ópticas entrantes. En esta tesis doctoral, proponemos un algoritmo de identificación de formatos de modulación y tasas de símbolos preciso y de baja complejidad basado en el análisis de señales ópticas cuando las señales recibidas se decodifican con varios formatos de modulación y tasas de símbolos disponibles. Los extensos resultados de la simulación muestran una precisión notable en varios escenarios realistas, basados en diferentes tipos de fibra, incluida la interferencia de ruido lineal y no lineal, así como en sistemas ópticos de portadora única y múltiple.

El objetivo final de esta tesis doctoral es el despliegue de un gemelo digital basado en aprendizaje automático para el análisis y modelado de constelaciones ópticas. Una señal óptica a lo largo de su trayectoria en la red óptica se ve afectada por varios efectos. Estos efectos pueden ser lineales, por ejemplo, el ruido proveniente de la amplificación óptica, o no lineales, por ejemplo, los efectos Kerr de la propagación de la fibra. Las constelaciones ópticas son una buena fuente de información sobre estos efectos, tanto lineales como no lineales. Por lo tanto, mediante un análisis preciso y profundo de las señales ópticas recibidas, visualizadas en constelaciones ópticas, podemos extraer información útil de ellas para comprender mejor los diversos impactos a lo largo del camino propagado. Además, al aprender los diferentes impactos de los diferentes elementos de la red óptica en la señal óptica, podemos modelarla con precisión para crear un gemelo digital parcial de la camada física óptica. El gemelo digital propuesto muestra resultados precisos en conexiones que incluyen ruido de interferencia tanto lineal como no lineal, en varias configuraciones basados en diferentes tipos de enlaces ópticos, potencias ópticas y parámetros de fibra óptica. Además, el gemelo digital propuesto puede ser útil para predecir la calidad de las métricas de transmisión, como la tasa de error de bit, así como para detectar posibles errores de configuración en los elementos de la red óptica mediante la cooperación con el controlador de red, el monitoreo y agentes de análisis de datos.

Table of Contents

List of Figures

List of Tables

Page

Chapter 1

Introduction

This chapter is an introduction to this PhD thesis. Specifically, Section 1.1 motivates and introduces optical communications and its last advances. The main goals of this PhD thesis are introduced and explained in Section 1.2 and, in Section 1.3, the methodology to be used throughout this PhD thesis is introduced. Finally, the thesis outline is described in Section 1.4 and the explanation how we labeled the references from the literature is provided in Section 1.5.

1.1 **Motivation**

The global exponential increasing of the digital data traffic shows no signs of slowing down. With the current advances on 5G and beyond technologies, cloud-based services, high-speed access networks, etc., in the coming years, is expected an even faster and higher increasing of the global internet protocol (IP) data traffic generated around the world [Cisco]. Optical transport network (OTN) [ITU-T G.709] is nowadays responsible to transport all data traffic and, to face that unprecedent increasing has been pushed to further innovative developments in both data and control plane, from advanced digital signal processing (DSP) blocks to real-time management tasks based on monitoring and data analytics (MDA) [EON16], [Ag16]. Furthermore, current data traffic is based on ultra-reliable and low-latency (URLL) services and, particularity, in optical communications, the coexistence of heterogeneous data traffic, i.e., in terms of modulation formats (MF), symbol rates (SR), channel spacings, originated from optical transponders (TRx) differently configurated, etc., leads to highly flexible and dynamic optical networks environments [EON16]. In this regard, optical communications needed to be continuously reconsidered, retuned, managed and monitored in an efficient, reliable, and cost-effective way to deal with this continuous increasing and variability of current and future data traffic demands.

Optical communications are mostly characterized by high-bandwidths (common C band of about 5 THz), low insertions losses (around 0.25 dB/km in C-band) and low bit error rate (BER) enabled by robust forward error correction (FEC) algorithms [FCS12]. Nowadays, wavelength and polarization division multiplexing (WDM and PDM) technologies are extensively deployed in optical communication to increase their overall performance [Mu00], [Gn08]. Further improvements are enabled by advanced DSP blocks at the TRx, e.g., for compensation of signal impairments [Zh19]. All those advanced technologies together are paramount for the success of high-capacity and high-speed optical communications systems and networks [Ta04], [Fl20].

Another technology enabling the overall improvement of the spectral efficiency (SE) allowing more flexibility and programmability in optical networks management is the elastic optical networks (EON) [EON16] alongside with decentralized and disaggregated software-defined networking (SDN) solutions at the control plane [Th16], [Pa18]. SE is improved by managing most efficiently the available spectrum to transmit the data traffic demands, for example, regarding the spectral assignment, in EON the data traffic can be assigned in different spectral widths with spectral granularity of 12.5 GHz, instead of a conventional fixed spectral width of 50 GHz [ITU-T G.694.1]. Sliceable bandwidth variable transponders (S-BVT) [Sa15], which are evolved in the recent year and currently are complex sub-systems capable to work based on different channel configuration, e.g., high-order MF and high SR [Pe20], and the programmable wavelength selective switches (WSS) [Ba06][St10] are key optical network elements enabling the implementation of advanced and most efficient EONs.

In recent years, to face the arising of URLL services, there are some interests and investigations to increase the presence of advanced optical communications technologies closer to the end users, e.g., in the access and metro aggregation network segments, instead of only concentrate on core ones. At the same time, is desired to avoid higher physical layer impairments (PLI) that can degrade the overall quality of transmission (QoT) of an optical signal, such as optical filtering and nonlinear interference (NLI) noise effects. One technology which enables that is the hybrid filterless and filtered optical networks [Sh18], [Uz21]. In these kinds of networks, the metro core segments are composed by few active optical nodes based on WSS, known as reconfigurable optical add/drop multiplexers (ROADMs), while in the metro aggregation ones, i.e., closer to the access, the optical nodes are based on passive optical splitters/couplers, allowing to avoid the optical filtering effects from the WSS, as well as reducing the latency inside these nodes. Because of the absence of optical filters in the filterless segments, the planning and management of these networks must be done carefully, for example, to avoid channel overlapping and noisily systems. Most recently, another technology proposed to cost-effectively face the data traffic increasing and changing in data patterns, enabling most efficient point-to-multipoint (P2MP) connections, and minimizing the impact of the NLI noise, is the digital subcarrier multiplexing (DSCM) [Qi14], [Infinera.1]. The key

aspect of DSCM technology is the propagation of multiple digital subcarriers (SC), e.g., 4, 8 or more, in the same optical channel instead of only one optical carrier to transport the data traffic. One of the main benefits of using DSCM in optical communications is to keep high capacities while using lower SR per SC, which, in general, leads to lower penalties caused by optical fiber propagation impairments, such as chromatic dispersion (CD) and nonlinear Kerr effects [NFO13].

In order to improve even more the overall performance of the optical communications networks and systems, the management, monitoring and prediction of the several parameters involved in these complex environments is important. In the recent years, with the emerging of numerous artificial intelligence (AI) and machine learning (ML) applications, optical communications can also be enhanced by AI/ML-based methods and algorithms [Ra18], [Po21]. Network operation tasks, such as data analytics of the monitored optical networks parameters and QoT estimation, are nowadays receiving special attention and efforts to future implementation and improvements by AI/ML-based solutions. One of the most interesting fields in that regard is the development of digital twin (DT) of the optical communications networks and systems [Wa21]. Nowadays, accurate and low-computational DTs are useful for several use-cases, from QoT estimation to fault management, and can be a key tool to real-time operations, to reduce design margins in optical networks design, to support the coordination with SDN controllers and MDA agents, etc.

In summary, to face the anticipated high and variable data traffic demands, current and future optical communications networks and systems are expected to be highly flexible and dynamic based on SDN solutions and supported by MDA strategies. Optical technologies complexity is moving from the optical fiber links, to the TRx by advanced DSP blocks, and to ROADM nodes based on advanced architectures and able to perform monitoring tasks. Future optical networks are expected to have several network architectures and parameters to be managed and monitored. In this context, this PhD thesis is manly motivated by the deeper understanding, modeling and analysis of these highly dynamic systems and networks taking advantage of AI/ML-based methods, as well as traditional analytical models.

1.2 Goals of the Thesis

In line with the above motivation, the main and final objective of this PhD thesis is a comprehensive investigation about the future highly flexible and dynamic optical communications systems and networks, realizing how we can develop useful models and algorithms to improve and deeper understand these networks and their overall performance based on analysis of the optical signals. Specifically, monitoring, analysis, and exploitation of the PLIs affecting the optical signals, such as optical filtering and NLI noise, are key aspects throughout this investigation work.

Specifically, three goals are defined to achieve the main and final goal.

G.1 – Investigate the applicability of high-order MF and high SR in future optical networks scenarios

This goal focuses on understand if high-order MF and high SR can be applied to future optical metro network scenarios, specifically, in hybrid filtered and filterless optical metro networks, where the optical filtering impact is expected to be lower.

G.2 – Develop methods for blind and autonomous MF and SR identification at Receiver (Rx) side

This goal targets at developing accurate and low-complex identification method to implement in the DSP blocks at Rx side, making possible an autonomous operation of the Rx in terms of MF and SR identification. This goal will be applied to both high-order MF and high SR regimes, as well as to low-order MF and low SR regimes, e.g., based on DSCM systems.

G.3 – ML-based DT for optical constellations (OC) modeling and analysis

This goal aims at investigating ML-based models to extract useful information from the monitored OC to further analysis, as well as to ML-based modeling of the OCs for accurately abstract the optical physical layer in different optical networks scenarios. Particularity, differentiate between linear interference (LI) and NLI noise modeling, to acquire useful knowledge from both.

1.3 Methodology

Regarding the methodology to be used during this PhD thesis we can divide it in two groups: *i*) method to generate accurate synthetic datasets, i.e., a MATLAB-based simulator of the optical physical layer, i.e., data plane, to generate optical signal data samples; and *ii*) ML-based methods for modeling and analysis purposes, using the generated synthetic datasets, for example, to assist the control plane with MDA tasks.

Let us first focus on the methods to obtain optical signal data samples by modeling the desired optical physical layer, i.e., the optical network elements crossed by the optical signal along its lightpath from Tx to Rx. This method is commonly used to investigate optical system performance [Mo15][Se18], since it allows to investigate in a controlled environment the several PLI impact on the desired lightpath scenarios. Furthermore, the investigation through simulation setups allows taking the first conclusions regarding the possibility of implementation of such optical connections based on their performance, by varying a large set of network parameters and features that would be impossible and highly expensive to perform experimentally. In that regard, there are some works where accurate analytical models are presented to model the optical physical layer. In general, optical communications systems can be modeled by discrete-time statistical models. For example, the fiber optic propagation, including propagation effects like CD and NLI

noise, can be modeled by solving the nonlinear Schrödinger equation (NLSE) using the split-step Fourier method (SSFM) [NFO13], [Sh14]. Realistic physical modeling of the optical filters' response is also possible as it is proposed in [Pu11], instead of using less realistic analytical optical filters models, like the super-Gaussian filters. Additionality, in this PhD thesis, experimental optical filter measurements are averaged out to obtain even more realistic filter shapes [Ra18]. In this way, we can conclude that an accurate model of the optical physical layer is accessible. Consequently, after an accurate model of the desired lightpath scenarios to investigate, we can perform Monte-Carlo (MC)-based simulations, for example, to obtain optical signal samples in order to generate accurate data sets to test the algorithms and models proposed throughout this PhD thesis. Additionally, by doing MC-based simulations, we are able to estimate the optical system performance by estimation of common QoT metrics, such as BER and optical signal-to-noise ratio (OSNR). The model of the optical physical layer, to generate signal samples as well as to BER and OSNR predictions, is supported by the MATLAB software.

Regarding the ML-based methods, in this PhD thesis, we focus on low-complexity feedforward artificial neural networks (ANN) and concatenation of several ANN creating a kind of deep neural networks (DNN). ANNs are computational models able to learn complex and nonlinear relationships between the input and output layers, supported by a given number of hidden layers. In line with the introduced thesis goal G.3, we consider ANN and DNN to develop the proposed DT for OC modeling and analysis. Additionality, we take advantage of Gaussian mixture models (GMM) to characterize OC samples and then fed the proposed ANN models. After generating the desired datasets, by using the developed MATLAB-based optical physical layer simulator, we can perform data analytics and, consequently, acquire useful knowledge to train different ANN to several use-case, such as ANN concatenation for optical physical layer modeling. The training, test and validation of the several ML-based methods proposed throughout this PhD thesis is mainly supported by Python programming language, taking advantage of advanced libraries, such as TensorFlow.

1.4 Thesis Outline

The remainder of this thesis is organized as follows.

Chapter 2 provides the needed background to easier understand the research carry out throughout this PhD thesis. Specifically, background on optical communication, optical network architectures and transmission technologies as well as OC and ML is provided.

Chapter 3 reviews the state-of-the-art associated with the defined thesis goals. The optical filtering impact in optical networks, MF and SR identification and DT for optical communications are briefly reviewed.

Chapter 4 is related with the G.1 and investigates the applicability of high-order MF and high SR in hybrid filterless and filtered optical networks. This chapter is focused on the congress publication [ACP21].

Chapter 5 focuses on G.2 and proposes an accurate and low complex algorithm to MF and SR identifications. This chapter is partially based on the journal publication [JSAC21].

Chapter 6 concentrates on G.3 and proposes OCATA. OCATA is a partial DT for OC modeling and analysis in disaggregated and mixed disaggregated-proprietary optical networks. This chapter is based on the journal publication [JOCN22] and on conference papers [ECOC21.1] and [ECOC21.2].

Chapter 7 concludes this PhD thesis.

1.5 Contributions and References from the Literature

For the sake of clarity and readability, references contributing to this PhD thesis are labelled using the following criteria: [<conference/journal acronyms> <Year (*yy*)[.autonum]>], e.g., [ECOC21.1] or [JSAC21]; in case of more than one contribution with the same label, a sequence number is added.

The rest of the references to journal articles and conference papers are labelled with the initials of the first author's surname and year of publication, e.g., [Ra18] or [Pe20]. Additionality, references for books are labelled with the initials of the book title and year of publication, e.g., [EON16]. Finally, references to norms or standards are labeled with its identification, e.g., [ITU-T G.709].

Chapter 2

Background

This chapter has the main goal to familiarize about the several technologies investigated throughout this PhD thesis. Most precisely, in Section 2.1, we provide a brief background on optical communications and its developments in the last years. Section 2.2 describes innovative optical networks architectures, focus on hybrid filterless and filtered optical networks, as well as advanced optical transmission technologies, such as DSCM. Section 2.3 explains about OC and its main characteristics. In Section 2.4, a summary of the ML-based methods and algorithms applied to the optical communications are introduced. Finally, Section 2.5 concludes this chapter.

2.1 Optical Communications

In this section, we provide a background about optical communications, how they have been evolved in the last years and why investigations about PLI in optical signals are important to better understand and manage their impact on overall optical communication system performance.

Since the invention of the laser and the propose to propagate it over an optical fiber, the optical communications systems and networks have been advanced in all vectors [Ag16]. Firstly, the development of pure silica optical fibers allows to transmit the light, i.e., an optical signal, over low attenuation regimes, around 0.2 dB/km, over optical systems based on C-band. Additionally, in the middle of 1980's, the development of dispersion compensating fibers (DCF) to compensate to the CD effect in optical fiber propagation, lead to a significant improvement in optical communications systems [Gr05]. The early commercially available optical communications systems were mainly based on intensity modulation and direct detection (IM/DD) [Ag16]. However, the need to push forward the optical

communications system capacity and reach, leads to developments of the optical coherent detection technology, enabling PDM and most advanced DSP blocks, as well as advances in optical amplification, mainly based on erbium-doped fiber amplifiers (EDFA) [EDFA02]. The latter together with additional optical network elements, such as optical splitters/couplers, wavelength blockers and WSSs, bring the possibility of transparent optical networks, where there is no need optical-electrical-optical (OEO) conversion of a signal along its route, i.e., lightpath, between the transmitter (Tx) and receiver (Rx).

WDM is a key technology in optical communications since it enables the propagation of different wavelengths in the same optical fiber [OWN06]. For example, typical WDM signals, based on fixed frequency grid of 50 GHz and C-band of 4.8 THz, can carry up to 96 wavelengths, i.e., 96 different optical channels, which can be independently routed, managed, and configured to transport the data traffic most efficiently.

In order to further improvements in optical communications networks, i.e., not only increasing the capacity and reach but also to increase flexibility and dynamism, optical network elements to multiplex and route the optical signal are paramount. In that regard, ROADMs play an important role, and in the last years they have been advanced from fixed add/drop structures to colorless, directionless and contentionless (CDC) flexible ROADMs, bringing the desired flexibly and more freedom to route transparently the WDM signals [Co15]. In general ROADMs blocks are built with typical optical network elements, such as optical splitters/couplers, WSSs and EDFAs. Most recently, the advances on coherent TRx technology turns possible more sophisticated DSP blocks, also based on advanced ML-based algorithms, able to compensate for the transmission impairments, for example by digital backpropagation (DBP) techniques [Na14], [Zh18], avoiding the use of DCF and reducing the overall capital and operational expenditures (CapEx and OpEx).

In the recent tears, one of the most challenges in coherent optical communications systems and networks are the mitigation and monitoring of both LI and NLI noise [Zhu19], since the higher number of optical channels differently configured, in terms of optical power, MF, SR, etc., propagated through different types of optical fibers originate multiples signal-to-noise interactions along a lightpath. The nonlinear optical fiber propagation in single mode fibers (SMF) can be modeled by derivate the NLSE for one polarization by [Ma97]

$$\frac{dE_x}{dz} = -\frac{\alpha}{2}E_x + \frac{j\beta_2}{2}\frac{d^2}{dt^2}Ex - j\gamma\frac{8}{9}(|Ex|^2 + |Ey|^2)Ex \tag{2-1}$$

where Ex and Ey are the complex envelope of the propagated signal at x and y polarization, respectively, at time t and distance z. The α, β_2 and γ are the fiber attenuation, dispersion, and nonlinear parameters, respectively. Considering the averaged random evolution of the polarization effects over long optical fiber distances, as well as ideal TRx, the Eq. (2-1) can be generalized for both polarizations

[Zhu19]. Additionality, because Eq. (2-1) does not have analytical solution, the SSFM was proposed to numerically solve it at the cost of high computational requirements due to large number of fast Fourier transform (FFT) and inverse FFT (IFFT) at each step considered. To face the trade-off between accuracy and computational requirements, other models have been proposed mainly focus on the estimation of the accumulated NLI noise power, such as the Gaussian noise (GN) model [Po12], [Po14], at the cost of assuming some uncertainties and thus, leading to higher optical network design margins. However, recently, some studies show the good accuracy of GN model in multi-vendor optical networks [Fe20], [Cu22].

By Eq. (2-1), we can conclude that in optical fiber propagation, an optical signal is mainly impacted by three effects: the fiber losses, chromatic dispersion, and Kerr nonlinearities effects. Firstly, to compensate the fiber losses while increasing the communication reach, optical amplifiers (OA) are used. Particularity, in C-band, EDFA are employed to amplify the optical signal at the cost of introducing amplified spontaneous emission (ASE) noise to it [EDFA02]. The ASE noise accumulation can be modeled by additive white Gaussian noise (AWGN), and it is considered LI noise. Regarding the CD, it is the effect associated with the propagation of different wavelengths at different speeds in an optical fiber which depends on the employed fiber type by different chromatic dispersion parameters. Nowadays, the CD is mainly compensated by DBP-based DSP methods at the coherent Rx side, without implementation of in-line DCF, these kinds of systems are referred as dispersion uncompensated optical systems [Po12]. Finally, the Kerr nonlinearities effects are the main source of NLI noise in nowadays optical communications, the two main contributors are the self-phase modulation (SPM) and the cross-phase modulation (XPM) [NFO13]. Consequently, the NLI noise is highly associated with the number of wavelengths in a WDM signal, as well as, the optical power, transmission distance, signal configuration, etc., of each transmitted channel. All those parameters will lead to several signal-nonlinear interaction that will lead to several impacts of the NLI noise in an optical system. Recently, ML-based methods have been proposed to NLI noise mitigation due to their natural capability of understand nonlinear interactions between inputs and outputs.

2.2 Innovative Optical Network Architectures and Transmission Technologies

In this section, we offer the needed background for two innovative optical transmission technologies and network architectures investigated in this PhD thesis. Specifically, in Section 2.2.1, hybrid filterless and filtered optical networks are introduced and, in Section 2.2.2, the main characteristics, advantages and disadvantages of DSCM systems are briefly explained.

2.2.1 Hybrid Filterless and Filtered Optical Networks

To face the expected high demands by emerging of 5G and beyond communications, optical networks must advance and readapt in a cost-effective way. Recently, one of the optical network architectures capturing interest from the telecom industry are the filterless optical networks (FON) [Ay22]. Mainly, FONs are valuable to cost-effectively aggregate the data traffic in metro aggregation network segments, by employing bus/horseshoe network topologies, and then propagate the aggregated traffic to the core network segments, commonly, based on mesh and ring topologies. In FONs, as the name suggest, there is no filter capabilities at the intermediate nodes, usually, the nodes based on optical filters, i.e., WSSs, are placed at the edges of the bus/horseshoe topologies, closer to the core network, to avoid undesired signal propagation beyond the horseshoe and laser loops [Ay22]. Consequently, it leads to hybrid filterless and filtered optical networks. Notice that, from the edge nodes of the bus/horseshoe to a destination node in the core network, the optical signal can transit more than one WSS-based ROADM.

One of the main advantages of the FONs is the cost-effectiveness by using passive optical elements, such as optical splitter/couplers, instead of active ones, like the WSSs. By that replacement, the networks operators are saving not only in CapEx but also in OpEx, because the optical splitter/couplers are cheaper than the WSSs and they have lower power consumption and footprint. Moreover, the network agility as well as the multilayer networks are improved since the intermediate nodes in filterless segments do not require reconfigurability [Ar16].

On the other hand, by avoiding the WSSs in the filterless segments, the wavelength reuse cannot be performed and, consequently, all the wavelength in a WDM signal will be propagated beyond their destination node until reach an optical node with filter capabilities. Additionality, in presence of OA in the filterless segments, the ASE noise will be propagated through all the transmitted band, affecting all the propagated channels. In this way, it is desired that the employed OA at those segments is low-noise OA, which can be challenger since the high-port optical splitter/couplers can lead to high insertion losses.

2.2.2 Digital Subcarrier Multiplexing (DSCM)

DSCM systems have been proposed to increase the overall capacity and flexibility of the optical networks [Su20], [We21], [JSAC21], [Infinera.1]. It is a key technology to provide more freedom in coherent optical networks design and management, offering the possibility to propagate an optical channel with multiple digital independent SC. DSCM shows significant benefit in both point-to-point (P2P) transmission (working at high symbol rates [Su20]) and in P2MP network topologies [We21]. The former, by employing an 8 DSCM signal powered by advanced parallelized DSP blocks and probabilistic constellation shaping (PCS) for 800G and beyond applications. The latter, taking advantage of the multiple independent SC, can offer a finer granularity

in optical network resources assignment and management, from 25G up to 400G, which play an important role in future metro aggregation optical networks where the data traffic mainly follows a hub-and-spoke pattern [Ho22].

Following, we introduce Eqs. (2-2), (2-3) and (2-4) which govern the DSCM systems. The total DSCM signal is defined by [Ra17]

$$S_{DSCM}(t) = \sum_{i=1}^{N} S_i(t).\, e^{j2\pi f_i t} \tag{2-2}$$

where N is the total number of SC and f_i is the frequency shift applied for a given generated signal S_i to create a SC, it is defined by:

$$f_{i=1\ldots N} = \left[(i-1) - \frac{N-1}{2}\right].\Delta f_{SC} \tag{2-3}$$

where Δf_{SC} is the spectral width of one single SC, given by:

$$\Delta f_{SC} = \frac{R_s}{N}\,(1+\beta) \tag{2-4}$$

where R_s is the total symbol rate and β is the roll-off factor of the digital root-raised-cosine filter (RRC) used to optical signal shaping. Notice that, by Eq. (2-3), we considered a SC spectral assignment from low to high frequencies.

