

# POLITECNICO DI TORINO

Master's Degree in Communications and Computer  
Networks