Fig. 2-1 illustrate the basic step by step digital process at the optical Tx to create a DSCM signal. Firstly, generation of N pseudo-random binary sequences (PRBS) mapped by quadrature amplitude modulation (QAM) formats and shaped by RRC filters is performed. Then, applying the frequency shift f_i computed by Eq. (2-3) and, finally, combining all the SC by means of an optical multiplexer, creating the DSCM signal, S_{DSCM}, to be propagated through the optical fiber.

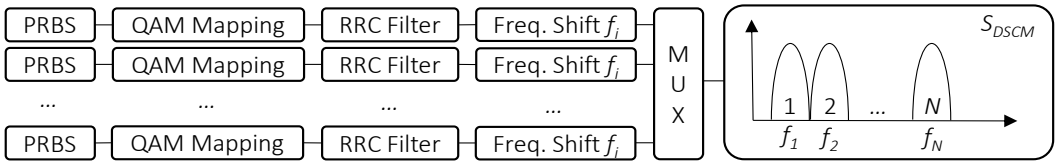


Fig. 2-1. DSCM signal generation process at the optical coherent Tx.

DSCM signals are characterized to be more robust against the NLI noise because, normally, they are working inside the NLI "sweet spot" in term of symbol rates [Infinera.1], i.e., from 4GBd to 16GBd per SC. However, the impact of the NLI noise is lower, leading to mild to low nonlinear optical transmission regimes, other PLI, such as optical filtering plays a critical role in these kinds of optical systems by the fact that channel narrowing is affecting more the SC placed at the edges of the optical channel, so-called the external SC, than the internal SC [Ra16].

2.3 Optical Constellations (OC)

In this section, we concentrate on OC and its main characteristics. OC are one of the most common ways to represent an optical signal modulated by a digital modulation scheme, such as *M*-QAM signals. Typically, at the optical Tx, the optical signal is generated by using the optical laser and an in-phase/quadrature (IQ) modulator [Xi17], [Se19]. The IQ modulator is responsible to modulate the waveform generated by the optical laser in accordance with the input data bits to be transmitted, following a constellation mapping. The main function of the mapping is to assign the proper amplitude levels corresponding with the I and Q components of the respective symbols. For that purpose, the input sequence of bits to be transmitted is grouped by a number of bits creating the symbols. The number of available symbols is given by the modulation order M and the number of bits transmitted by each symbol is given by the $\log_2 M$. For example, for 16QAM signals, they are characterized by 16 symbols and each one carrying 4 bits. Normally, the mapped symbols are represented by complex numbers, where the real part corresponds to the I component and the imaginary part to the Q component of the optical signal. Thus, an optical signal is represented by a sequence of complex numbers, representing the respective symbols. Assuming an ideal Tx and the Gray mapping, Table 2-I and Table 2-II represents an example of the mapping for the quadrature phase shift keying (QPSK)/4QAM and 16QAM signals, respectively. Fig. 2-2 exemplifies the ideal IQ OC generated at optical Tx side to be propagated for QPSK, 8QAM, 16QAM, 32QAM and 64QAM signals in Fig. 2-2a), b), c), d) and e), respectively.

Table 2-I. QPSK mapping.

Symbol	Bits	Mapped symbol
1	00	$-1-i$
2	01	$-1+i$
3	11	$1+i$
4	10	$1-i$

Table 2-II. 16QAM mapping.

Symbol	Bits	Mapped symbol	Symbol	Bits	Mapped symbol
1	0000	$-3-3i$	9	1000	$3-3i$
2	0001	$-3-i$	10	1001	$3-i$
3	0010	$-3+3i$	11	1010	$3+3i$
4	0011	$-3+i$	12	1011	$3+i$

5	0100	$-1-3i$	13	1100	$1-3i$
6	0101	$-1-i$	14	1101	$1-i$
7	0110	$-1+3i$	15	1110	$1+3i$
8	0111	$-1+i$	16	1111	$1+i$

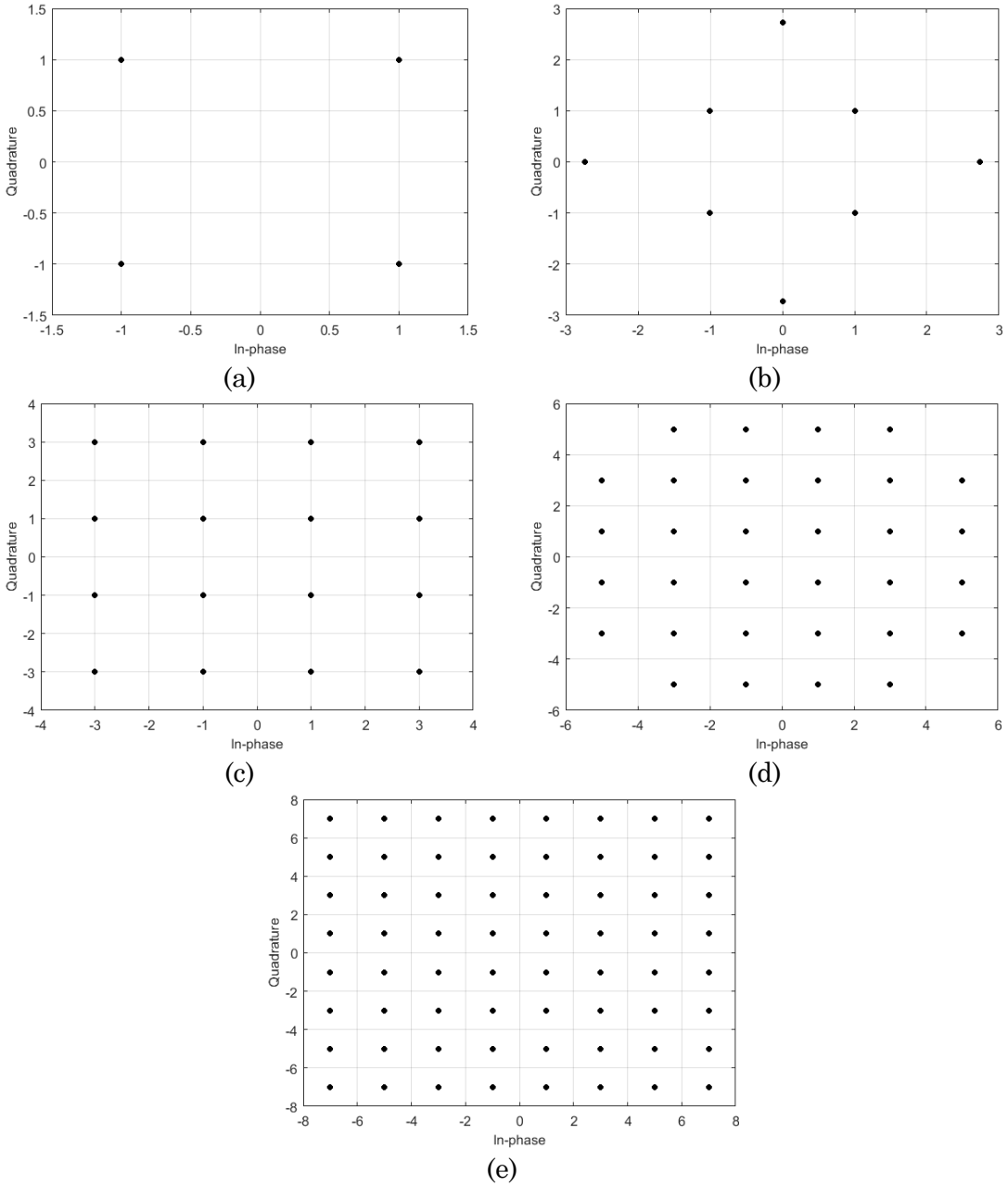
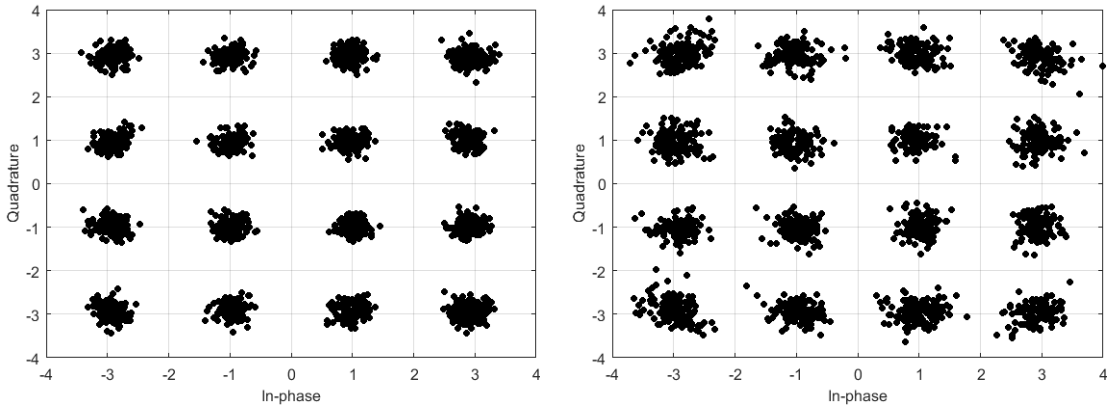


Fig. 2-2. Ideal OC generated at optical Tx for QPSK (a), 8QAM (b), 16QAM (c), 32QAM (d) and 64QAM (e).

Furthermore, as the IQ OC are a representation of the optical signal, they are a valuable source of information regarding the several impacts on the signal along its lightpath, both the LI and the NLI noise. For example, it is well known that the optical signal amplification by EDFAs, will introduces ASE noise to the optical signal, that effect is normally visible in OC and can be investigated to extract useful information regarding the OA in an optical line system, since the ASE noise added to the optical signal is directly dependent on the gain and noise figure of the OA. Also, the nonlinear Kerr effects, arising from the optical fiber propagation, lead to more impacts on the optical signal and, consequently, on the OC. It is known that the NLI noise is harmful in high power constellation symbols, i.e., higher amplitude levels. Thus, by a deep analysis of the OC, we can extract valuable knowledge from it for several use cases, such as optical performance monitoring (OPM) and lightpath characterization, from fiber type to OA gain and noise figure.

For better understanding of the several effects on the IQ OC, Fig. 2-3 shows an example of simulated 16QAM OC considering the same optical power and after 400km of different types of fiber, standard SMF (SSMF) and large effective area fiber (LEAF), in Fig. 2-3a) and Fig. 2-3b), respectively. Clearly, we can observe different impacts by different types of fiber. One of the main goals of this PhD thesis is investigate and extract useful knowledge from the IQ OC to lightpath modeling, analysis, and characterization.



(a) (b)

Fig. 2-3. Simulated 16QAM OC after 400km of SSMF (a) and LEAF (b).

2.4 **Machine Learning (ML)**

In this section, we concentrate on the background regarding the ML-based methods considered throughout this PhD thesis. Particularly, in Subsection 2.4.1, we provide background on ML-based to feature engineering, specifically, GMMs and its main characteristics. In Subsection 2.4.2, the feedforward ANN, as well as its concatenation creating a kind of DNN are explained.

2.4.1 Feature Engineering

Generally, feature engineering is the process of using the knowledge from the raw data to extract useful features from it in order to develop most accurate ML-based algorithms and to better and deeper analysis of the available data. For example, data clustering is a kind of feature engineering organizing the raw data by clusters or groups that, in theory, have similar properties/features.

Differently from the *k*-means, a hard-clustering algorithm, GMMs [MMA20] are soft-clustering probabilistic models, that describe each cluster in a dataset by a Gaussian distribution characterized by its mean and variance. Notice that, when we are working two independent random variables, GMMs output bivariate Gaussian distributions characterized by its mean, variance, and covariance. The covariance focuses on the relationship between the deviation of both variables instead of the deviation from the mean of each variable. Consequently, after applying the GMM to a given dataset, we can do deeper analysis of the features extracted, i.e., mean, variance and covariance, to better understand the dataset and how we can perform ML-based model development using the extracted features. In this PhD thesis, we focus on GMM to characterize and extract useful features from the IQ OC.

2.4.2 Artificial Neural Networks (ANN)

ML-based algorithms are considered a strong and interdisciplinary tool and, recently, they are applied to the optical communications as well to turn them more intelligent systems [Ra18]. ANN are information processing systems that simulate the biological behavior of neural networks [Ja96], and, naturally, they can learn nonlinear interaction between network inputs and outputs. They are characterized by the input layer, following a given number of hidden layers and, finally, an output layer. Each layer has a number of neurons and based on their architecture they can be grouped into two main groups, feedforward and recurrent ANN. Feedforward ANN are characterized by non-loops between the layers or neurons, just propagating forward the inputs through the multiple hidden layers until the output layer. Fig. 2-4 illustrates an example of the feedforward ANN with the respective layers and the neurons represented by the circles.

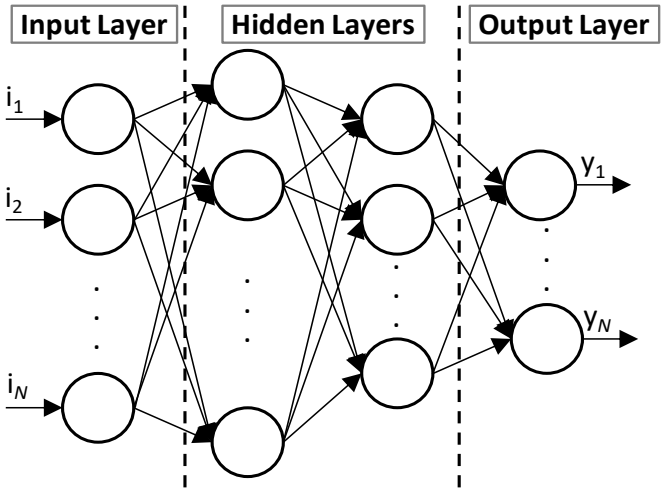


Fig. 2-4. General structure of a feedforward ANN.

In this PhD thesis, we focus on the application of feedforward ANN for several use-cases applied to the optical communications networks modeling and analysis. Additionality, we investigate the methodology of concatenating different feedforward ANN to create an end-to-end (E2E) model of a given lightpath. The idea of ANN concatenation is shown in Fig. 2-5 and it is a kind of DNN. As we can observe, the main characteristic is that the outputs of one ANN will be the inputs of the follow concatenated ANN. Notice that, individual ANN models are trained separately and then they are concatenated as desired.

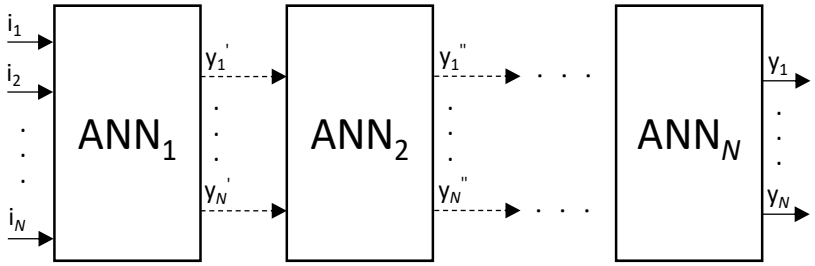


Fig. 2-5. ANN concatenation.

2.5 Conclusions

The purpose of this chapter has been offered the needed background about optical communications, innovative optical transmission technologies and networks, OC, as well as the ML-based methods investigated throughout this PhD thesis.

Nowadays, optical communications are paramount to transport all data traffic. Optical technologies such as WDM, EON, optical coherent receivers, advanced DSP blocks, etc., are extensively deployed in current optical communication systems and

networks. Consequently, continuous investigations in these complex infrastructures need to be considered to overall improvements and face the next traffic demands in a cost and energy-efficient way.

Innovative optical transmission technologies and networks have been proposed to increase the overall optical network performance. In hybrid filterless and filtered optical networks, trade-off investigations between network costs and performance should be carefully carried out in order to understand the full benefits of FON segments, as well as investigations of the PLI impact in these kinds of networks. Additionally, DSCM is a promise transmission technology to bring more freedom to the optical networks management while reducing the overall costs.

OC are an important representation of the optical signals and can be a useful source of information to understand the several PLI impact on the signals. In highly flexible optical networks, there are several signal-to-noise interactions and optical parameters to be analyzed and monitored. A deeper and careful analysis of the OC can give valuable knowledge about the PLI effects in optical signals.

Finally, ML-based methods to be used in this PhD thesis were explained. Specifically, background on GMM to feature engineering by soft-clustering and the idea of ANN concatenation was provided.

Chapter 3

Review of the State-of-the-Art

This chapter aims to review the state-of-the-art related with the technologies investigated in this PhD thesis. Section 3.1 reviews the state-of-the-art regarding the optical filtering impact in optical networks. In Section 3.2, the state-of-the-art of MF and SR identification is introduced. Section 3.3 focuses on the state-of-the-art of DT and ML-based techniques applied to optical communications. Finally, Section 3.4 concludes the revision of the state-of-the-art.

3.1 Optical Filtering Impact in Optical Networks

The commercial WSSs availability [Finisar], [Lumentum] enables the deployment of EON [EON16], together with the support of S-BVTs [Sa15] – operating at SR beyond 32 GBd and high-order MF, e.g., 64QAM [Za19.1], [Za19.2], [Pe20]. In addition, this led to an increasing of the SE and allows to extend the optical layer towards the network edge [Ve13]. Next generation of TRx might be soon capable of generating and transmitting optical signals based on SR as high as 100 GBd with MF employing high-order QAM formats enhanced by PCS, DSCM, as well as dynamic bandwidth allocation (DBA) and other innovative capabilities [Infinera.2]. Based on current commercially available WSSs, the flexible grid concept is possible, and a frequency granularity of 12.5 GHz is in line with the standardizations [ITU-T G.694.1], leading to several options for what concerns spectral assignments. Furthermore, in the future, to further improvements in SE, a spectrum management based on a tighter frequency granularity of 6.25 GHz is expected, offering an even wider range of optical channels configurations [Kh15]. Notice that commercially available WSS already supports frequency granularities of 6.25 GHz and even tighter ones [Finisar].

Therefore, as the frequency granularity to spectral assignment is becoming tighter, the optical filtering performed inside the WSS can lead to higher QoT degradation.

For example, due narrowing filter effects [Fa16], if the trade-off between channel spacing, MF and SR assignments is not carefully considered. Investigations about the impact of the optical filtering in optical networks can be found in the literature. The authors in [Za19.1] studied the impact of optical filtering assuming 75 GHz channel spacing on optical signals that pass through several ROADMs, owing to the fact of the PLIs in these scenarios, lower order MF must be used. In [Se18], the authors investigated the optical filtering impact in 100G QPSK signals for 50 GHz optical channels and CDC ROADM-based optical networks, leading to negligible OSNR penalties due to the optical filtering for reasonable optical reaches. In addition, in the literature there are some works proposing the detection, identification and localization of failures related with the optical filtering. For example, the authors in [Sh19], taking advantage of optical spectrum analyzers (OSA), proposed ML-based methods to detect, identify and localize optical filter-related soft-failures, such as optical filter shifts. Most recently, in [Ca22], a fusion model to optical filtering impairments monitoring was proposed.

In this PhD thesis, we concentrate our investigations on hybrid filterless and filtered optical metro networks scenarios and we investigate the optical filtering impact considering 32 GBd and 64 GBd, 32QAM and 64 QAM formats and channels spacing from 37.5 GHz to 112.5 GHz with a frequency granularity of 12.5 GHz. The main objective, in line with the thesis goal G.1, is to find the minimum channel spacing to avoid higher OSNR penalties due to the optical filtering in these types of networks, while improving the SE and reducing the number of needed TRx, consequently reducing the overall costs.

3.2 Modulation Format and Symbol Rate Identification

Digital coherent optical transmission, enabled by programmable TRx alongside with advanced DSP algorithms, is nowadays prevalent in optical networks [Sa15], [Zh19]. It not only increases the spectral efficiency and the flexibility, as well as it also simplifies dynamic network operation [Lo20]. Advanced DSP algorithms improve the overall system performance, e.g., by NLI equalization, and they can extract useful information from the received signals. Furthermore, coherent S-BVT enable to configure optical signals with different MF and SR to deal with different traffic demand scenarios. It is worth noting that, by optimizing the MF and SR, large energy efficiency can be achieved, as reported in [Fr17].

In the simplest scenario, at the optical connection set-up time, the centralized SDN controller finds the route, the spectrum allocation, and the TRx selecting the best configuration within the Tx and Rx [Da15]. A centralized MDA system [Ve19] can be also in charge of collecting monitoring data from the different nodes analyzing the data and work together with the SDN controller to perform the needed tuning of

parameters on the devices [Ve18], [Gi18]. Nonetheless, due to the low latency required, e.g., in 5G and beyond services, optical connectivity of the access to datacenters located in metro/core segments is needed; such connectivity imposes high dynamicity to provide the required capacity and entails the coordination among SDN controllers of the different network segments involved. To relieve the SDN controller and other centralized elements from real-time operation, BVTs can make local decisions autonomously, e.g., change the signal's configuration in terms of MF and SR to adapt to the actual needs [Sp19], [JSAC21]. Although these solutions are highly energy efficient, they require a more complex Rx capable of identifying the configuration of the received optical signal.

In the literature, several works have proposed different approaches for the MF identification and some of them also include the identification of the SR (see e.g., [Sa20]), as well as OSNR estimation [Xi19]. In particular, most of the previous works use ML-based techniques [Ra18], either supervised, which entails a training phase, or unsupervised, e.g., using clustering algorithms [Ja99]. The authors in [Ch19] presented a DNN for MF and SR identification considering up to 16QAM and slower SRs. In [Ch21], the authors proposed a method based on random forest to MF identification but also focusing on slower SRs, e.g., up to 16 GBd. As the application of DNNs might increase the complexity of the Rx, the authors in [Ta19] proposed a blind MF identification method based on phase distribution and average amplitude ratio analysis. Finally, the authors in [Zha19] proposed a MF identification algorithm based on clustering for direct detection systems.

Even though some approaches achieve 100% accuracy in the identification of the MF and SR, they do not show that such level of accuracy can be obtained in higher nonlinear optical transmission scenarios, e.g., in the presence of NLI noise on systems working with high-order MFs and high SRs. In this PhD thesis, we propose an accurate and low-complexity algorithm for MF and SR identification. The performance of the proposed algorithm is evaluated in realistic lightpath scenarios, including LI and NLI noise, working with high-order MF and high SR, as well as in single and multicarrier optical systems.

3.3 Digital Twin and Machine Learning for Optical Communications

Recently, DT solutions for optical communications networks and systems have been proposed [Wa21], [Vi22], [Ja22]. DTs are basically taking advantage of the data from the optical physical layer, i.e., network elements configuration and monitoring data from MDA agents, extracting valuable knowledge from it to create digital models in order to support the physical layer operations. For example, it can be useful for TRx configuration by selecting the best available configuration in terms of MF and SR. In addition, ML-based methods are important in DT to process most efficiently the

data acquire from the optical physical layer and extract useful knowledge to apply in the digital models. ML-based methods are widely investigated in several fields, as well as in optical communications [Ra18], [Ro18].

In the literature, there are plenty of works where AI/ML-based models are proposed to improve the overall performance of the optical transmission networks. For example, in [Fr21] the authors investigated the implementation of ANN to NLI noise equalization in order to improve the optical networks performance. In [JSAC21], we proposed an autonomous lightpath operation based on DSCM technology showing significant improvements in both power and spectrum savings, enabled by reinforcement learning algorithms to traffic prediction and then configure most properly the DSCM signal. ML-based approaches are also useful to optical networks parameters prediction, for example, it is useful for soft-failures prediction, identification, and localization, as reported in [Sh19] and [Ba21]. Moreover, nowadays, most common is the use of the ML-based models to QoT estimation in optical networks [Ro18], [Mo18], [Po21], because they can perform accurately at the cost of low computational requirements. Differently, the well-known analytical model SSFM [Sh14] provides higher accuracy but requires high computational requirements, which limits it utilization for real-time operations and in highly dynamic optical network scenarios with high number of optical channels. Another interesting application of ML-based methods is to analyze OC diagrams, for example, it can be useful to MF and SR identification, as it was explained in previous Section 3.2. For example, in [Wa17], several ML-based methods are investigated to MF identification and OSNR estimation, concluding that convolutional neural networks lead to the best performance.

3.4 Conclusions

This chapter focused on the state-of-the-art related with the main objectives of this PhD thesis.

Regarding the optical filtering impact in optical networks, associated with the thesis goal G.1, the investigations carried out are mainly focus on long reaches optical networks, based on advanced ROADM architecture, since in these scenarios the optical signal crosses a higher number of nodes. Consequently, due to the high PLI in these network scenarios, low order MF and low SR must be considered to avoid higher QoT degradation due to the optical filtering. On the other hand, applying hybrid filterless and filtered approaches to the optical networks, to cost-effectively spread the optical technologies to the metro and access networks segments, requires investigations about the optical filtering impact in these scenarios, and more interesting understand the possibility to propagate optical signals based of high-order MF and high SR to overall SE improvements.

Autonomous identification of MF and SR at optical Rx is paramount to the autonomous lightpath operation. With the expected availability of several optical signal configurations at optical Tx to face all type of data traffic, the capacity of the Rx identify the MF and SR of the incoming optical signals will greatly improve the autonomous lightpath operation. Even more if the Tx can autonomously adapt its configuration for the actual data traffic needs. In that regard, several works proposed different methods to MF and SR identification at Rx, mainly based on pre-trained ML-based models. Most of the proposed methods show good accuracy but mainly for low order MF, i.e., 4 and 16QAM, and low SR, i.e., up to 16 GBd. By the foreseen use of high-order MF and high SR in future optical networks, new accurate and low-complexity methods to MF and SR identification are essential.

Recently, DT approaches have been proposed for the optical communications. There are several use-cases where DT can improve the overall optical communications performance, from lightpath provisioning to QoT estimation. In fact, accurate and low-complexity optical transmission and networking DT can lead to significant improvements in optical network management and, consequently, in OpEx. In addition, a better use of the optical networks' resources enabled by accurate DT, lead also to more cost and energy-efficient optical networks. Although nowadays, while the paradigm of disaggregated optical networks is under considerations, optical networks mainly follow a partially disaggregated approach, leading to more challenges to create accurate DT for the optical network. These challenges can be addressed by use ML-based models, one of the pillars of the DT approaches.

Chapter 4

Filtering Impact on High-Order Modulation Formats and High Symbol Rates in Hybrid Filterless and Filtered Optical Metro Networks

In this chapter, a comprehensive study on the impact of optical filtering in hybrid filterless and filtered optical metro networks is carried out through MC-based simulation. The goal of this work is to determine the minimum channel spacing to avoid optical filtering penalties. We consider symbol rates of 32 and 64 GBd with 32 and 64QAM in channel spacing scenarios from 37.5 GHz to 81.25 GHz with 12.5 GHz of granularity. The results show an OSNR penalty higher than 1 dB – with respect to back-to-back (B2B) – due to the optical filtering only for 64QAM format in the worst-case studied, i.e., after 10 cascaded ROADMs, which might enable its application for future metro optical networks.

4.1 Hybrid Filterless and Filtered Metro Networks Scenarios and Case-Study

Although WSSs, inside of ROADMs, play an important role to transparently route an optical signal along its lightpath, they impact network performance due to filter cascading effects (optical bandwidth narrowing entails OSNR penalties), as well as being a point of failure that causes QoT degradation [Sh19]. The narrow filtering occurs because of the non-ideal transfer function of the WSS, which results to be narrower than the nominal selected value in commercial WSS [Ra14]. Further

degradation can be caused by the frequency shift between WSS and channel, as well as by WSS ageing effects.

One important feature for 5G and beyond transport networks, alongside with higher capacities, is that of low latency for connectivity services between the access and the services running in metro and core data centers (DCs). In that regard, filterless segments (e.g., in the form of horseshoes) can used to extend the filtered metro network to reach the access one [Pa20], [Sh20]. Fig. 4-1 represents the targeted scenario with the filtered mesh metro network topology implemented with ROADMs and a filterless horseshoe topology to extend the optical network to the edges with filterless nodes based on optical splitters/couplers (and OAs if needed). The internal architecture of each node is shown in the figure. Additionally, let us assume that metro DCs can be collocated with the ROADMs.

We are interested in studying the QoT experienced by signals connecting a node in the horseshoe, typically an aggregation node connecting one or more access networks (including radio and fixed access), and a metro DC collocated with one of the ROADMs.

Therefore, in the case that a metro DC is collocated to one of the end ROADMs in the local horseshoe (ROADM-1A/B in Fig. 4-1), a lightpath connecting a filterless node (R1..4 in Fig. 4-1) to that DC1 would only traverse one ROADM, i.e., two WSSs. However, since not all metro nodes could be equipped with computation capabilities, the lightpath might traverse additional ROADMs in their route.

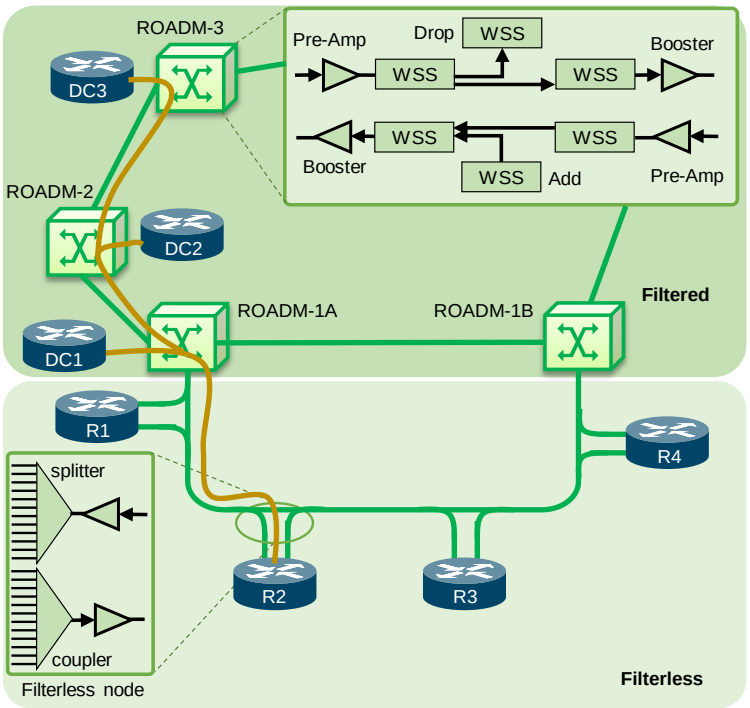


Fig. 4-1. Hybrid filterless and filtered-based optical metro network.

4.2 Simulation Setup

To evaluate the system performance in those scenarios, we carried out extensive MC-based simulations. At the Tx side, we generated 2^{13} PRBS shaped by a RRC filter with 0.15 roll-off factor considering the SRs and MFs defined in the previous section. The optical filtering impact is investigated on two different scenarios that corresponds to the case where the DC is collocated with the ROADM in the local horseshoe or with a different metro ROADM, denoted as Scenario 1 and 10 depending on the number of ROADMs that the signal traverses, e.g., Scenario 1: R2-DC1 and Scenario 10: R2-DC10, where the signal crosses 10 ROADMs.

To model the WSS spectral shape, we averaged out real spectral measurements from 4 different optical channels for each considered channel spacing [Ra14]. Fig. 4-2 depicts the respective transfer function for the 50, 87.5 and 112.5 GHz channels the −3dB bandwidth is about 46.3, 83.6 and 110.2 GHz, respectively. Note that, as the filter transfer functions were obtained by real spectral measures, they also include filter ripples, which accumulate with the cascaded filters, and filter insertion losses.

Regarding the optical amplification, all OAs are modeled as EDFAs with a noise figure of 6 dB. In this chapter, we consider a filterless segment with 40 km, a filtered one with 100 km and a fiber attenuation factor of 0.25 dB/km.

Finally, we assume an ideal optical coherent receiver and direct error counting to estimate the BER, considering a total of 1000 errors as threshold and a target pre-FEC BER of 4×10^{-3}.

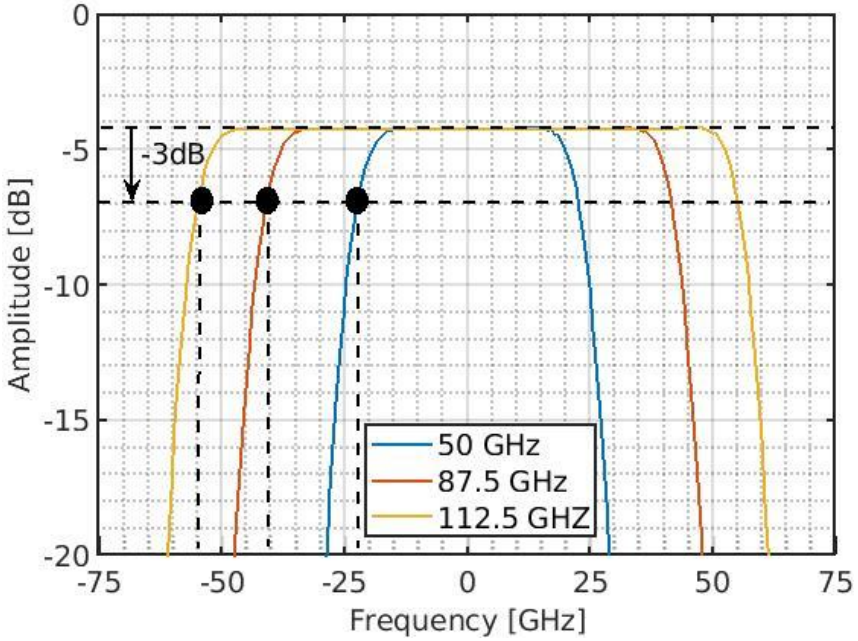


Fig. 4-2. Passband filter transfer functions for 50 GHz, 87.5 GHz, and 112.5 GHz optical channel spacing.

4.3 **Illustrative Simulation Results**

In this section, we analyze the OSNR penalty due to optical filtering for different channels spacing and for the 2 scenarios presented in the previous section, i.e., when an optical signal crosses 1 or 10 ROADM from source to destination. To analyze the OSNR penalties due to the optical filtering in the two scenarios, we considered 37.5 and 50 GHz channel spacing for 32 GBd signals, 75 and 87.5 GHz for 64 GBd ones and 100 and 112.5 GHz for 96 GBd signals. Fig. 4-3 depicts the BER as a function of the OSNR for the two scenarios, and for the B2B one, which does not suffer from optical filtering impairments. The required OSNR obtained for the threshold pre-FEC BER is about 22 and 26.5 dB for 32 GBd signals, 25.5 and 29.5 dB for 64 GBd, and 26.7 and 31 dB for 96 GBd, all cases for 32 and 64QAM, respectively. Fig. 4-3a) plots the results for 32 GBd signals in a 50 GHz optical channel, Fig. 4-3b) for 64 GBd in an 87.5 GHz, and Fig. 4-3c) for 96 GBd in a 112.5 GHz, in all for both 32 and 64QAM.

We can observe that, the channels spacing considered reduce the optical filtering penalties to just less than 0.05 dB for almost all investigated scenarios, i.e., practically negligible penalties. The only scenarios where some filtering penalties are observed was that for 64QAM 32/96 GBd signals in 50/112.5 GHz channel spacing after 10 ROADMs, with penalties about 1 and 0.3 dB, respectively. Notice that, for the other channels spacing considered, the two tighter ones for 32 and 96 GBd signals (37.5 and 100 GHz) impose large filtering penalties (> 4 dB) for Scenario 1 and 32QAM signals. For 64 GBd signals in the tighter channels (75 GHz), we obtained penalties lower than 1 dB (~0.7 dB) for Scenario 1 and both MFs.

Table 4-I summarizes the SE as a function of the MF and SR considering a FEC overhead of 20% and 25% for 32 and 64QAM, respectively. We observe that the SE increases when increasing of MF and SR, achieving the better one for 96 GBd signals, where the use of 112.5 GHz channel spacing results in SE about 6.83 and 7.68 b/s/Hz for 32 and 64QAM, respectively.

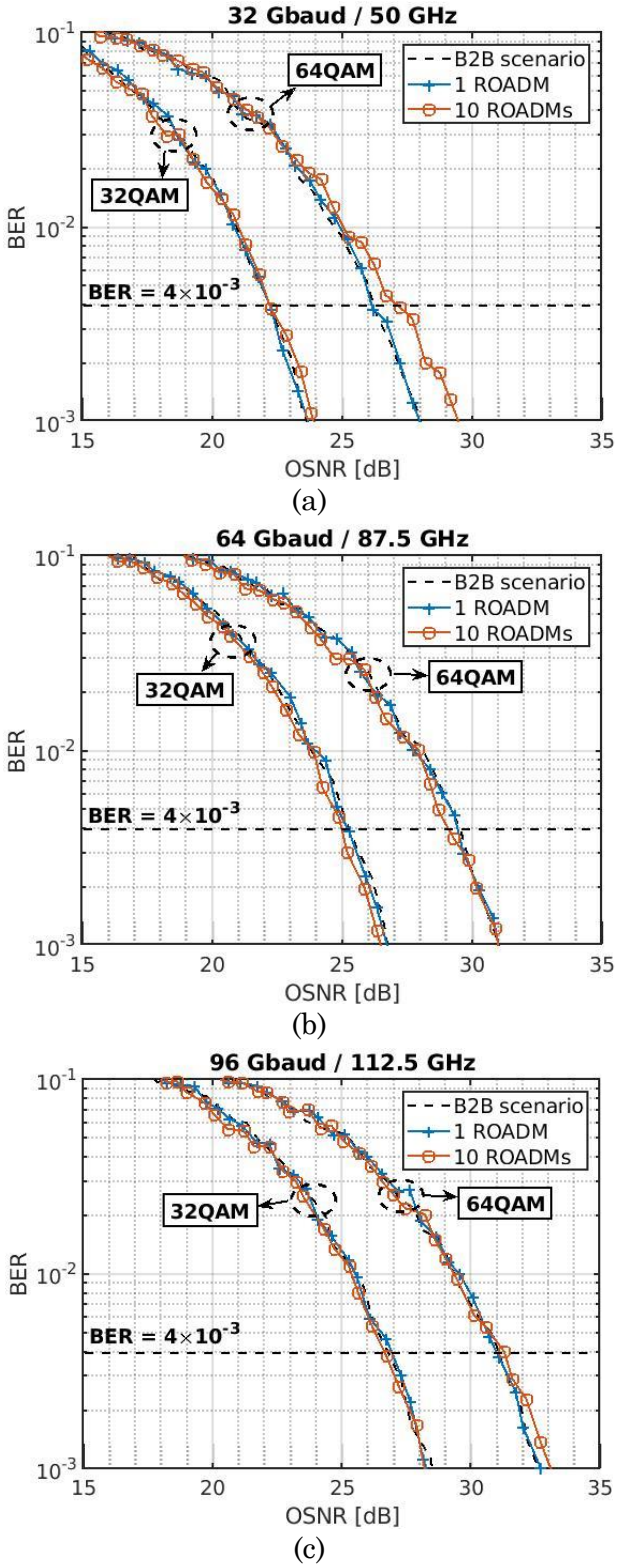


Fig. 4-3. BER as a function of the OSNR for 32 GBd signals in 50 GHz channel spacing (a), 64 GBd in 87.5 GHz (b) and 96 GBd in 112.5 GHz (c).

Table 4-I. SE as a function of the MF, SR and channel spacing.

MF	SR [GBd]	Ch. [GHz]	Bit rate [Gb/s]	Net bit rate [Gb/s]	SE [b/s/Hz]
32QAM	32	50	320	256	5.12
	64	87.5	640	512	5.85
	96	112.5	960	768	6.83
64QAM	32	50	384	288	5.76
	64	87.5	768	576	6.58
	96	112.5	1152	864	7.68

4.4 Concluding Remarks

The impact of the optical filtering in hybrid filterless and filtered optical metro networks was investigated for high-order MF, high SR, and several optical channels spacing. Two scenarios where selected to capture filtering effects, where in scenario *i* a lightpath crossed *i* (1 or 10) ROADMs from source to destination.

The results showed that optical filtering leads to negligible OSNR penalties, lower than 1 dB, for 32 GBd signals in 50 GHz channels, 64 GBd in 87.5 GHz and 96 GBd in 112.5 GHz. Consequently, we foresee the use of such SRs and MFs to increase the transport capacity and improve the overall SE in future optical metro networks. The availability of such SRs and MFs will allow reducing the number of TRx, thus reducing not only the capital but also operational expenses for network operators.

We also foresee the use of a frequency granularity of 6.25 GHz to increase even more the spectral efficiency, as well as the use of fast symbol rates and high-order modulation formats to increase the transport capacity and provide low latency to different applications. For example, the number of channels in a full C-band can be increased from 96 to 109 using 43.75 GHz channel spacing instead of 50 GHz, a trade-off investigation between PLI and capacity improvements should be carried out.

In view of the feasibility of the type of scenarios considered in this chapter, faster symbol rates, up to 120 GBd, and 128QAM modulation format will be considered as future work. Additionally, the implication of filter-related failures, such as filter shift and filter tightening in this kind of networks should be investigated.

Chapter 5

Accurate Low Complex Modulation Format and Symbol Rate Identification for Autonomous Lightpath Operation

Network automation promises to reduce costs while guaranteeing the required performance. This is paramount for the highly dynamic traffic expected soon. In optical networks, the autonomous lightpath operation entails that the optical Rx can identify the configuration of a received optical signal without being specifically configured from the network controller. This allows relieving the network controller from real-time operation, and it can simplify the operation in multi-domain scenarios, where an optical connection spans more than one domain. In this chapter, we propose a blind and low complex MF and SR identification algorithm. The algorithm is based on studying the effects of decoding an optical signal with different MFs and SRs. Extensive simulation results, considering a coherent wavelength division multiplexed system based on high-order MF and high-speed SRs, show remarkable identification accuracy in the presence of NLI noise for a wide range of feasible configurations.

5.1 Introduction

In this chapter, we propose a novel algorithm that can recognize the configuration of the received optical signal, in terms of MF and SR, by analyzing the received IQ OC. Its low complexity and simple operation make it ideal for real-time applications and it could be implemented directly in the in-operation TRx. Because high-order MFs and high-speed SRs are more affected by NLIs, the proposed algorithm has been tested under realistic scenarios considering signal configurations up to 64QAM and

96 GBd, both SSMF and LEAF types, and BVTs supporting single and multiple carrier configurations.

The reminder of the chapter is organized as follows. Section 5.2 introduces the rationale behind the proposed identification method with an example. Section 5.3 studies the effects of decoding the received signal with different SRs. An algorithm that exploits the main observation is proposed to identify the configuration of the received signal. The results of the algorithm evaluation carried out by simulation are presented in Section 5.4 for different signal configurations and scenarios. Specifically, the proposed recognition algorithm is tested in both single carrier and DSCM scenarios, in Subsection 5.4.1 and 5.4.2, respectively. Finally, Section 5.5 draws the main conclusions of this chapter.

5.2 MF and SR Identification based on the Analysis of the Received Constellation

Firstly, we analyze the characteristics of an OC when it is received after the optical transmission, i.e., after the mux/demux filters, OAs, and fiber spans. Fig. 5-1 shows several examples of constellations of a 64QAM/64GBd signal received after different number of spans; SSMF spans in Fig. 5-1a) and LEAF spans in Fig. 5-1b). The effect of NLI noise (specifically, SPM and XPM) can be clearly observed, since the received symbols not only are more disperse around the *expected constellation point*, as it happens due to the ASE noise, but also their typical round shape turns into elliptical. This can be clearly noticed in the outer constellation points, higher-power symbols, e.g., $7+7i$, as opposite to inner ones, e.g., $1+1i$. Such NLI noise effects make the identification of the MF more challenging.

However, before the identification of the MF can be carried out, one must estimate the SR. Fig. 5-1c) illustrates the effects of decoding signals with different SRs; for the sake of clarity, only the decoded symbols for the constellation point ($7+7i$) are shown. Each row in Fig. 5-1c) shows the symbols when they are decoded with a particular SR ∈ {32, 64, 96} GBd, whereas the columns show the effect of distance in terms of number of SSMF spans. We observe that the symbols were more dispersed when the SR used for decoding the signal at the Rx (SR_{Rx}) was different from the one used for generating the signal at the Tx side (SR_{Tx}). Such dispersion can be quantified by computing the Euclidean distance of every received symbol with respect to the expected constellation point (we call it as *centroid* because of its similarity to clustering algorithms). The average Euclidean distance $\bar{d}_2(S_{SR}, C_{MF})$ of a set of symbols S decoded with a given SR, with respect to the set of expected centroids C for a given MF, can be computed as:

$$\bar{d}_2(S_{SR}, C_{MF}) = \frac{1}{|S_{SR}|} \sum_{s \in S_{SR}} \min_{c \in C_{MF}} (\|s - c\|_2) \quad (5\text{-}1)$$

The average Euclidean distance can be used to determine the correct SR at the Rx side without previous knowledge, as it will be detailed in the next section.

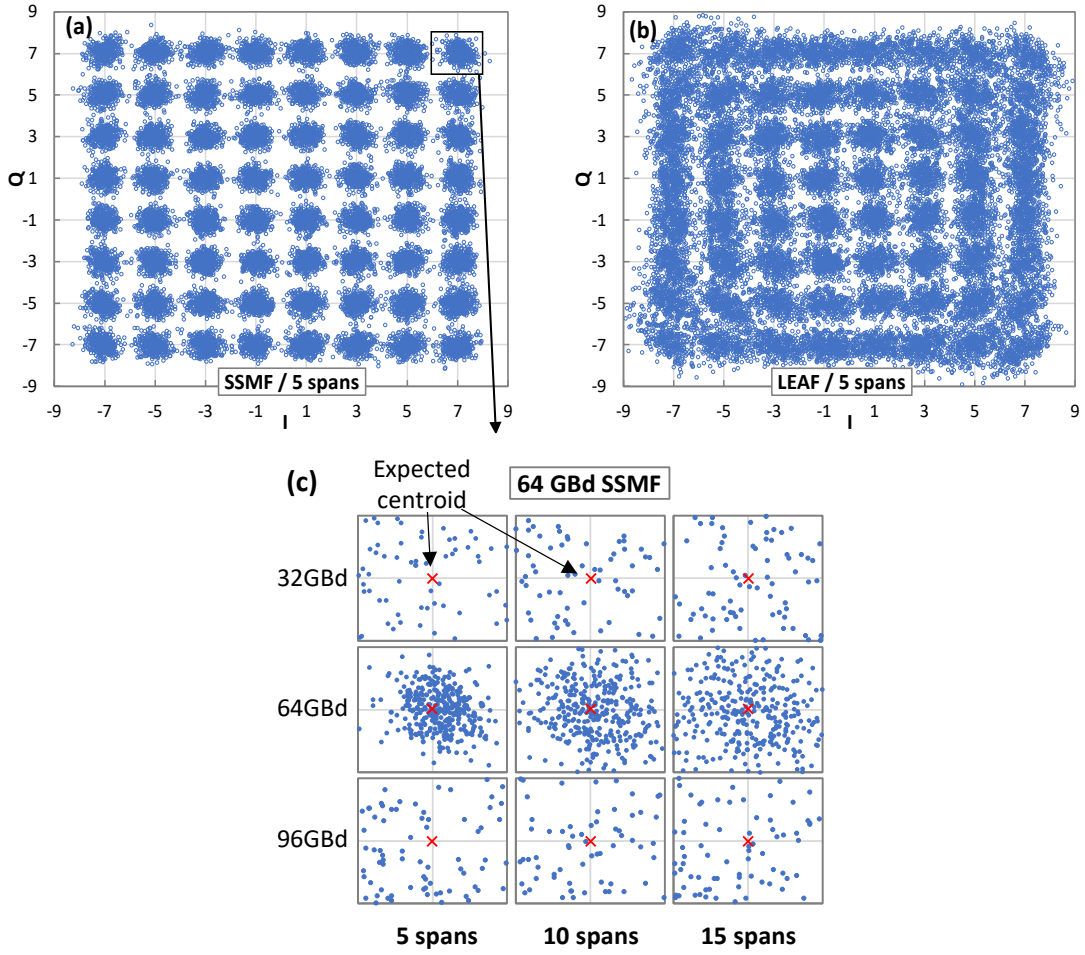


Fig. 5-1. Received 64QAM constellations after 5 SSMF (a) and LEAF (b) spans. Symbols decoded with several SRs and for several distances (c).

5.3 Decoding Signals with Different Symbol Rate

In this section, we analyze the effects of decoding an incoming optical signal at the Rx with a SR_{Rx} different from the one used at the Tx. For the sake of clarity in the discussion, let us assume a B2B scenario, i.e., without any noise impairments.

The first effect is related to the matched digital filter applied at the optical Rx. In particular, state-of-the-art coherent optical communications systems use RRC filters for Nyquist - WDM spectral shaping [Wa13]; by using a RRC filter at the Tx and the matched RRC filter at the Rx, inter-symbolic interference (ISI) is significantly reduced, and the OSNR is maximized. The frequency response of the RRC filter is $H_{RRC}(f|\beta, T_s)$, where β is the roll-off factor and $T_s = SR^{-1}$ is the symbol period [DC01]. Thus, assuming ideal matched filters and $SR_{RX} = SR_{TX}$, then $H_{RRC_{Tx}}(f|\beta, T_s) = H_{RRC_{Rx}}(f|\beta, T_s)$, and the power spectral density (PSD) of the output optical signal, $S_{out}(f)$, is:

$$S_{out}(f) = S_{in}(f) \cdot H_{RRC}(f|\beta, T_s)^2 = S_{in}(f) \cdot H_{RC}(f|\beta, T_s), \quad (5\text{-}2)$$

where $S_{in}(f)$ is the PSD of the input optical signal and $H_{RC}(f)$ is the frequency response of the raised-cosine (RC) filter.

However, when $SR_{Rx} \neq SR_{Tx}$:

$$S_{out}(f) = S_{in}(f) \cdot H_{RRC}(f|\beta, T_{sTx}) \cdot H_{RRC}(f|\beta, T_{sRx}) \approx S_{in}(f) \cdot H_{RC}(f|\beta, T_{max}) \quad (5\text{-}3)$$

$$T_{max} = max(T_{sTx}, T_{sRx}), \quad (5\text{-}4)$$

which introduces high ISI, especially when $T_{sRx} > T_{sTx}$.

The second effect is related to the sampling step. When $SR_{Rx} \neq SR_{Tx}$, a different symbol period is used at the Rx (i.e., $T_{sRx} \neq T_{sTx}$), which produces misalignments in time, thus leading to higher average Euclidean distances from the received symbols to the expected ones. Special cases are when SR_{Rx} is a divisor of SR_{Tx}, i.e., $SR_{Rx} = 1/a \times SR_{Tx}$, $a \in \mathbb{Z}^+$, e.g., SR_{Tx} = 64 GBd and $SR_{Rx} \in \{8, 16, 32\}$ GBd. In those cases, there will be no misalignment in time and the obtained symbols will be correctly decoded, thus leading to average Euclidean distances equivalent to that obtained when $SR_{Rx} = SR_{Tx}$. However, when SR_{Rx} is a multiple of SR_{Tx}, some of the decoded symbols come from inter-symbol times, which will be observed as increased Euclidean distance. Clearly, in both cases, the number of symbols that will be obtained at the Rx will not match, i.e., some symbols will be lost when $SR_{Rx} < SR_{Tx}$ and more symbols will be obtained when $SR_{Rx} > SR_{Tx}$.

The plots in Fig. 5-2 show the average Euclidean distance in the analyzed B2B scenario as a function of SR_{RX}, for 32, 64, 96 GBd. The combined effect and the two individual ones (different RRC filters and the sampling step) are shown separately for a better understanding of the conducted analysis. Two cases can be separately analyzed: *i*) $SR_{Rx} < SR_{Tx}$ and *ii*) $SR_{Rx} > SR_{Tx}$. When $SR_{Rx} < SR_{Tx}$, the RRC filter at the receiver is too narrow and ISI appears; consequently, the decoded symbols are far from their original position on the constellation and the Euclidean distance to the closest centroids increases. On the contrary, if $SR_{Rx} > SR_{Tx}$ the RRC filter is wider, so the different sampling effect dominates.

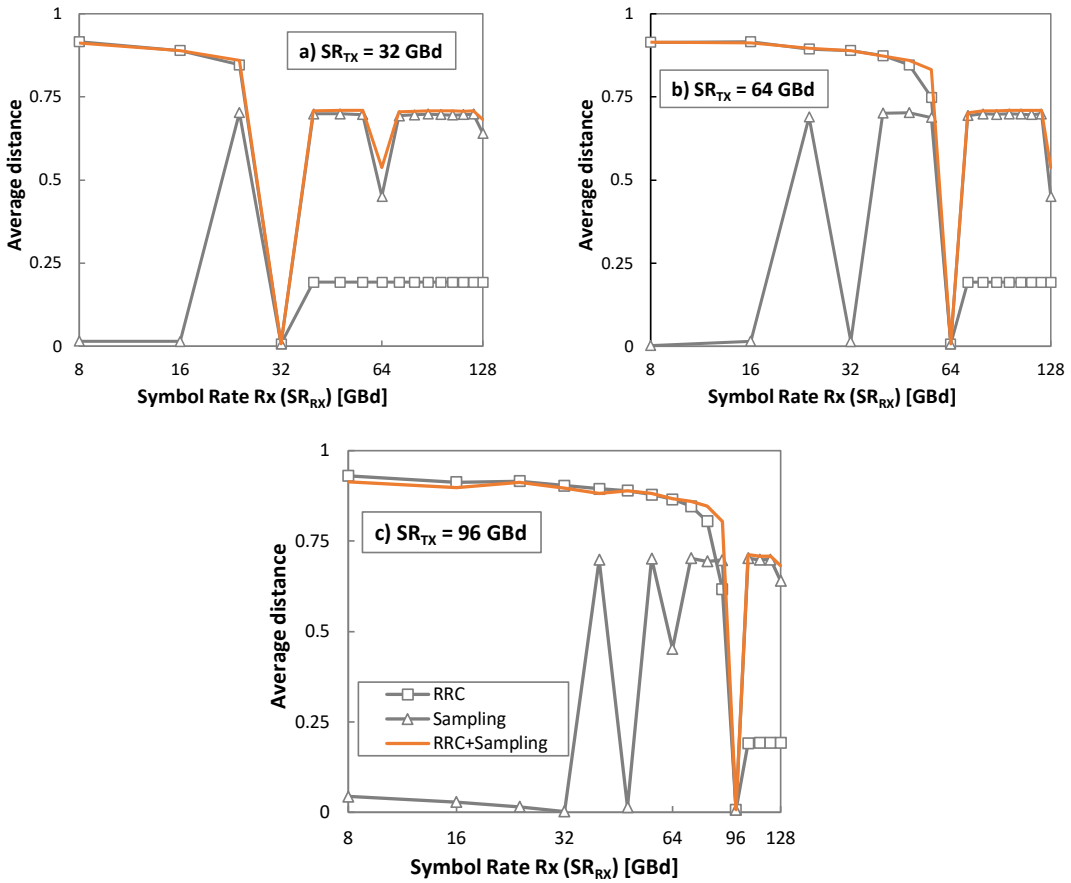


Fig. 5-2. Average Euclidean distance as a function of the SR_{Rx} for three different SR_{Tx} considering a B2B scenario.

From the results in Fig. 5-2, one can observe that only when $SR_{Rx} = SR_{Tx}$ the average Euclidean distance will be minimal. In addition, as the expected centroids are defined for every MF, the average Euclidean distance will be minimized when the right one is selected. Our algorithm for identifying the SR and the MF of a received optical signal can be summarized as:

Hypothesis: given a signal at the Rx side, its configuration, in terms of SR and MF, can be identified by finding the combination that minimizes the Euclidean distance from the symbols decoded with every supported SR and the centroids defined by the supported MFs, i.e.:

$$\min_{\{<SR,MF>\}} \bar{d}_2(S_{SR}, C_{MF}) \tag{5-5}$$

The flowchart presented in Fig. 5-3 defines the algorithm that identifies the configuration of the received signal in terms of SR and MF. It receives the configurations supported by the transponder with a list of supported SRs and MFs. After initializing variables, the algorithm exploits the hypothesis presented above

by finding the minimum distance of the symbols for all supported configurations (pairs <SR, MF>). The algorithm iterates over the supported SRs and decodes the received signal with the selected SR to obtain the set S_{SR} of decoded symbols; the average Euclidean distance of the decoded symbols with the centroids C_{MF} defined by that MF is computed and compared with the minimum obtained so far.

Assuming the optical Tx has selected a configuration that is supported by the optical Rx, the identified configuration is correct provided that the hypothesis is also true. Note that although the hypothesis has been demonstrated for B2B scenarios, the symbols will suffer from distortion induced by LI and NLI noise, and therefore, the conditions for which the above hypothesis is met should be investigated.

The time complexity of the algorithm is related to the number of configurations <SR, MF> supported by the BVT and to the complexity of Eq. (5-1). Therefore, the complexity can be approximated using the big O notation, as:

$$O(|SR| \cdot |MF| \cdot |S_{SR}| \cdot |C_{MF}|) \tag{5-6}$$

In the next section, the algorithm presented in Fig. 5-3 is exhaustively evaluated for several scenarios that include LI and NLI noise for high-order MF and high-speed SRs.

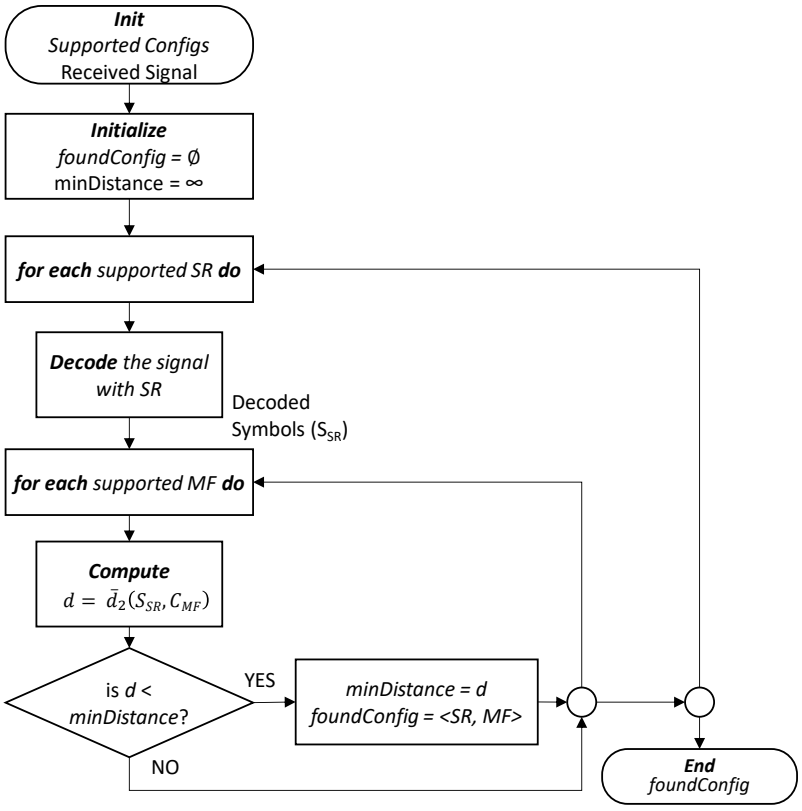


Fig. 5-3. MF and SR identification algorithm.

5.4 **Illustrative Simulation Results**

5.4.1 Single Carrier Optical Systems

The evaluation of the proposed identification algorithm has been performed through MC-based simulations carried out in MATLAB. We modeled a WDM system with 11 single-carrier channels and analyzed the results on the central one, which is impaired by the adjacent channels. 20 equiprobable 2^{13} PRBS were generated, which translates into 1638 and 1365 symbols when 32QAM and 64QAM is used, respectively. Next, the modulated signal is shaped by a RRC filter with 0.06 roll-off factor. Single polarization 32 and 64QAM signals, and SR of 32, 64, and 96 GBd, which have small tolerance to NLI noise, were considered. The SSMF and LEAF spans are modeled as shown in Table 5-I [Fr21] and the pulse propagation is modeled by solving the NLSE using the SSFM with a step size of 100 meters. We considered fiber spans of 80 km followed by an EDFA at the end of each one, which fully recovers the fiber losses, and it is characterized by a noise figure of 4.5 dB. An ideal optical coherent Rx is considered, with ideal CD compensation and carrier phase recovery. To evaluate the system performance, we measured the BER by direct error counting.

Table 5-I. Parameters for SSMF and LEAF spans

	SSMF	**LEAF**
Attenuation factor (α) [dB/km]	0.21	0.225
Dispersion parameter (D) [ps/(nm·km)]	16.8	4.2
Nonlinear parameter (γ) [(W·km)$^{-1}$]	1.14	1.3

Before evaluating the proposed MF and SR identification algorithm, we investigate the maximum reach that each configuration can support, aiming at finding the proper range of applicability leading to useful and fair accuracy analysis. Fig. 5-4 presents the obtained pre- FEC BER as a function of the optical power for: a) 32QAM signals and SSMF spans; b) 32QAM/LEAF; c) 64QAM/SSMF; and d) 64QAM/LEAF. Incremental number of spans and 75 GHz (dashed lines) and 100 GHz (solid lines) channel spacings are considered. In view of these results, we can conclude that higher BER is obtained when LEAF is used, with respect to using SSMF, due to the smaller CD ($D_{SSMF} >> D_{LEAF}$) and, consequently, higher impact from NLI noise. In addition, the overall optimal optical power is around -1 dBm for SSMF and LEAF, so we consider this optical power hereafter.

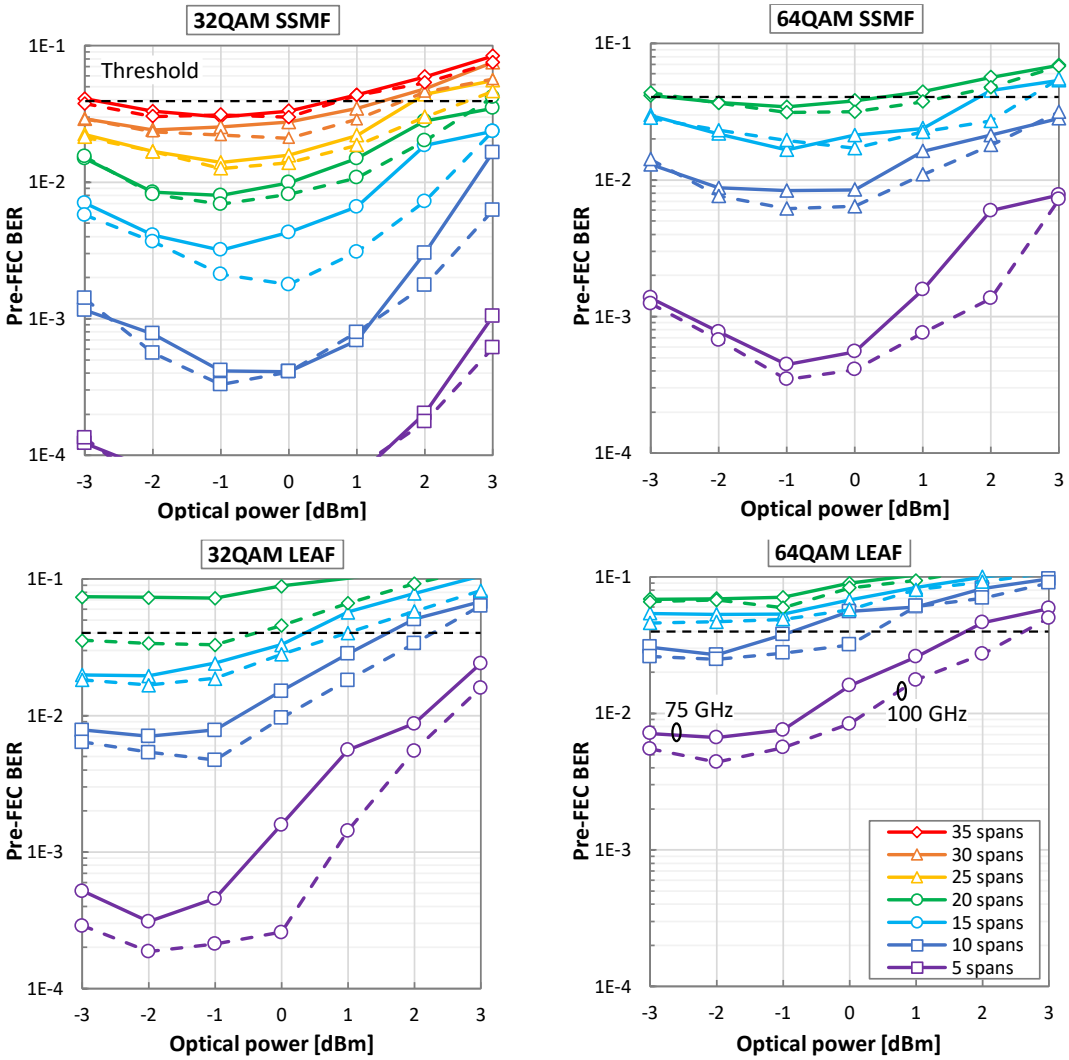


Fig. 5-4. Pre-FEC BER as a function of optical power for 32 and 64QAM, SSMF and LEAF over different optical fiber lengths.

Fig. 5-5 shows the pre-FEC BER versus the number of spans for the several cases investigated. Considering a pre-FEC BER threshold of 4×10^{-2}, we can conclude that for 32QAM and LEAF spans, the signal can reach up to 18 and 24 spans with 75 and 100 GHz channels spacing, respectively. Although, the pre-FEC BER remains below the threshold even after transmission along 35 SSMF spans. In the case of 64QAM, the signal can reach up to 22 and 23 spans of SSMF, and up to 11 and 13 spans of LEAF, when setting the channel spacing to 75 and 100 GHz channels, respectively.

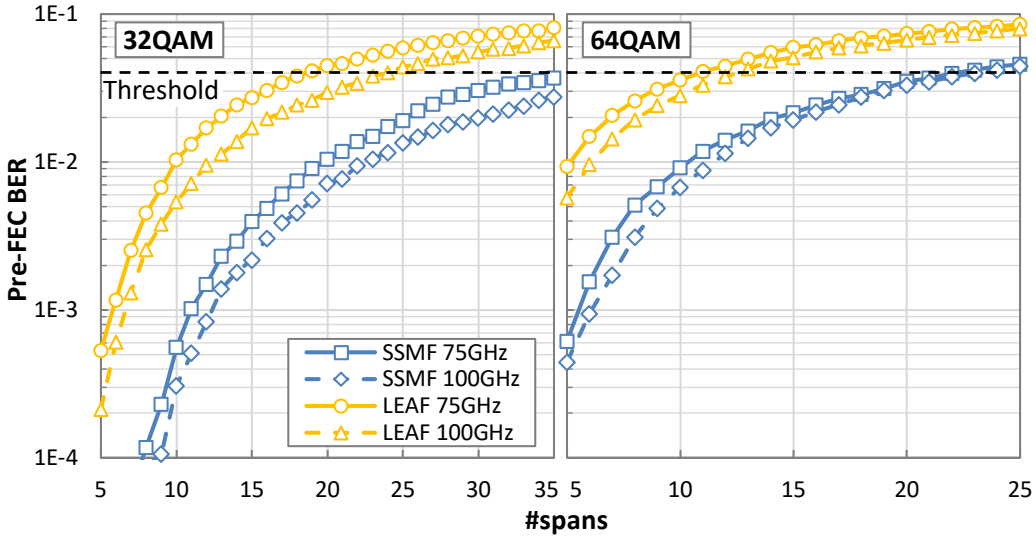


Fig. 5-5. Pre-FEC BER as a function of the number of spans.

We now evaluate the accuracy of the proposed algorithm, which entails identifying the scenarios where the hypothesis stated in Section 5.3 can be confirmed. The algorithm was applied to identify the signal configuration (i.e., 32QAM or 64QAM and 64GBd for the generated sequences). It is worth mentioning that the algorithm processes a received signal after collecting enough number of symbols to produce accurate identification results; recall that the complexity of the identification algorithm is related to such number. In particular, with the number of symbols generated for 32QAM and 64QAM, and assuming that the BVT supports 3 different SRs and 3 different MFs, the number of operations that the identification algorithm will perform (Eq. (5-6)) is in the order of 10^6, which makes it suitable for real-time implementation.

Fig. 5-6 shows the obtained accuracy of the algorithm as a function of the number of spans for the optical power of -1 dBm. We can observe that the algorithm achieves perfect identification (100% accuracy) for SSMF spans in all cases investigated. For LEAF spans and 32QAM signals, the accuracy only decreases from 100% after transmission along 32 spans with a channel spacing of 75 GHz; in the case of 100 GHz channel spacing, it achieves perfect performance up to transmission along 35 spans. When 64QAM signals are considered, the accuracy is 100% until 11 and 13 spans for 75 and 100 GHz, respectively. Note that this is the maximum number of spans imposed by the pre-FEC BER threshold.

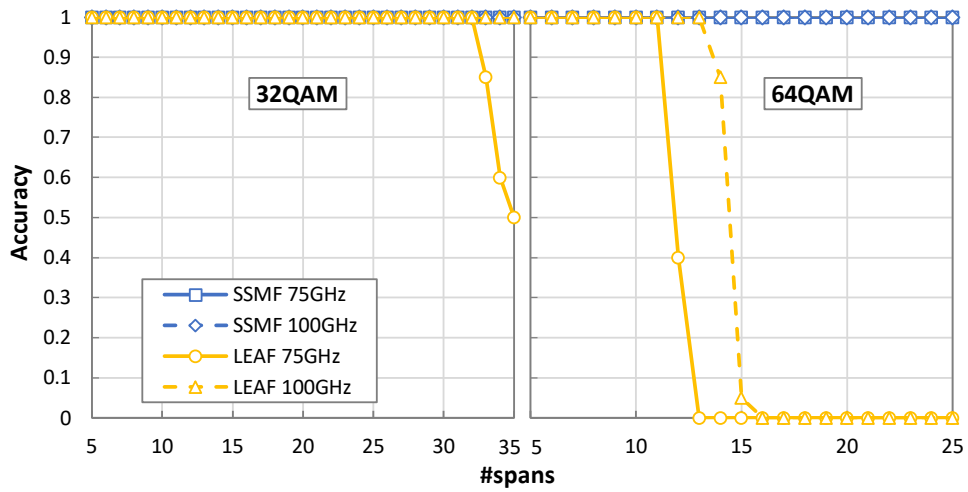
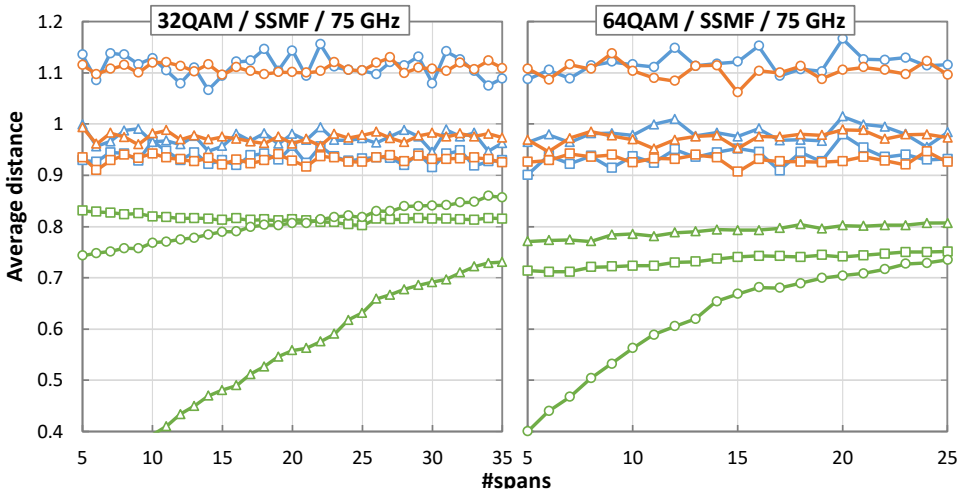


Fig. 5-6. Accuracy as a function of the number of spans.

We now analyze the obtained average Euclidean distances for the different scenarios to identify the conditions where our hypothesis is true, i.e., the minimum average Euclidean distance from the received symbols to the centroids defined by the selected MF is obtained when the signal is decoded with the SR and MF used at the Tx side. Fig. 5-7 plots the average Euclidean distance for 32QAM and 64QAM, and for the more restrictive 75 GHz channels spacing case, when the received signal is decoded with 32, 64 and 96 GBd and computed for 16, 32, and 64 QAM. We observe that the Euclidean distance for the correct SR/MF is minimal in the entire range of spans until the noise (mainly NLI noise) causes the received signal to exceed the pre-FEC BER threshold. Importantly, the hypothesis is always true when the pre-FEC BER is below the threshold, which empirically demonstrates our hypothesis for a wide range of application scenarios.



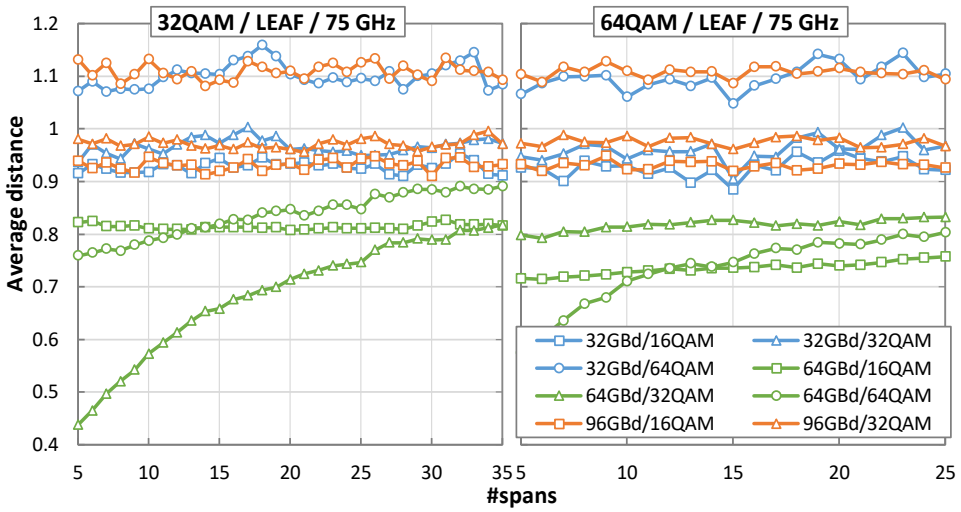


Fig. 5-7. Average distances as a function of the number of spans for 32 and 64QAM, SSMF and LEAF, and 75GHz optical channels.

Finally, Table 5-II studies the number of symbols that need to be analyzed for the highest MF supported under the worst-case scenario. As observed, 1365 and 2048 symbols need to be collected for SSMF and LEAF spans, respectively, for 100% accuracy.

Table 5-II. Accuracy [%] as a function of the number of received symbols for 64QAM/75GHz.

#Symbols	**Collecting time @32GBb**	**SSMF (25 spans)**	**LEAF (11 spans)**
512	16 ns	83%	40%
1024	32 ns	99%	75%
1365	42.6 ns	100%	89%
2048	64 ns	100%	100%

5.4.2 DSCM Optical Systems

Regarding the accuracy of the proposed algorithm in DSCM systems, let us first study the minimum OSNR for each configuration defined by the pair <*MF*, *SR*> assuming a B2B scenario. We have considered QPSK, 8QAM and 16QAM, and 8 and 11 GBd per SC, assuming no guard-band between SCs. The results have been obtained through MC-based simulation carried out in MATLAB, where a 2^{13} PRBS was generated and shaped by a RRC filter with 0.15 roll-off factor; no NLI was considered since the DSCM signals are low SR. Fig. 5-8 presents the obtained BER, computed through direct error counting, as a function of the OSNR. Assuming a pre-

FEC BER threshold of 4×10^{-3}, which includes about one order of magnitude to cope with system margins, the minimum OSNR for every combination is also highlighted.

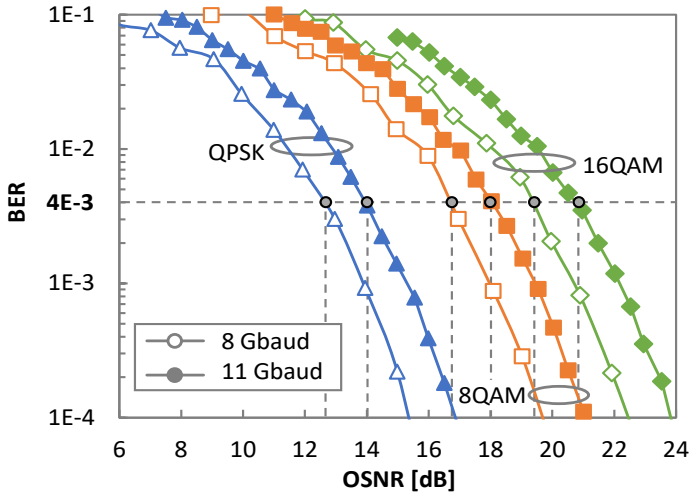


Fig. 5-8. BER as a function of the OSNR for all considered SC configurations.

Table 5-III. Minimum OSNR for threshold BER

Config. (<MF, SR>)	**OSNR [dB] (BER = 4×10^{-3})**
QPSK, 8	12.7
QPSK, 11	14.0
8QAM, 8	16.8
8QAM, 11	18.1
16QAM, 8	19.5
16QAM, 11	20.8

Let us now analyze the accuracy of proposed algorithm for MF and SR configuration recognition in SC signals. Note that the target is to reach the maximum accuracy of the algorithm for scenarios with OSNR values summarized in Table 5-III, for the different configurations.

Comparing the average distances between the received constellation and the sets of expected constellation points representing three different MFs will reveal at the Rx side the actual MF used by the Tx. However, one key aspect of the algorithm is the assumption that two configurations with the same MF and different SR will produce different average distance between the received constellation and the expected constellation points, being that with the lowest distance the one actually selected by the optical Rx. Table 5-IV shows the distances resulting from evaluating the received

constellation under the conditions in Table 5-III. We observe significant differences in the computed distances between the SRs, being the average distance for the right one the shortest in all cases. Note that the minimum difference is as high as 11% for <QPSK, 8> evaluated under OSNR = 12.7 dB and 10 ROADMs with 50 GHz slot width, which allows for accurate recognition even in the worst conditions. Note also that the difference increases when the OSNR increases, e.g., the difference of the distances is almost double for <16QAM, 8> evaluated under OSNR = 19.5 dB.

Fig. 5-9 presents the results of the accuracy of the proposed algorithm obtained for all configurations as a function of the OSNR. We observe that the algorithm recognizes the received configuration with 100% accuracy well below the minimum required OSNR in Table 5-III.

Table 5-IV. Average Distances for SR Detection

TX / RX [GBd]	**QPSK**			**8QAM**			**16QAM**		
	8	**11**	**diff**	**8**	**11**	**diff**	**8**	**11**	**diff**
8	0.57	0.63	11%	0.53	0.60	13%	0.46	0.55	20%
11	0.75	0.63	19%	0.81	0.60	35%	0.76	0.54	41%

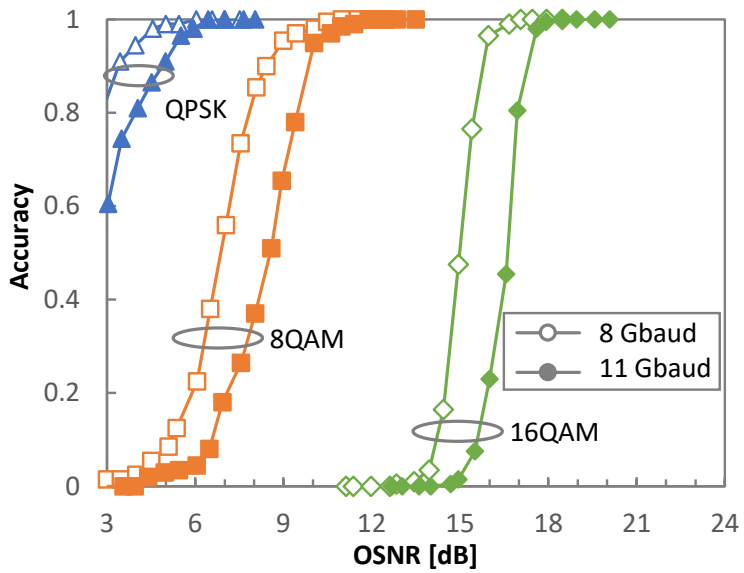


Fig. 5-9. Accuracy of SC configuration recognition as a function of the OSNR.

5.5 Concluding Remarks

By analyzing the extensive simulation results presented in this chapter, we can conclude that the proposed algorithm for MF and SR identification is accurate and can in run real-time. The algorithm is based on a hypothesis that relates the decoded symbols with the Euclidean distance to the expected centroids defined by each supported MF, and states that such distance should be minimal for the right SR and MF used at the Tx side. The hypothesis is supported by the two main effects that impact on the Euclidean distance when an optical signal is decoded at the Rx side with SRs and MFs different than those used at the Tx side: *i*) narrower RRC filter when $SR_{Rx} < SR_{Tx}$; and *ii*) sampling step when $SR_{Rx} > SR_{Tx}$.

Simulation results of a coherent WDM system for realistic scenarios were presented, where the optical signal was impaired by both LI and NLI noise. 32QAM and 64QAM signals, SSMF and LEAF spans and 75 and 100 GHz channels spacing were considered. The results showed 100% accuracy for the maximum reach imposed by the pre-FEC BER threshold for all the cases investigated.

Additionally, we have investigated the proposed algorithm in DSCM systems, where the effect of the NLI noise is lower. For that purpose, we simulated DSCM signals based on QPSK, 8QAM and 16QAM, working at 8 and 11 GBd. In line with the results for single carrier optical systems, the results showed 100% accuracy in the feasible lightpath range for all cases investigated.

Chapter 6

OCATA: A Deep Learning-based Digital Twin for the Optical Time Domain

The development of DT to represent the optical transport network might enable multiple applications for network operation, including automation and fault management. In this chapter, we propose a deep learning-based DT for the optical time domain, named OCATA. OCATA is based on the concatenation of DNN modeling optical links and nodes, which facilitates representing lightpaths. The DNNs model linear and nonlinear noises, as well as optical filtering. Additional DNN-based models are proposed to extract useful lightpath metrics, such as lightpath distance, number of optical links and nonlinear fiber parameter. OCATA exhibits low complexity, which makes it ideal for real-time applications. Illustrative results for the application of OCATA to disaggregated and mixed disaggregated-proprietary optical networks scenarios reveal remarkable accuracy.

6.1 Introduction

During the last years, optical networks have evolved from static and manually operated networks to more dynamic and flexible ones to support traffic dynamicity requirements. This evolution has been possible thanks to technologies like advanced ROADMs architectures, in the data plane, and SDN and MDA, in the control plane [EON16], [Ve19].

In addition, with the introduction of 5G and beyond services, more autonomous, efficient, and reliable optical networks are required in all network segments, from core to metro and access [Ha21]. In that regard, ML has received special attention and several works have demonstrated their benefits for optical communication improvements and network automation (see e.g., tutorials [Ra18], [Ve22]). The low

computational requirements of ML models make them very attractive to solve hard computational problems. For example, the authors in [Fr21] proposed DNN for NLI noise equalization, and the authors in [Mo18] proposed ML-based models for QoT estimation in single domain, and the authors in [Pr19] in multidomain network scenarios. In contrast, the well-known analytical model SSFM [Sh14] solves the NLSE equation with high accuracy but requires high computational requirements, which limits its utilization for real-time operations and in dynamic optical network scenarios.

ML, SSFM, and other methods can help in the control and management of optical networks, and provide high performance when solving dedicated, isolated network problems. However, they fail to provide a holistic representation of the network with high accuracy and low computational complexity. In this context, the use of a DT has been proposed to improve lightpath provisioning, QoT estimation, fault management, etc. [Wa21]. Some initiatives already exist, like the open-source project GNPy [Cu22], which implements the GN model [Po14] to estimate the signal's LI and NLI noise powers. GNPy provides independent models for fiber propagation, OAs, and ROADMs, which can be concatenated to estimate the QoT of an optical connection. The use of GNPy as a DT of the optical network was suggested by the authors in [Ba21], where the output of the models was analyzed and used for soft-failures prediction, localization, and identification.

In fact, the possibility to have independent models enables developing specific models for devices with different characteristics and it is key for the application of GNPy in open disaggregated scenarios [Gi18]. Independent modeling allows also to obtain results not only in the optical Rx, but also in any intermediate location of a lightpath. In addition, different models can work together to represent the coexistence of different optical fiber types, optical channels working on non-flat optical powers profiles and heterogenous data traffic, based on different modulation formats, symbol rates, channel spacing, etc. Note that such heterogeneity originates different signal-to-noise interactions along the lightpaths. Not only the LI noise contribution, but also signal-to-nonlinear noise interactions play an important role in current and future optical networks.

However, in addition to estimating the QoT, a DT should produce expected signals that can be compared with those obtained from the network [Vel18], [Sh19]. In that way, deviations between the observed and the expected signals can be detected and used to, e.g., soft-failure detection. In our published article in [JOCN22], we proposed DNNs to model the LI noise contribution of ROADMs and optical links in the time domain i.e., to IQ OC. Like in the case of GNPy, such specific models can be concatenated to represent lightpaths. In addition to the benefits discussed above regarding specific models, in the case of ML this allows to have a database (DB) of pre-trained models for elements with different characteristics like fiber length and number of spans in the case of optical links. Then, creating a DT for a lightpath consists in concatenating models for the specific elements in the route of the

lightpath. In this chapter, we introduce models for both LI and NLI noises and filtering impairments. So, optical signals include LI noise from OAs, NLI noise from fiber Kerr effects, and optical filtering impairments from the WSS inside ROADMs. We call such deep learning-based DT for the optical time domain as OCATA. We evaluate OCATA under two relevant scenarios for disaggregated networks and mixed disaggregated-proprietary networks.

The remaining of the chapter is organized as follows. Section 6.2 motivates the use of an optical DT and presents the considered network scenarios. Section 0 introduces the features defined to extract useful characteristics from IQ OC and explain how we develop, train, and concatenate feedforward DNN-based models for IQ OC. Section 6.4 explains the proposed models for OC analysis focusing on lightpath modeling under the considered optical network scenarios. Illustrative numerical results are showed and discussed in Section 6.5, from the generation of the datasets to the OCATA performance. Finally, Section 0 draws the main conclusions of this chapter.

6.2 Optical Time Domain Digital Twin and Application Scenarios

Developing an accurate and low-complex optical transmission and networking time domain DT can lead to significant improvements in network management, from lightpath provisioning to network maintenance, which would have significant impact on operational costs.

Throughout this work, we propose an optical time domain DT named OCATA, and investigate two different optical network scenarios: 1) disaggregated optical networks [Gi18], with TRx, ROADMs and OAs from multiple vendors (Fig. 6-1a)); and 2) mixed disaggregated-proprietary, with single-vendor proprietary islands providing transparent optical connectivity to disaggregated segments. The example in Fig. 6-2a) represents two disaggregated segments (S1 and S3) connected through a vendor-proprietary one (S2).

In the disaggregated scenario, information regarding the network topology, the type of fibers, etc., as well as the configuration and monitoring data from every optical component is accessible. In this scenario, a lightpath from site A to site Z can be modeled by concatenating models for the different components supporting such lightpath, i.e., TRx, ROADMs, and optical links [JOCN22]; see the example in Fig. 6-1b). In brief, the concatenated model is based on propagating IQ OC features. A Tx model generates the initial constellation features following the Tx configuration. Models for ROADMs and optical links are concatenated in the same order that the respective network components appear in the route of the lightpath. Each single model modifies the input features according to the noise that the specific physical network component introduces. The different models are concatenated, so the output

features of one model are the input features of the following one. Finally, to reduce complexity, only the features of few selected constellation points are propagated. In consequence, a constellation reconstruction (CR) module generates the features of the non-propagated constellation points based on the received features to complete IQ OC.

If the models are accurate, the samples collected from the optical transponder in Z would match the expected OC obtained with OCATA. As stated in the introduction section, in this chapter we consider both LI and NLI noise, so extensions to the concatenated model proposed in [JOCN22] are needed.

In the mixed scenario, we assume that vendor-specific optical network management systems (NMS) provide software interfaces for provisioning but hide relevant information of the underlying optical line systems, e.g., the route of the lightpath. Since no information about the ROADMs and optical links supporting the lightpath in the proprietary domain are available in this case, a concatenation model cannot be built in the same way than in the disaggregated scenario. A possible solution is to create an ad-hoc DNN model for the whole vendor-proprietary segment, as suggested in Fig. 6-2b) for segment S2 and use such model to create concatenated model for the E2E lightpath.

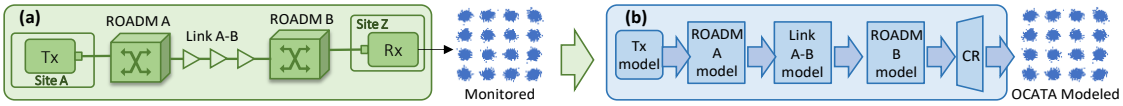


Fig. 6-1. a) Example of lightpath in disaggregated scenarios and b) DNN-based concatenation model for lightpath abstraction.

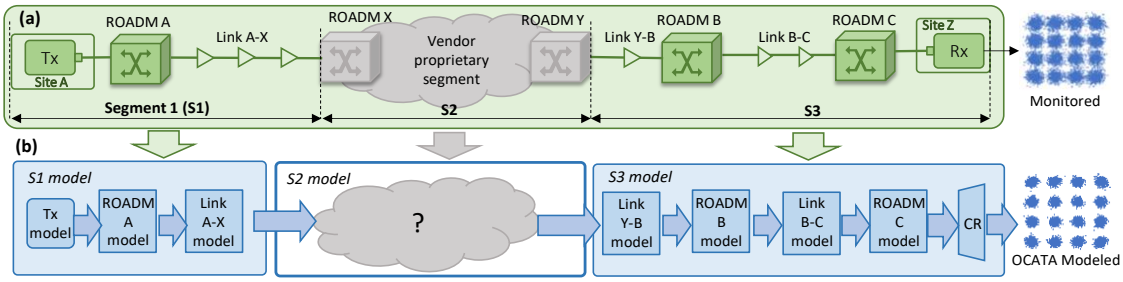


Fig. 6-2. a) Example of lightpath in mixed disaggregated-proprietary scenarios and b) DNN-based concatenation model with per-segment analysis.

Different NMSs will provide different partial information about the internals of the segment, e.g., the distance from input port to output port for the optical connection or the number of intermediate ROADMs. A possibility would be to infer the model for segment S2 by analyzing the monitored IQ OC and finding a model for S2 that results into similar observed and expected OC. Next sections introduce lightpath IQ

OC modeling including both LI and NLI noises, OCATA and some use cases for the defined network scenarios.

6.3 Deep Learning-based Constellation Modeling

In this section, we first elaborate on IQ OC feature extraction; recall that features are propagated along the concatenated DNN models representing the lightpath. Since different network components impair differently the optical signal traveling along the lightpath, different DNN-based models are used to model links and ROADMs network components as well as LI and NLI noises.

6.3.1 IQ OC Features Engineering

Discrete optical signal samples X are defined by a sequence of symbols defined as complex numbers, where the real and imaginary parts represent the I and Q components of the optical signal, respectively. Such sequence is commonly represented as an IQ OC, where every symbol belongs to one of the constellation points M, being m the number of constellation points, i.e., $m=|M|$.

Fig. 6-3 left an input sample X from a 16QAM optical signal, i.e., m=16. We apply GMM [MMA20] to characterize a given OC sample as multiple bivariate Gaussian distributions, i.e., one distribution for each constellation point. Then, for 16QAM signals, we obtain 16 different distributions.

Fig. 6-3 right represents the GMM fitting of the OC sample X. The objective is to generate the set of semi-supervised constellation features Y that summarizes X with a number of clear, unequivocal, and predefined statistical characteristics. Let us denote Y^i the vector of features characterizing constellation point i, i.e., $Y^i=[y^k]^i$, where y^k represents the k-th feature. Then, Y is formally defined as the vector that specifies all constellation point features, i.e., $Y = [Y^i, \forall i \in M]$.

In particular, GMM fitting is used to characterize every constellation point i by means of a two-component vector $[\mu^I,\mu^Q]$ representing the mean position in the constellation and a three-component vector $[\sigma^I,\sigma^Q,\sigma^{IQ}]$, which captures the real and imaginary variance and symmetric covariance terms that the symbols belonging to the constellation point i experience around the mean. Therefore, each constellation point i is characterized by 5 features, i.e., $Y^i=[y^k]^i=[\mu^I,\mu^Q,\sigma^I,\sigma^Q,\sigma^{IQ}]^i$. For an m-QAM signal, the GMM fitting is initialized to find m different bivariate Gaussian distributions and mean values to be the respective expected constellation centroids, e.g., (−3,3).

Note that in the case of highly impaired optical signals, symbols would be very dispersed, which makes constellation characterization more challenging. In addition, by forcing the constellation points to be modeled as bivariate Gaussian distributions and initializing the GMM fitting with the expected constellation

centroids, the obtained features are strongly conditioned. For this reason, we compute the loglikelihood value $\mathcal{L}$ as a goodness-of-fit metric to estimate GMM fitting accuracy [MMA20].

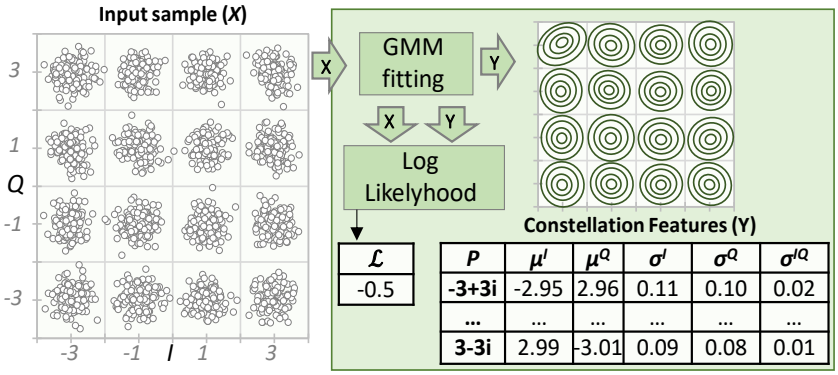


Fig. 6-3. GMM fitting to optical IQ constellation features extraction.

6.3.2 Linear Links and ROADMs Models

Linear DNN-based models target to propagate the LI features Y_{LI}. Hence, to develop the DNN-based models, we use the features obtained after GMM fitting considering the LI noise only.

To reduce the complexity of the DNNs while keeping the good performance, we limit the constellation points that are propagated to a selected subset. In particular, for 16QAM signals, we selected 4 out of 16 constellations points, which greatly reduces the number of input neurons. Specifically, two inner and two outer constellation points are selected to obtain knowledge from higher and lower power constellation points. Consequently, each DNN for 16QAM signals receives 4 constellations points × 5 features = 20 inputs. Different DNNs need to be trained for different optical link configurations, i.e., length and number of spans. In addition, we consider two models for the ROAMs, one for the add and transit ROADMs that include a booster OA at the output and another one for the drop ROADM.

After each individual DNN is trained, the main approach is to concatenate the desired models to emulate the desired lightpath. For example, we model the lightpath in Fig. 6-1a) by concatenating 3 DNN models: add ROADM A, link A-B, and drop ROADM B. It is worth noting that we can extract the propagated Gaussian features at the output of each model for OC reconstruction purposes (see details in Subsection 6.3.4).

6.3.3 Nonlinear Residuals Models

For the sake of clarity in the description, let us assume that we can get the features of an IQ OC that includes the NLI noise, Y. The features related to the LI noise, i.e., without nonlinear impairments, Y_{LI}, can be obtained as defined in the previous subsection. Then, we compute what we call NLI residuals features, ΔY_{NLI}, as:

$$\Delta Y_{NLI} = \mathrm{Y} - Y_{LI} \tag{6-1}$$

In line to the linear models, DNNs are used to model NLI residuals features. We assume that ROADMs do not add any NLI noise. In addition, NLI noise is related to the distance that the optical signal traverses and therefore, DNNs modeling NLI noise for optical fibers require the distance as additional input. However, to follow the concatenation model, we need models for both ROADMs and optical links, and hence: *i*) ROADM models just copy the inputs to the outputs; and *ii*) DNNs for the optical links require two additional inputs for the accumulated distance until the link, denoted as d, and for the link length, denoted as l. Fig. 6-4 represents the concatenated model for a lightpath considering both NLI and LI noises; for simplicity, we only represent a Tx, a ROADM, and a link.

After obtaining both, the NLI residuals and the linear features, the predicted features Y_{pred} for the signal are computed as:

$$Y_{pred} = Y_{LI} + \Delta \mathrm{Y}_{NLI} \tag{6-2}$$

Details of the differentiated propagation to predictions of X_{LI} and ΔX_{NLI} along the lightpath is shown in Fig. 6-4. For the sake of simplicity, we represent a segment with the three main network elements that compose it, i.e., Tx, ROADMs, and links.

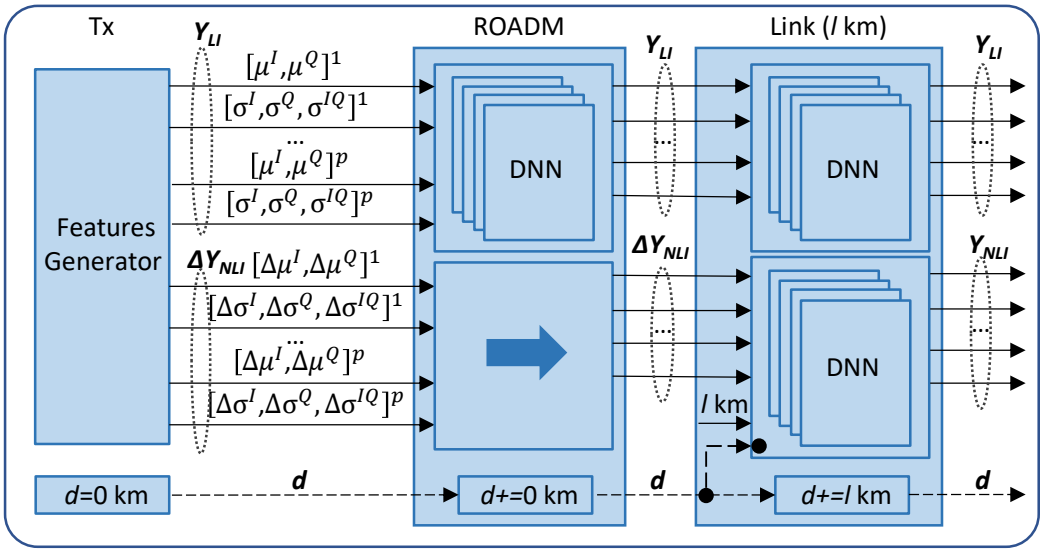


Fig. 6-4. Details on LI and NLI features propagation.

6.3.4 Constellation Reconstruction

As previously stated, to reduce the size of the DNNs, as well as their complexity, we select and propagate a selected subset of constellation points $C \subseteq M$ from a given IQ OC sample of an m-QAM signal. CR is the process that predicts the features of every constellation point as a function of the features of the propagated ones, i.e.:

$$Y^i = F^i\left(\left[Y^j, \forall j \in C\right]\right) \quad \forall i \in M \tag{6-3}$$

Note that $F^i(\cdot)$ for the constellation points in C can be defined as:

$$F^i\left(\left[Y^j, \forall j \in C\right]\right) = Y^i \quad \forall i \in C \tag{6-4}$$

Recall that $Y^i=[\mu^I,\mu^Q,\sigma^I,\sigma^Q,\sigma^{IQ}]^i=[y^k]^i$. In [JOCN22], we proposed finding functions $F^i(\cdot) = [f^k(\cdot)]^i$ as linear combinations of the propagated LI features. Consequently, we need to train $(|M|-|C|)\times 5$ linear models. For instance, for 16QAM, this method requires training $(16-4)\times 5=60$ $f^k(\cdot)$ functions and it is not clear that linear models will provide accurate results when NLI noise is considered. For these very reasons, in this work, we propose a different approach based on one single DNN that takes as input propagated features $[Y^j, \forall j \in C]$ and outputs non-propagated features $[Y^i, \forall i \in M \setminus C]$. For instance, for 16QAM, the DNN has 20 inputs and 60 outputs.

6.4 Digital Twin for the Optical Time Domain and Constellation Analysis

Once the methodology for IQ OC modelling and features propagation, including LI and NLI noises, has been detailed in the previous section, let us now present the algorithms to create a time domain DT (OCATA) of a given lightpath under the two different optical network scenarios defined in Section 6.2. Next, some examples of constellation analysis that take advantage of OCATA are presented.

6.4.1 OCATA for Disaggregated Optical Networks

Algorithm 6-1 presents the general algorithm of OCATA for a given lightpath. OCATA receives: *i*) the complete or a partial route of the lightpath (*Route*); it includes all the needed information for the Tx, the ROADMs, and the optical links, and returns the IQ OC *X'* that we could expect at the last element in the route; *ii*) the reference to DB that contains objects with already trained DNN models for the different elements in the route of a lightpath (*ModelDB*); and *iii*) the length of the PRBS to be generated for the initial OC (*n_bits*). Individual models are trained for every element and with specific characteristics: *i*) Txs with different configurations of modulation format, symbol rate, power, etc.; *ii*) optical links with different lengths and spans; and *iii*) add-transit and drop ROADMs.

The algorithm assumes that the first element in the route is a Tx. Therefore, the description of the Tx is used to find a model in the model DB that fits the characteristics of the Tx detailed in the configuration. Such model is then used to generate the initial features to be propagated (lines 1-3 in Algorithm 6-1). Then, the current distance of lightpath (d) and the length added by the last element (l) are initialized (line 4).

Next, for each element in the route (lines 5-15), d is updated with the length added by the last element and l is reset (line 6); it will be updated only in the case of an optical link (line 11). Then, the specific model for current element in the route is retrieved from the model DB, i.e., an add-transit or drop ROADM model, or an optical link model with the defined configuration in terms of length and number of spans (lines 7-13). Once a model has been found, the LI and NLI features are propagated separately from the current ones (lines 14-15). Note that the propagation of the NLI residuals features depends on the fiber length l of the element and accumulated distance d. In the case of a ROADM, the NLI residuals features propagation consist of just forwarding the input values without modification.

After propagating the features through all elements in the route, the CR model is retrieved from the models DB and used to predict the complete set of constellations point features Y' from the propagated ones Y_{pred}. Next, features Y' are sampled and the IQ OC X' is obtained (lines 16-18). Next, a model of feasible loglikelihood values for lightpaths matching the characteristics of the given one is retrieved from the DB (line 19). Finally, the IQ constellation, its features, and the range of loglikelihood values for the lightpath are returned (line 20).

Algorithm 6-1. OCATA main algorithm.

INPUT: *Route, ModelDB, n_bits*
OUTPUT: X', Y', $[\mathcal{L}_{min}, \mathcal{L}_{max}]$

```
 1: e ← Route.pop()
 2: Txmodel ← ModelDB("Tx", e.config)
 3: Y_LI, ΔY_NLI ← featuresGen(Txmodel, e.config, n_bits)
 4: d, l ← 0
 5: for each e in Route do
 6:     d ← d + l; l ← 0
 7:     if e.type = "add-transit" then
 8:         model ← ModelDB("ROADM", "add-transit")
 9:     else if e.type = "link" then
10:         model ← ModelDB("link", e.config)
11:         l ← e.config.length
12:     else
13:         model ← ModelDB("ROADM", "drop")
14:     Y_LI ← model.propagateLI(Y_LI)
15:     ΔY_NLI ← model.propagateNLI(ΔY_NLI, d, l)
16: CRmodel ← ModelDB("CR", null)
17: Y' ← CRmodel(Y_pred = Y_LI + ΔY_NLI)
18: X' ← sampling(Y')
```

19: *LLHmodel* ← *ModelDB*("LogLikelihood", *route*)
20: **return** *X', Y', LLHmodel*.getRange()

Algorithm 6-2 presents the generation of the initial LI and NLI residuals features that will be propagated afterwards. The algorithm receives as input the Tx model, the required Tx configuration, and the length of the PRBS to be generated. After generating the bit sequence (line 1 in Algorithm 6-2), the initial IQ OC X is obtained by sampling, shaping and modulating the bit sequence [DC01] following the received Tx configuration (line 2). Next, GMM fitting is performed to compute the set of bivariate Gaussian distribution that better fits the generated OC X (line 3). Since such features come from an IQ OC without NLI noise, they are stored in Y_{LI}, whereas the NLI residuals features are initialized to 0 (line 4). Finally, subset C of constellation points features to be propagated is selected and returned (line 5).

Algorithm 6-2. Features Generator.

INPUT: *Txmodel, Txconfig, n_bits*
OUTPUT: Y_{LI}, ΔY_{NLI}

1: *bitSequence* ← PRBS(*n_bits*)
2: X ← *Txmodel*.generateConst (*bitSequence, Txconfig*)
3: Y_{LI} ← GMMfitting(X)
4: ΔY_{NLI} ← [0,..,0]
5: **return** selectFeatures(Y_{LI}, ΔY_{NLI})

6.4.2 Tuning OCATA with Received Samples

In this subsection, we present a methodology that allows to find unknown network configuration by tuning OCATA to force that the generated IQ OC are as closed as possible to the received ones. One possible application of this methodology is in mixed disaggregated-proprietary optical networks, where the configuration of the proprietary segment is not completely known.

Let us assume the example of mixed disaggregated-proprietary optical networks in Fig. 6-2, where the only known characteristic of segment S2 is the distance. Conversely, the number of optical (hops) and their characteristics in terms of spans, gamma parameter, etc. are unknown. Our target thus, is to find a configuration of segment S2 so that when IQ constellations are generated with OCATA following such configuration, results into IQ OC that are similar to those measured. We compare two IQ OC X_1, X_2 by computing the Euclidean distance of the difference of their constellation point features Y_1, Y_2, i.e.,

$$\mathrm{diff}(X_1, X_2) = \|Y_1 - Y_2\|_2 \tag{6-5}$$

To find the configuration of segment S2, we can train DNN models that use features Y as input and predict the number of hops of the lightpath and the nonlinear gamma parameter. Hence, such DNNs have $|M|\times 5$ inputs (e.g., for 16QAM, the DNN have 80 inputs) and one single output.

Algorithm 6-3 presents the procedure that finds the most likely proprietary segment configuration. The algorithm receives a sampled constellation *X*, the route, an initial proprietary segment with two ROADMs and one single optical link with the total length of the proprietary segment, the models DB, and a list with feasible values for the non-linear gamma parameter. First, the features *Y* of the received sample are extracted using GMM fitting (line 1 in Algorithm 6-3) and the loglikelihood value is computed to estimate GMM fitting accuracy (line 2). Predicted number of hops of the lightpath and end-to-end non-linear gamma parameter are compute using DNN models from the models DB (lines 3-6). The closer feasible value for the nonlinear gamma parameter to that predicted for the whole lightpath is selected for all the hops of the segment and the proprietary segment in the *Route* is expanded with the predicted number of hops of the segment and the value of the non-linear gamma parameter (lines 7-9). OCATA is then used to generate the estimated constellation *X'* and its features *Y'* using a PRBS of the same length than the bit sequence of the received constellation (lines 10-11) and the difference between the constellation sample and the generated one is computed (line 12). Because it is possible that some of the optical links in the proprietary segment have different fibers, and thus different non-linear gamma parameter, a search is started to find the configuration that minimizes the difference between the sampled and generated IQ constellations (lines13-20). The loglikelihood computed for the constellation sample is verified to be in the range of loglikelihood values for lightpaths of the found characteristics; otherwise, an error is generated (lines 21-22). Finally, the found configuration for the proprietary segment is returned (line 23).

Algorithm 6-3. Find proprietary segment model.

INPUT: *X, Route, propSegment, ModelDB, GammaList*
OUTPUT: *propSegment*

```
 1: Y ← GMMfitting(X)
 2: ℒ ← logLikelihood(X, Y)
 3: LPNHopmodel ← ModelDB("LPNHops", Route)
 4: nHopsSample ← LPNHopmodel (Y)
 5: LPGammamodel ← ModelDB("LPGamma", Route)
 6: LPgammaSample ← LPGammamodel(Y)
 7: gammaSegment ← get(GammaList, LPgammaSample)
 8: nHopsSegment ← nHopsSample-Route.nHops
 9: propSegment ← expandSegment(propSegment, nHopsSegment,
                    gammaSegment)
10: n_bits ← computeBits(X)
11: X', Y', ℒmin, ℒmax ← OCATA(Route, ModelDB, n_bits)
12: diffX ← diff(X, X')
13: while true do
14:     gammaSegment ← getGammaConfig(GammaList,
                    propSegment, LPgammaSample)
15:     newPropSegment ← expandSegment(propSegment,
                    nHopsSegment, gammaSegment)
16:     X', Y', ℒmin, ℒmax ← OCATA(Route, ModelDB, n_bits)
```

```
17:     newDiffX ← diff(X, X')
18:     if newDiffX > diffX then break
19:     diffX ← newDiffX
20:     propSegment ← newPropSegment
21: if ℒ not in [ℒ_min, ℒ_max] then
22:     return ERROR("loglikelihood value out of range")
23: return propSegment
```

6.4.3 Constellation Analysis

The opportunities for the analysis of OC that OCATA opens are broad. For instance, similarly to the ideas presented in the previous subsection, one could devise algorithms that compare predicted lightpaths' metrics from measured OC to the ones predicted from constellations generated with OCATA, e.g., to detect and localize soft failures. In this case, DNN models to predict metrics can be extended to include, in addition to the number of hops and nonlinear gamma parameter, the total length of the lightpath and the optical power. Note that the evolution of such metrics can be used to detect soft-failures and degradations.

Algorithm 6-4 compares lightpaths' metrics predicted from the features of received and generated OC and returns the results. The algorithm receives a sampled constellation *X*, the route, and the models DB. Four metrics will be predicted, and the names of the DNN models are defined in a vector (line 1 in Algorithm 6-4). Next, the features *Y* of the received sample are extracted using GMM fitting (line 2) and the loglikelihood value is computed to estimate GMM fitting accuracy (line 3). OCATA is the used to generate the estimated constellation *X'* and its features *Y'* using a PRBS of the appropriated length (lines 4-5). The loglikelihood computed for the constellation sample is verified to be in the expected range; otherwise, an error is generated (lines 6-7). Next, the lightpaths' metrics are computed for the received constellation sample and for that generated with OCATA using the specific DNNs trained to predict each metric (lines 8-11) and the obtained results are eventually returned together with the difference between the sampled and generated OC (line 12).

Algorithm 6-4. Constellation Analysis.

```
INPUT: X, Route, ModelDB
OUTPUT: results, diffX
 1: lpModels ← ["LPLength", "LPNHops", "LPPower",
                "LPGamma"]
 2: Y ← GMMfitting(X)
 3: ℒ ← logLikelihood(X, Y)
 4: n_bits ← computeBits(X)
 5: X', Y', ℒ_min, ℒ_max ← OCATA(Route, ModelDB, n_bits)
 6: if ℒ not in [ℒ_min, ℒ_max] then
 7:     return ERROR("loglikelihood value out of range")
 8: results ← {}
```

```
 9: for each lpModel in lpModels do
10:     model ← ModelDB(lpModel, Route)
11:     results.add(lpModel, <model(Y), model(Y')>)
12: return results, diff(X, X')
```

6.5 Illustrative Numerical Results

In this section, we firstly introduce the simulation scenario and the IQ OCs synthetic data generation for the numerical evaluation of OCATA. Then, we focus on evaluating the key aspects involved on the OC characterization and analysis. Finally, the performance of OCATA methodology for both disaggregated and mixed disaggregated-proprietary scenarios is provided.

6.5.1 Optical Signal Samples Generation

To evaluate OCATA, we generated optical signal samples by means of a coherent digital optical system simulator developed in MATLAB based on SSFM. Specifically, an optical system consisting in 11 WDM optical channels was considered where all use 16QAM at 64GBd and 75 GHz channel spacing. The optical power per channel was set either to −1 dBm, around the optimal one, or to 2 dBm, working in a higher nonlinear regime. The fiber was modeled as standard single mode fiber characterized by attenuation factor of 0.21 dB/km, chromatic dispersion parameter of 16.8 ps/(nm·km) and nonlinear parameter gamma (denoted as γ) of 1.14 $(W \cdot km)^{-1}$. In addition, we considered $\gamma = 0$, i.e., without NLI noise, and $\gamma = 2$ $(W \cdot km)^{-1}$, i.e., high nonlinear transmission regime. The OAs were modeled as EDFA considering a noise figure of 4.5 dB and ideal gain, introducing ASE noise to the optical signal. The WSSs' shape inside the ROADMs were modeled by averaging real spectral transfer function measurements, including filter losses and ripples [Ra14]. At the optical coherent Rx, digital signal processing able to perform ideal chromatic dispersion compensation and carrier phase recovery was considered.

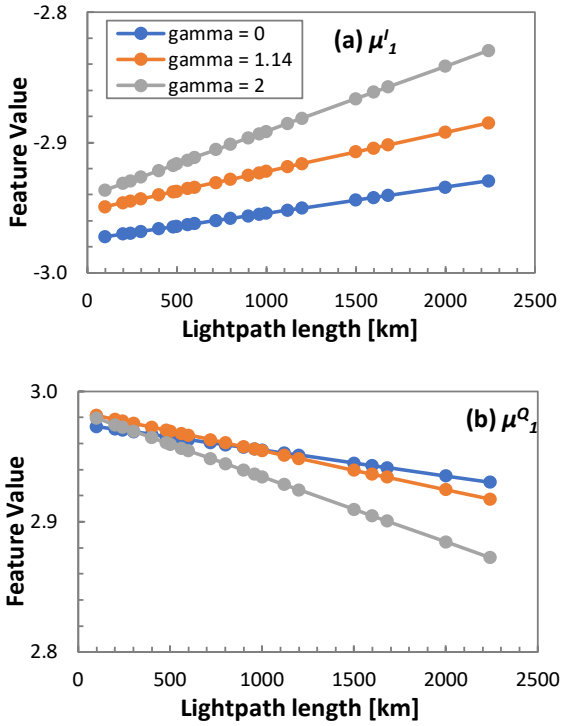
Regarding links and lightpaths configuration, we considered 7 different types of optical links changing the total length, span length, and number of spans, i.e., *i*) 100km (2 spans of 50km); *ii*) 200km (4×50km); *iii*) 240km (4×60km); *iv*) 300km (4×75km); *v*) 400km (5×80km); *vi*) 500km (5×100km); and *vii*) 560km (7×80km). Then, by configuring from 2 to 5 cascaded ROADMs, i.e., from 1 to 4 hops, optical signal samples for lightpaths with total length ranging from 100 to 2,240 km were generated.

Following the optical system configuration described above, we defined a number of scenarios changing the link and lightpath configuration (note that different type of links were assumed in the same lightpath), non-linear parameter γ (0 or 2 $(W \cdot km)^{-1}$), and optical power per channel (−1 and 2 dBm). For every scenario, 10 signal samples

were generated using 2^{13} PRBSs and shaped by a RRC filter with roll-off factor of 0.06.

6.5.2 Constellation Features and Loglikelihood Analysis

Fig. 6-5 shows an example of the evolution of the OC features as a function of the total lightpath length for constellation point 1 ($-3+3i$) and for the several transmission regimes considered, i.e., LI regime and both NLI regimes with nonlinear parameter γ of 1.14 or 2 (W·km)−1. We observe the strong correlation between the feature value and the lightpath length, as well as noticeable differences among transmission regimes. As expected, for LI regime the feature value range (defined as the difference between values at shortest and longest lightpath) is smaller than that for the NLI ones. Fig. 6-5 also suggests the importance of accurate models not only for LI noise, but also for the NLI one, i.e., to clearly distinguish between different NLI regimes.



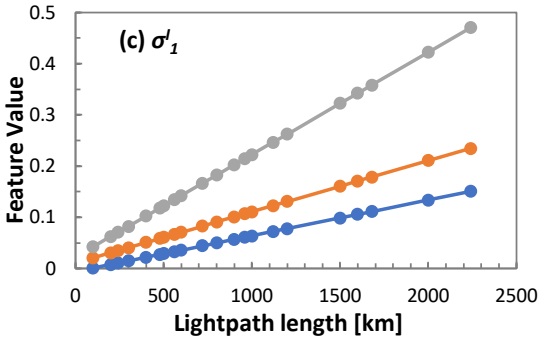


Fig. 6-5.Lightpath length effect on constellation features μ^I_1 (a), μ^Q_1 (b), and σ^I_1 (c).

Additionally, Fig. 6-6 shows the evolution of the features as a function of the number of hops for four lightpath scenarios. Every curve is for a single lightpath scenario with the same total lightpath length and nonlinear parameter γ, and connects points where the lightpath consisted in 2, 3, or 4 hops. Table 6-I summarizes the specific configuration of every lightpath in terms of the length on the links. Although correlation is not so strong than with the lightpath distance in Fig. 6-5, there are evidence of differences in the constellation features among the different scenarios. In consequence, we conclude that the selected bivariate Gaussian features provide an OC characterization that allows distinguishing among lightpaths with different configuration in terms of length and NLI regime.

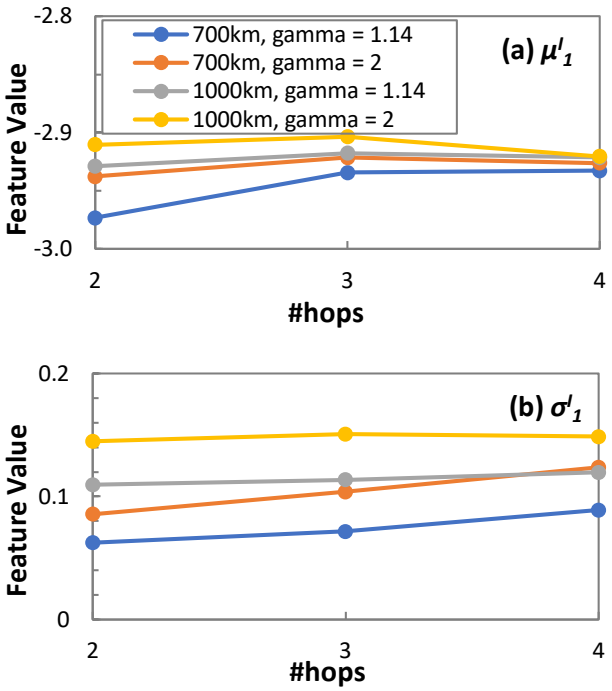


Fig. 6-6. Number of hops effect on constellation features μ^I_1 (a) and σ^I_1 (b).

Table 6-I. Links lengths [km] for the different lightpath scenarios.

Lightpath Length [km]	**#hops**		
	2	**3**	**4**
700	400, 300	300, 200, 200	400, 100, 100, 100
1,000	500, 500	400, 400, 200	400, 400, 100, 100

Another important parameter for OCATA is the loglikelihood $\mathcal{L}$ of the bivariate Gaussian distribution that characterizes a given optical signal sample. Fig. 6-7 shows the evolution of $\mathcal{L}$ as a function of the number of symbols collected, averaging all constellation points, for 3 different lightpath length scenarios: short (200km), medium (1,000km), and long (2,080km). As expected, $\mathcal{L}$ decreases with the lightpath length due to the loss of gaussianity caused by NLI effects that distort constellation points shape. In other words, expected $\mathcal{L}$ clearly varies with lightpath scenario, and therefore, it can be successfully used to perform constellation analysis validation like, e.g., that in lines 6-7 of Algorithm 6-4. We observe that $\mathcal{L}$ reaches a stationary state after around 1,000 symbols for all the considered scenarios. Note that the optical samples were generated with 2,048 symbols each, which is well in the stationary region.

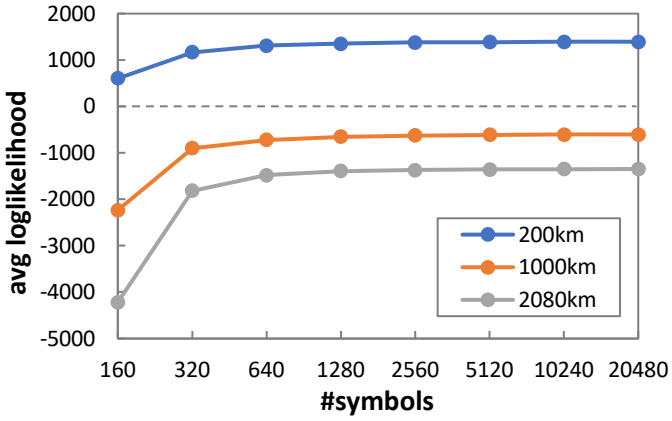


Fig. 6-7. Average loglikelihood as a function of the number of symbols.

Finally, Fig. 6-8 evaluates the performance of the constellation comparison function *diff*, detailed in Eq.(6-5), as a function of several lightpath parameters including total length, number of hops and nonlinear parameter γ. Each curve belongs to a reference sample X_1 that is compared against other samples X_2 with the characteristics defined in the x axis. For instance, the blue curve in Fig. 6-8a) presents the results of the comparison between a reference X_1 sample of a 200km lightpath and X_2 samples of

lightpath with 200km, 1,000km, and 2,080km. We observe that the minimum value (*diff*~0) is achieved when X_1 and X_2 are from lightpaths with the same configuration. This fact is evident for all configurations of total length (Fig. 6-8a)), number of hops (Fig. 6-8b)), and nonlinear parameter γ (Fig. 6-8c)). These results validate the usefulness of the *diff* function, which is essential when evaluating and finding proper lightpath configuration for proprietary segments in mixed disaggregated-proprietary optical network scenarios.

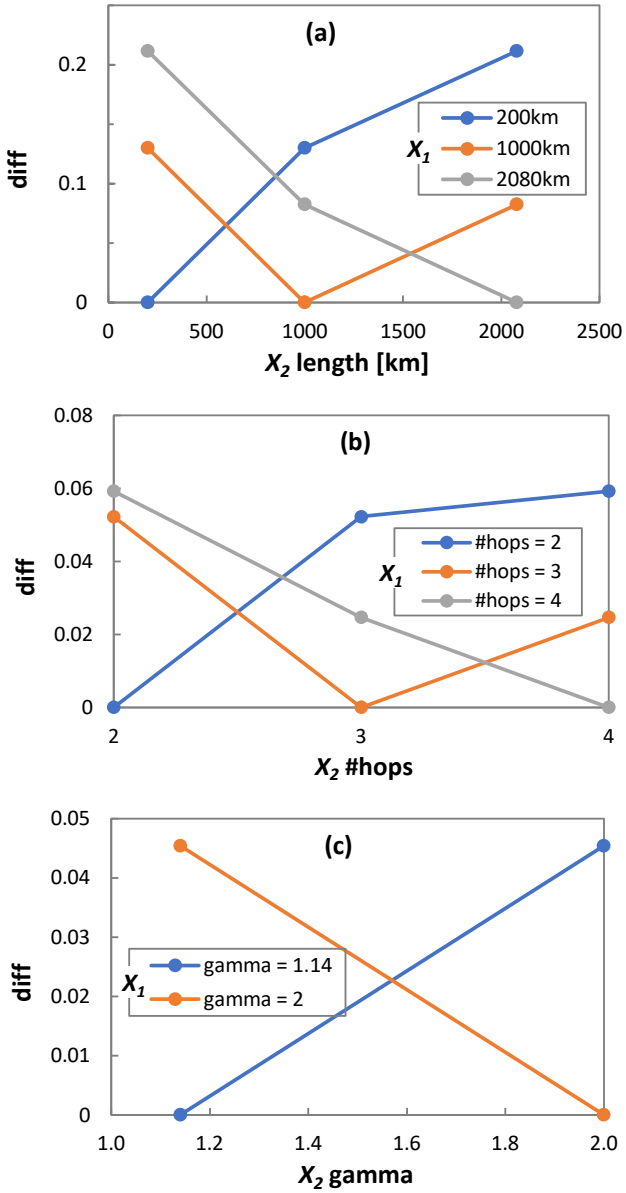


Fig. 6-8. Evaluation of the OC differences (*diff*) as a function of lightpath parameters.

6.5.3 OCATA Performance Evaluation

To evaluate the performance of the OCATA algorithm detailed in Algorithm 6-1, let us concentrate on studying the accuracy of the models involved in the whole process, i.e., the DNN models in charge of characterizing optical components and the constellation reconstruction models. As detailed in Subsection 6.3.4, aiming at reducing DNN models' complexity, we selected a subset C of constellation points from the whole range of M points to be propagated. In particular, for 16QAM signals, we selected two inner and two outer ones, i.e., constellation points 1 ($-3+3i$), 7 ($1+1i$), 10 ($-1-1i$) and 15 ($1-3i$).

Considering a disaggregated network scenario like the one represented in Fig. 6-1, we assume that full information about all network elements and its configuration is available and, consequently, we can concatenate pre-trained DNN models to model a given lightpath with precision. Table 6-II shows the average relative error as a function of the model type for different optical components. The results are obtained by averaging all the predicted mean μ and variance σ features of the propagated constellation points. As for μ, the average relative error is small for all the DNN models considered (< 2%). Regarding σ, errors are below 20% in all components, except for the ROADM add/transit where the error increases.

Table 6-II. Average relative error of optical components models.

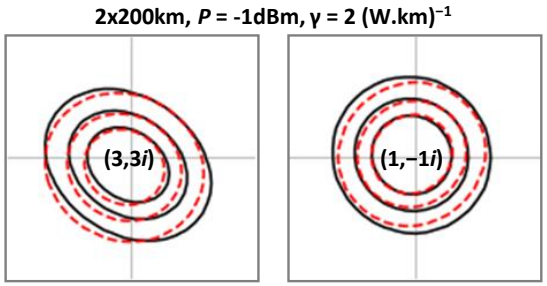
Model	**Average relative error**			
	μ^{I}	μ^{Q}	σ^{I}	σ^{Q}
Optical link (100km)	0.004	0.003	0.160	0.160
Optical link (200km)	0.004	0.003	0.097	0.112
Optical link (400km)	0.006	0.006	0.130	0.137
Optical link (560km)	0.008	0.007	0.202	0.209
ROADM add/transit	0.012	0.011	0.324	0.332
ROADM drop	0.019	0.016	0.055	0.048

After training the individual DNN-based models and performing concatenation and propagation of the selected constellation features, the reconstruction of the non-propagated constellation points was carried out using the CR model. To show the performance of CR model, Table 6-III summarizes the average relative error as a function of the total lightpath length. We observe that errors are pretty similar to those obtained with the individual DNN models in Table 6-II. Therefore, we conclude that concatenation and reconstruction processes do not add significant incremental error in the received constellation, which is an essential result to validate the OCATA methodology.

For illustrative purposes, Fig. 6-9 depicts the bivariate Gaussian distributions of constellation points 4 ($3+3i$) and 11 ($1-1i$) obtained with OCATA (red dashed lines) and by configuring the same lightpath in the MATLAB simulator detailed in Subsection 6.5.1 (black solid lines). Note that these constellation points are not propagated by optical component models; therefore, this figure allows to visually evaluate the overall performance of the propagation and reconstruction processes in the OCATA methodology. Two different lightpath scenarios were considered with different lightpath length, optical power, and nonlinear parameter γ. Specifically, a lightpath of 400km (2×200km), power of −1 dBm and nonlinear parameter $\gamma = 2$ $(\text{W}\cdot\text{km})^{-1}$ is shown in Fig. 6-9 top and a lightpath of 1,600km (4×400km), power of 2 dBm and $\gamma = 1.14$ $(\text{W}\cdot\text{km})^{-1}$ is shown at bottom. The similarity of Gaussian distributions eventually validates the accuracy and usefulness of the OCATA methodology.

Table 6-III. Average relative error of the CR model.

Lightpath length [km]	μ	σ
100	0.01	0.36
200	0.01	0.30
300	0.01	0.25
560	0.01	0.23
960	0.02	0.19
1,200	0.02	0.16
2,000	0.04	0.30



avg rel error	μ	σ
2x200km	~ 1.4%	~ 20%
4x400km	~ 1.9%	~ 15%

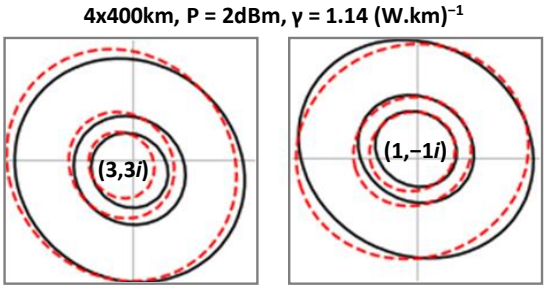


Fig. 6-9. OCATA vs MATLAB-based simulator performance.

Let us now move to the mixed disaggregated-proprietary optical network scenario illustrated in Fig. 6-2. We focus on the performance evaluation of the DNN-based models detailed in Algorithm 6-3 and used to predict lightpath characteristics from OC features. To this aim, we have trained different DNN models to predict the lightpath length, the number of hops and the optical power of the lightpath as a function of the OC features. All DNNs models followed the same structure: 80 input neurons, i.e., 5 features × 16 constellations points, 3 hidden layers with 160, 80, and 20 neurons, respectively, and one single output neuron predicting the specific characteristic of that model. In addition, all the models used: *i*) hyperbolic tangent (*tanh*) as activation function; *ii*) *root mean squared propagation* as optimization algorithm; *iii*) training stage up to 5,000 epochs; and *iv*) mean square error as loss function.

Fig. 6-10 shows the performance of the several models considered in this work. Regarding the DNN to predict the total lightpath length (Fig. 6-10a)), the average relative error is below 35% for all the whole length range and below 5% for distances above 500km. We tested the proposed models with the predicted constellation features only with the LI noise and considering both LI and NLI noises. In both cases we obtained similar performances, concluding that the NLI residuals models are not introducing additional error for the predicted IQ OC features. Concerning the DNN to predict the number of hops in the route of a lightpath (Fig. 6-10b)), the model

shows remarkable accuracy, providing average relative error below 4% for all the cases. Once again, both LI and LI + NLI residuals features are providing similar performances. Finally, for the DNN to predict the optical power (Fig. 6-10c)), the model produces average relative errors below to 7% only considering the LI noise and below than 2% considering both LI and NLI noise.

By analyzing the remarkable performance of the proposed DNNs to characterize a given lightpath given the OC features, we conclude that a good enough characterization is possible to understand the crossed proprietary segments and its parameters, as well as to compare the expected with the observed OC features.

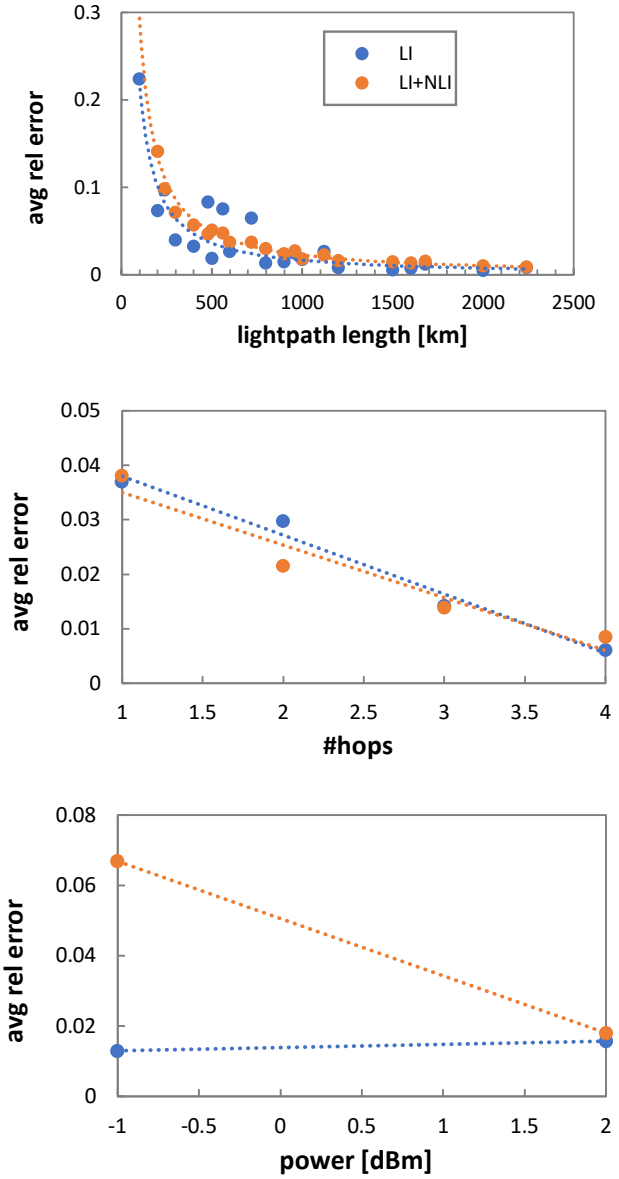


Fig. 6-10. Performance of lightpath metrics DNN-based predictors.

To further illustrate this, let us evaluate the performance of Algorithm 6-3 to find the most likely proprietary segment model. Considering a reference 1,300km lightpath sample X_1 crossing a proprietary segment with 2 hops: of 300km and 200km; the results are show in Fig. 6-11. We observe that the minimum diff value is obtained exactly when the reference sample X_1 is compared against one modeled sample X_2 with similar configuration, i.e., 2 number of hops and nonlinear parameter $\gamma = 1.14$ $(W \cdot km)^{-1}$ (Fig. 6-11a)) or $\gamma = 2$ $(W \cdot km)^{-1}$ (Fig. 6-11b)).

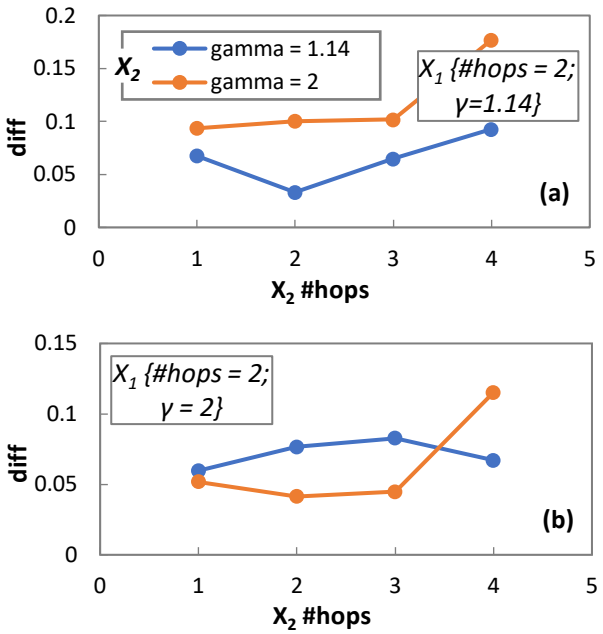


Fig. 6-11. Performance of proprietary segment model algorithm.

6.6 Concluding Remarks

A low complexity deep learning-based digital twin for the optical time domain, named OCATA has been presented. DNN models for IQ OC include LI noise coming from optical amplification, optical filtering coming from the ROADMs and NLI noise coming from the optical fiber propagation. OCATA is based on concatenation of several DNNs models specific network elements, i.e., ROADMs and optical links. The proposed DNN-based models are fed by features characterizing IQ OC.

Two optical network scenarios have been analyzed for disaggregated and mixed disaggregated-proprietary domains. For each network scenario, several lightpaths with different configuration in terms of span and link length, number of hops, optical power, and nonlinear parameter, have been used to obtain datasets to train, test and validate the proposed models. Exhaustive results show the accuracy of OCATA for optical constellation modeling including constellation features propagation and

reconstruction in both LI and NLI optical transmission regimes. Additionally, DNN-based models for E2E lightpath modeling in such mixed disaggregated-proprietary optical networks have been proposed. Such models can predict typical lightpath parameters, such as number of hops, optical power, and nonlinear parameter gamma, from OC features.

With these models, OCATA can be used as a tool for network automation and failure management, including anticipated degradation detection, identification, and localization.

Chapter 7

Closing Discussion

7.1 Main Contributions

The main contributions of this thesis are:

- In Chapter 4, we investigated the optical filtering impact in hybrid filterless and filtered optical networks. The main objective is understanding the minimum channel spacing needed to avoid higher OSNR penalties due to the optical filtering in order to increase the overall SE for future metro optical networks. The simulation results obtained suggest the use of high-order MF, i.e., 32 and 64QAM, as well as high SR, up to 96 GBd, in future optical metro networks avoiding OSNR penalties from the optical filtering when an optical signal crosses up to 10 cascaded ROADMs. In the scenarios investigated, SE between 5.12 and 7.68 b/s/Hz is obtained, foreseen the use of high-order MF and high SR in future highly efficient optical metro networks.
- Chapter 5 proposes an accurate and low complex algorithm to MF and SR identification for autonomous lightpath operation. The possibility of network elements, such as the optical TRx, work autonomously reduces the overall complexity of the SDN solutions. For instance, if the TRx are able to detect the MF and SR of the received signal, without interaction with the SDN controller, the complexity is reduced. The proposed algorithm is based on the Euclidean distance between the received symbols and the supported configurations of the TRx. The observed behavior when an optical signal is decoded with the different configurations, lead to the hypothesis that the correct configuration is the one that minimizes the Euclidean distance. The illustrative simulation results show remarkably accuracy in both single carrier optical systems, based on high-order MF and high-SR, and DSCM optical systems, based on low-order MF and low SR. Several lightpath scenarios are investigated, based on different type of fibers and channel spacing.

- In Chapter 6, OCATA is proposed and evaluated. OCATA is a deep learning-based DT for the optical time domain, specifically, for modelling and analysis of IQ OC. The idea of concatenate DNN-based models to abstract typical lightpaths composed by TRx, optical links and ROADMs is introduced. Additional DNN-based models are proposed to predict lightpath metrics as a function of the OC features. The OCATA is proposed not only for disaggregated optical network scenarios but also for mixed disaggregated-proprietary where there is not enough information for modeling proprietary network segments. The overall OCATA performance is evaluated showing good accuracy for OC modeling and analysis. OCATA can be used as a DT of the optical physical layer in order to, for example, predict QoT or detect misconfigurations in optical network elements.

7.2 List of Publications

7.2.1 Publications in Journals

[JSAC21] L. Velasco, S. Barzegar, D. Sequeira, A. Ferrari, N. Costa, V. Curri, J. Pedro, A. Napoli, and M. Ruiz, "Autonomous and Energy Efficient Lightpath Operation Based on Digital Subcarrier Multiplexing," IEEE Journal on Selected Areas in Communications, vol. 39, pp. 2864-2877, 2021

[JOCN22] M. Ruiz, D. Sequeira, and L. Velasco, "Deep Learning -based Real-Time Analysis of Lightpath Constellations [Invited]," in IEEE/OSA Journal of Optical Communications and Networking, vol. 14, pp. C70-C81, 2022

[PNET] D. Sequeira, M. Ruiz, N. Costa, A. Napoli, J. Pedro, and L. Velasco, "Accurate Low Complex Modulation Format and Symbol Rate Identification for Autonomous Lightpath Operation," submitted to Photonic Network Communication

[JOCN] D. Sequeira, M. Ruiz, N. Costa, A. Napoli, J. Pedro, and L. Velasco, "OCATA: A Deep Learning-based Digital Twin for the Optical Time Domain," submitted to IEEE/OPTICA Journal of Optical Communications and Networking

7.2.2 Publications in Conferences

[ACP21] D. Sequeira, M. Ruiz, A. Napoli and L. Velasco, "Filtering Impact on High-Order Modulation Formats in Hybrid Filterless and Filtered Optical Metro Networks," OSA Advanced Photonics Congress, 2021

[ECOC21.1] D. Sequeira, M. Ruiz, N. Costa, A. Napoli, J. Pedro, and L. Velasco, "Lightweight Optical Constellation Modelling by Concatenating Artificial Neural Networks," European Conference on Optical Communications, 2021

[ECOC21.2] M. Ruiz, J. Morales, D. Sequeira and L. Velasco, "An Autoencoder-Based Solution for IQ Constellation Analysis," European Conference on Optical Communications, 2021

7.3 List of Research Projects

REAL-NET: Real-time monitoring and mitigation of nonlinear effects in optical networks, European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie (G.A. 813144).

7.4 Other Achievements

In the scope of the REAL-NET project, this PhD thesis contribute with several deliverables:

- "D3.1 Real-time optical performance monitoring for meshed optical networks using EON architecture";
- "D4.1 Analysis of QoT and methods to anticipate service degradations";
- "D3.2 Implementation of machine learning techniques for real-time optical performance monitoring";
- "D3.3 Experimental demonstration of advanced optical nonlinear real-time performance monitoring techniques for EON";
- "D4.2 Report on the applications of the methods to improve optimization algorithms for the network".

Additionality, the synthetic data generated to obtain the results in [JOCN22] is available in the open data repository [OCATA].

7.5 Topics for Further Research

Following, we suggest some interesting topics for further research:

- Regarding the optical filtering impact, we consider that a tighter channel frequency granularity of 6.25 GHz instead of 12.5 GHz can be investigated in the future to further SE improvements. For example, for 64GBd signals, instead of going from 75 GHz optical channels to 87.5 GHz ones, the 81.25

GHz channels can be considered and the optical filtering impact in these scenarios must be investigated.

- Take into consideration PCS solutions for the several methods proposed and investigate how it will affect both the proposed algorithm for MF and SR identification and the OCATA.
- To turn the proposed OCATA more robust, training of different DNN-based models for different fiber types, optical powers, MF, etc. can be considered.
- The proposed OCATA is based on the time domain, i.e., IQ OC. Extending the DT for the frequency domain, by propagate optical spectrum features, can be useful to have a more robust and complete DT.

List of Acronyms

AI	Artificial Intelligence
ANN	Artificial Neural Networks
ASE	Amplified Spontaneous Emission
B2B	Back-to-Back
BER	Bit Error Rate
CapEx	Capital Expenditures
CD	Chromatic Dispersion
CDC	Colorless, Directionless and Contentionless
CR	Constellation Reconstruction
DB	Database
DBP	Digital Backpropagation
DC	Data Center
DCF	Dispersion Compensating Fiber
DNN	Deep Neural Networks
DSCM	Digital Subcarrier Multiplexing
DSP	Digital Signal Processing
DT	Digital Twin
E2E	End-to-End
EDFA	Erbium-Doped Fiber Amplifiers
EON	Elastic Optical Networks
FEC	Forward Error Correction
FFT	Fast Fourier Transform

FON	Filterless Optical Networks
GMM	Gaussian Mixture Models
GN	Gaussian Noise
IFFT	Inverse Fast Fourier Transform
IM/DD	Intensity Modulation and Direct Detection
IP	Internet Protocol
IQ	In-phase and Quadrature
ISI	Inter-Symbolic Interference
LEAF	Large Effective Area Fiber
LI	Linear Interference
MC	Monte-Carlo
MDA	Monitoring and Data Analytics
MF	Modulation Format
ML	Machine Learning
M-QAM	*M*-ary Quadrature Amplitude Modulation
NLI	Nonlinear Interference
NLSE	Nonlinear Schrödinger Equation
NMS	Network Management System
OA	Optical Amplifier
OC	Optical Constellation
OEO	Optical-Electrical-Optical
OpEx	Operational Expenditures
OPM	Optical Performance Monitoring
OSA	Optical Spectrum Analyzer
OSNR	Optical Signal-to-Noise Ratio
OTN	Optical Transport Network
P2MP	Point-to-Multipoint
P2P	Point-to-Point
PCS	Probabilistic Constellation Shaping
PDM	Polarization Division Multiplexing
PLI	Physical Layer Impairments

PRBS	Pseudo-Random Binary Sequence
PSD	Power Spectral Density
QoT	Quality of Transmission
QPSK	Quadrature Phase Shift Keying
R	Router
RC	Raised-Cosine
ROADM	Reconfigurable Optical Add/Drop Multiplexers
RRC	Root-Raised-Cosine
Rx	Optical Receiver
S-BVT	Sliceable-Bandwidth Variable Transponders
SC	Subcarrier
SDN	Software Defined Networking
SE	Spectral Efficiency
SMF	Single Mode Fiber
SPM	Self-Phase Modulation
SR	Symbol Rate
SSFM	Split-Step Fourier Method
SSMF	Standard Single Mode Fiber
TRx	Optical Transponder
Tx	Optical Transmitter
URLL	Ultra-Reliable and Low-Latency
WDM	Wavelength Division Multiplexing
WSS	Wavelength Selective Switches
XPM	Cross-Phase Modulation

References

[Ag16] E. Agrell et al., “Roadmap of optical communications,” Journal of Optics, vol. 18, 2016

[Ar16] E. Archambault et al., "Routing and Spectrum Assignment in Elastic Filterless Optical Networks," IEEE/ACM Transactions on Networking, vol. 24, pp. 3578-3592, 2016

[Ay22] O. Ayoub, O. Karandin, M. Ibrahimi, A. Castoldi, F. Musumeci and M. Tornatore, "Tutorial on filterless optical networks [Invited]," IEEE/OSA Journal Optical Communications and Networking, vol. 14, pp. 1-15, 2022

[Ba06] G. Baxter *et al.*, "Highly programmable wavelength selective switch based on liquid crystal on silicon switching elements," IEEE/OSA Optical Fiber Communication Conference, 2006

[Ba21] S. Barzegar, M. Ruiz, A. Sgambelluri, F. Cugini, A. Napoli, and L. Velasco, “Soft-Failure Detection, Localization, Identification, and Severity Prediction by Estimating QoT Model Input Parameters,” IEEE Transactions on Network and Service Management, vol. 18, pp. 2627-2640, 2021

[Ca22] M. Cai, X. Liu, L. Liu, H. Lun, L. Yi, W. Hu, and Q. Zhuge, "Optical filtering impairment monitoring based on model fusion for optical networks," Optics Express, vol. 30, pp. 24639-24654, 2022

[Ch19] Y. Cheng, S. Fu, M. Tang and D. Liu, “Multi-task deep neural network (MT-DNN) enabled optical performance monitoring from directly detected PDM-QAM signals,” OSA Optics Express, vol. 27, pp. 19062-19074, 2019

[Ch21] J. Chai, X. Chen, Y. Zhao, T. Yang, D. Wang and S. Shi, “Joint Symbol Rate-Modulation Format Identification and OSNR Estimation Using Random Forest Based Ensemble Learning for Intermediate Nodes,” in IEEE Photonics Journal, vol. 13, pp. 1-6, Dec. 2021

[Cisco] "Cisco Annual Internet Report (2018–2023) White Paper", Cisco Annual Internet Report - Cisco Annual Internet Report (2018–2023) White Paper - Cisco

[Co15] B. Collings, "Advanced ROADM technologies and architectures," IEEE/OSA Optical Fiber Communication Conference, 2015

[Cu22] V. Curri, "GNPy model of the physical layer for open and disaggregated optical networking [Invited]," IEEE/OSA Journal of Optical Communication and Networking, vol. 14, pp. C92-C104, 2022

[Da15] M. Dallaglio, A. Giorgetti, N. Sambo, L. Velasco, and P. Castoldi, "Routing, Spectrum, and Transponder Assignment (RSTA) in Elastic Optical Networks," IEEE/OSA Journal of Lightwave Technology, vol. 33, pp. 4648-4658, 2015

[DC01] J. Proakis, Digital Communications, 4th ed., McGraw-Hill, 2001

[EDFA02] E. Desurvire, Erbium-Doped Fiber Amplifiers: Principles and Applications, Wiley, 2002

[EON16] Victor López and Luis Velasco, Elastic Optical Networks: Architectures, Technologies, and Control, in Optical Networks book series, ISBN 978-3-319-30173-0, Springer, 2016

[Fa16] J. Fabrega et al., "On the Filter Narrowing Issues in Elastic Optical Networks," IEEE/OSA Journal of Optical Communication and Networking, vol. 8, pp. A23-A33, 2016

[FCS12] G. Agrawal, Fiber-optic communication systems, John Wiley & Sons, 2012

[Fe20] A. Ferrari et al., "GNPy: an open source application for physical layer aware open optical networks," IEEE/OSA Journal of Optical Communication and Networking, vol. 12, pp. C31-C40, 2020

[Finisar] https://finisarwss.com

[Fl20] C. Fludger, "Performance Oriented DSP Design for Flexible Coherent Transmission," IEEE/OSA Optical Fiber Communication Conference, 2020

[Fr17] F. Frey, R. Elschner and J. Fischer, "Estimation of trends for coherent DSP ASIC power dissipation for different bitrates and transmission reaches," Photonic Networks, pp. 137-144, 2017

[Fr21] P. Freire, V. Neskornuik, A. Napoli, B. Spinnler, N. Costa, G. Khanna, E. Riccardi, J. Prilepsky, and S. Turitsyn, "Complex-Valued Neural Network Design for Mitigation of Signal Distortions in Optical Links," IEEE/OSA Journal of Lightwave Technology, vol. 39, pp. 1696-1705, 2021

[Gi18] Ll. Gifre, J.-L. Izquierdo-Zaragoza, M. Ruiz, and L. Velasco, "Autonomic Disaggregated Multilayer Networking," IEEE/OSA Journal of Optical Communications and Networking, vol. 10, pp. 482-492, 2018

[Gn08] A. Gnauck et al., "25.6-Tb/s WDM Transmission of Polarization-Multiplexed RZ-DQPSK Signals," IEEE/OSA Journal of Lightwave Technology, vol. 26, pp. 79-84, 2008

[Gr05] L. Gruner-Nielsen *et al.*, "Dispersion-compensating fibers," IEEE/OSA Journal Lightwave Technology, vol. 23, pp. 3566-3579, 2005.

[Ha21] H. Hallingby, S. Fletcher, V. Frascolla, A. Gavras, I. Mesogiti, and F. Parzysz, "5G Ecosystems," white paper, 5G-IA, 2021.

[Ho22] M. Hosseini, J. Pedro, A. Napoli, N. Costa, J. Prilepsky and S. Turitsyn, "Optimization of survivable filterless optical networks exploiting digital subcarrier multiplexing," IEEE/OSA Journal Optical Communications and Networking, vol. 14, pp. 586-594, 2022

[Infinera.1] Infinera, "The-Ultimate-Guide-to-Nyquist-Subcarriers-0208-WP-RevA-0719.pdf (infinera.com)"

[Infinera.2] Infinera, "Maximizing-the-Capacity-Reach-of-800G-Generation-Coherent-0271-WP-RevA-0920.pdf (infinera.com)"

[ITU-T G.694.1] Recommendation ITU-T G.694.1 "Spectral grids for WDM applications: DWDM frequency grid"

[ITU-T G.709] Recommendation ITU-T G.709 "Interfaces for the optical transport networks"

[Ja22] C. Janz, Y. You, M. Hemmati, Z. Jiang, A. Javadtalab and J. Mitra, "Digital Twin for the Optical Network: Key Technologies and Enabled Automation Applications," IEEE/IFIP Network Operations and Management Symposium, 2022

[Ja96] A. Jain, J. Mao and K. Mohiuddin, "Artificial neural networks: a tutorial," IEEE Computer, vol. 29, pp. 31-44, 1996

[Ja99] A. Jain, M. Murty, and P. Flynn, "Data clustering: A review," ACM Computing Surveys., vol. 31, pp. 264-323, 1999

[Kh15] P. Khodashenas, J. Moscoso, B. Shariati, D. Marom, D. Klonidis, and I. Tomkos, "Investigation of Spectrum Granularity for Performance Optimization of Flexible Nyquist-WDM-Based Optical Networks," IEEE/OSA Journal of Lightwave Technology, vol. 33, pp. 4767-4774, 2015

[Lo20] V. Lopez, W. Ishida, A. Mayoral, T. Tanaka, O. Gonzalez de Dios and J. Fernandez-Palacios, "Enabling fully programmable transponder white boxes [Invited]," IEEE/OSA Journal of Optical Communications and Networking, vol. 12, pp. A214-A223, 2020

[Lumentum] https://www.lumentum.com

[Ma13] G. Mantelet, A. Cassidy, C. Tremblay, D. Plant, P. Littlewood and M. Bélanger, "Establishment of dynamic lightpaths in filterless optical networks," IEEE/OSA Journal Optical Communications and Networking, vol. 5, pp. 1057-1065, Sept. 2013

[Ma97] D. Marcuse, C. Manyuk and P. Wai, "Application of the Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence," IEEE/OSA Journal of Lightwave Technology, vol. 15, pp. 1735-1746, 1997

[MMA20] N. Bouguila and W. Fao, Mixture Models and Applications, Springer, 2020

[Mo15] A. Morea et al., "Throughput Comparison Between 50-GHz and 37.5-GHz Grid Transparent Networks," IEEE/OSA Journal of Optical Communication and Networking (JOCN), vol. 7, pp. A293-A300, 2015

[Mo18] R. Morais and J. Pedro, “Machine Learning Models for Estimating Quality of Transmission in DWDM Networks,” IEEE/OSA Journal of Optical Communications and Networking, vol. 10, pp. D84-D99, 2018

[Mo21] R. Morais, “On the Suitability, Requisites, and Challenges of Machine Learning [Invited],” IEEE/OSA Journal of Optical Communications and Networking, vol. 13, pp. D1-D12, 2021

[Mu00] B. Mukherjee, “WDM optical communication networks: progress and challenges,” IEEE Journal on Selected Areas in Communications, vol. 18, pp. 1810-1824, 2000

[Na14] A. Napoli et al., "Reduced Complexity Digital Back-Propagation Methods for Optical Communication Systems," IEEE/OSA Journal of Lightwave Technology, vol. 32, pp. 1351-1362, 2014

[NFO13] G. Agrawal, *Nonlinear Fiber Optics*, Academic Press, 5th edition, 2013

[OCATA] M. Ruiz, L. Velasco, and D. Sequeira, “Optical Constellation Analysis (OCATA)”, https://doi.org/10.34810/data146, V2, 2022

[OWN06] B. Mukherjee, Optical WDM Networks, Springer, 2006

[Pa18] F. Paolucci, A. Sgambelluri, F. Cugini and P. Castoldi, "Network Telemetry Streaming Services in SDN-Based Disaggregated Optical Networks," IEEE/OSA Journal of Lightwave Technology, vol. 36, pp. 3142-3149, 2018

[Pa20] F. Paolucci, R. Emmerich, A. Eira, N. Costa, J. Pedro, P. Berenguer, C. Schubert, J. Fischer, F. Fresi, A. Sgambelluri, and F. Cugini, “Disaggregated edge-enabled C+L-band filterless metro networks,” IEEE/OSA Journal of Optical Communications and Networking, vol. 12, pp. 2-12, 2020

[Pe20] J. Pedro, N. Costa and S. Pato, “Optical Transport Networks Design beyond 100 Gbaud,” IEEE/OSA Journal of Optical Communication and Networking, vol. 12, pp. A123-A134, 2020

[Po12] P. Poggiolini, “The GN Model of Non-Linear Propagation in Uncompensated Coherent Optical Systems,” IEEE/OSA Journal of Lightwave Technology, vol. 30, pp. 3857-3879, 2012

[Po14] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, “The GN-Model of Fiber Non-Linear Propagation and its Applications,” IEEE/OSA Journal of Lightwave Technology, vol. 32, pp. 694-721, 2014

[Po21] Y. Pointurier, “Machine Learning Techniques for Quality of Transmission Estimation in Optical Networks,” IEEE/OSA Journal of Optical Communications and Networking, vol. 13, pp. B60-B71, 2021

[Pr19] R. Proietti, X. Chen, K. Zhang, G. Liu, M. Shamsabardeh, A. Castro, L. Velasco, Z. Zhu, and S. J. Ben Yoo, “Experimental Demonstration of Machine Learning-aided QoT Estimation in Multi-domain Elastic Optical Networks with Alien Wavelengths,” (Invited paper) IEEE/OSA Journal of Optical Communications and Networking, vol. 11, pp. A1-A10, 2019

[Pu11] C. Pulikkaseril et al., "Spectral Modeling of Channel Band Shapes in Wavelength Selective Switches," OSA Optics Express, vol. 19, pp. 8458-8470, 2011

[Qi14] M. Qiu et al., "Subcarrier multiplexing using DACs for fiber nonlinearity mitigation in coherent optical communication systems," IEEE/OSA Optical Fiber Communication Conference, 2014

[Ra14] T. Rahman, et al., "On the Mitigation of Optical Filtering Penalties Originating From ROADM Cascade," IEEE Photonics Technology Letters, vol. 26, pp. 154-157, 2014

[Ra16] T. Rahman, D. Rafique, B. Spinnler, A. Napoli, M. Bohn, A. Koonen, C. Okonkwo, and H. Waardt, "Digital Subcarrier Multiplexed Hybrid QAM for Data-rate Flexibility and ROADM Filtering Tolerance," IEEE/OSA Optical Fiber Communication Conference, 2016

[Ra17] T. Rahman, Flexible and high data-rate coherent optical transceivers. Diss. Ph. D. Thesis. Technische Universiteit Eindhoven, 2017

[Ra18] D. Rafique and L. Velasco, "Machine Learning for Optical Network Automation: Overview, Architecture and Applications," IEEE/OSA Journal of Optical Communications and Networking, vol. 10, pp. D126-D143, 2018

[Ro18] C. Rottondi, L. Barletta, A. Giusti, and M. Tornatore, "Machine-Learning Method for Quality of Transmission Prediction of Unestablished Lightpaths," IEEE/OSA Journal of Optical Communications and Networking, vol. 10, pp. A286-A297, 2018

[Sa15] N. Sambo et al. "Next generation sliceable bandwidth variable transponders." IEEE Communications Magazine, vol. 53, pp. 163-171, 2015

[Sa20] W. Saif, M. Esmail, A. Ragheb, T. Alshawi and S. Alshebeili, "Machine learning techniques for optical performance monitoring and modulation format identification: A Survey," IEEE Communications Surveys & Tutorials, vol. 22, pp. 2839-2882, 2020

[SCS06] M. Jeruchim, P. Balaban, and K. Shanmugan, Simulation of Communication Systems: Modeling, Methodology, and Techniques, Springer Science & Business Media, 2006

[Se18] D. Sequeira, L. Cancela and J. Rebola, "Impact of Physical Layer Impairments on Multi-Degree CDC ROADM-based Optical Networks," IEEE Optical Network Design and Modeling, 2018

[Se19] H. Sepehrian, J. Lin, L. A. Rusch and W. Shi, "Silicon Photonic IQ Modulators for 400 Gb/s and Beyond," IEEE/OSA Journal Lightwave Technology, vol. 37, pp. 3078-3086, 2019

[Sh14] J. Shao, X. Liang and S. Kumar, "Comparison of Split-Step Fourier Schemes for Simulating Fiber Optic Communication Systems," IEEE Photonics Journal, vol. 6, pp. 1-15, 2014

[Sh18] B. Shariati, M. Ruiz, A. Sgambelluri, F. Cugini, and L. Velasco, “Real-time Spectrum Surveillance in Filterless Optical Networks,” IEEE/OSA Optical Fiber Communication Conference, 2018

[Sh19] B. Shariati, M. Ruiz, J. Comellas, and L. Velasco, “Learning from the Optical Spectrum: Failure Detection and Identification,” IEEE/OSA Journal Lightwave Technology, vol. 37, pp. 433-440, 2019

[Sh20] B. Shariati, M. Ruiz, F. Fresi, A. Sgambelluri, F. Cugini, and L. Velasco, “Real-time Optical Spectrum Monitoring in Filterless Optical Metro Networks,” Photonic Network Communications, vol.40, pp. 1-13, 2020

[Sp19] B. Spinnler et al., “Autonomous intelligent transponder enabling adaptive network optimization in live network field trial,” IEEE/OSA Journal Optical Communications and Networking, vol. 11, pp. C1-C9, 2019

[St10] T. Strasser and J. Wagener, "Wavelength-Selective Switches for ROADM Applications," IEEE Journal of Selected Topics in Quantum Electronics, vol. 16, pp. 1150-1157, 2010

[Su20] H. Sun et al., "800G DSP ASIC Design Using Probabilistic Shaping and Digital Sub-Carrier Multiplexing," IEEE/OSA Journal of Lightwave Technology, vol. 38, pp. 4744-4756, 2020

[Ta04] M. Taylor, "Coherent detection method using DSP for demodulation of signal and subsequent equalization of propagation impairments," IEEE Photonics Technology Letters, vol. 16, pp. 674-676, 2004

[Ta19] Q. Tan, A. Yang, and P. Guo, “Blind Modulation Format Identification Using Differential Phase and Amplitude Ratio,” IEEE Photonics Journal, vol. 11, pp. 1-12, 2019

[Th16] A. Thyagaturu, A. Mercian, M. McGarry, M. Reisslein and W. Kellerer, "Software Defined Optical Networks (SDONs): A Comprehensive Survey," IEEE Communications Surveys & Tutorials, vol. 18, pp. 2738-2786, 2016

[Uz21] D. Uzunidis, M. Presi, A. Sgambelluri, F. Paolucci, A. Stavdas, and F. Cugini, "Bidirectional Single-Fiber Filterless Optical Networks: Modeling and Experimental assessment," IEEE/OSA Journal of Optical Communication and Networking, vol. 13, pp. C1-C9, 2021

[Ve13] L. Velasco, P. Wright, A. Lord, and G. Junyent, “Saving CAPEX by Extending Flexgrid-based Core Optical Networks towards the Edges [Invited],” IEEE/OSA Journal of Optical Communication and Networking (JOCN), vol. 5, pp. A171-A183, 2013

[Ve18] L. Velasco, A. Sgambelluri, R. Casellas, Ll. Gifre, J.-L. Izquierdo-Zaragoza, F. Fresi, F. Paolucci, R. Martínez, and E. Riccardi, “Building Autonomic Optical Whitebox-based Networks,” IEEE/OSA Journal of Lightwave Technology, vol. 36, pp. 3097-3104, 2018

[Ve19] L. Velasco et al., “Monitoring and Data Analytics for Optical Networking: Benefits, Architectures, and Use-Cases,” IEEE Network Magazine, vol. 33, pp. 100-108, 2019

[Ve22] L. Velasco, S. Barzegar, F. Tabatabaeimehr, and M. Ruiz, “Intent-Based Networking for Optical Networks [Invited Tutorial],” IEEE/OSA Journal of Optical Communications and Networking, vol. 11, pp. A11-A22, 2022

[Vel18] A. P. Vela, B. Shariati, M. Ruiz, F. Cugini, A. Castro, H. Lu, R. Proietti, J. Comellas, P. Castoldi, S. J. B. Yoo, and L. Velasco, “Soft Failure Localization during Commissioning Testing and Lightpath Operation [Invited],” IEEE/OSA Journal of Optical Communications and Networking, vol. 10, pp. A27-A36, 2018

[Vi22] R. Vilalta et al., "Architecture to Deploy and Operate a Digital Twin Optical Network," IEEE/OSA Optical Fiber Communication Conference, 2022

[Wa13] J. Wang, C. Xie and Z. Pan, “Matched Filter Design for RRC Spectrally Shaped Nyquist-WDM Systems,” IEEE Photonics Technology Letters, vol. 25, pp. 2263-2266, 2013

[Wa17] D. Wang, M. Zhang, J. Li, Z. Li, J. Li, C. Song, and X. Chen, "Intelligent constellation diagram analyzer using convolutional neural network-based deep learning," Optics Express, vol. 25, pp. 17150-17166, 2017

[Wa21] D. Wang, Z. Zhang, M. Zhang, M. Fu, J. Li, S. Cai, C. Zhang, and X. Chen, “The Role of Digital Twin in Optical Communication: Fault Managements, Hardware Configuration, and Transmission Simulation,” IEEE Communications Magazine, vol. 59, pp. 133-139, 2021

[We21] D. Welch et al., "Point-to-Multipoint Optical Networks Using Coherent Digital Subcarriers," IEEE/OSA Journal of Lightwave Technology, vol. 39, pp. 5232-5247, 2021

[Xi17] X. Xiao, M. Li, L. Wang, D. Chen, Q. Yang, and S. Yu, "High Speed Silicon Photonic Modulators," IEEE/OSA Optical Fiber Communication Conference, 2017

[Xi19] Q. Xiang, Y. Yang, Q. Zhang and Y. Yao, "Joint and Accurate OSNR Estimation and Modulation Format Identification Scheme Using the Feature-Based ANN," IEEE Photonics Journal, vol. 11, pp. 1-11, 2019

[Za19.1] T. Zami, I. Ruiz, A. Ghazisaeidi, and B. Lavigne, “Growing impact of optical filtering in future WDM networks,” IEEE/OSA Optical Fiber Communication Conference, 2019

[Za19.2] T. Zami, B. Lavigne, and M. Bertolini, “How 64 GBaud Optical Carriers Maximize the Capacity in Core Elastic WDM Networks With Fewer Transponders per Gb/s,” IEEE/OSA Journal of Optical Communication and Networking, vol. 11, pp. A20-A32, 2019

[Zh18] F. Zhang et al., "Blind Adaptive Digital Backpropagation for Fiber Nonlinearity Compensation," IEEE/OSA Journal of Lightwave Technology, vol. 36, pp. 1746-1756, 2018

[Zh19] J. Zhao, Y. Liu, and T. Xu, “Advanced DSP for Coherent Optical Fiber Communication, Applied Sciences, vol.9, 2019

[Zha19] Q. Zhang et al., “A Simple Joint Modulation Format Identification and OSNR Monitoring Scheme for IMDD OOFDM Transceivers Using K-Nearest Neighbor Algorithm,” MDPI Applied Sciences, vol. 9, pp. 1 – 12, 2019

[Zhu19] Q. Zhuge et al., "Application of Machine Learning in Fiber Nonlinearity Modeling and Monitoring for Elastic Optical Networks," IEEE/OSA Journal of Lightwave Technology, vol. 37, pp. 3055-3063, 2019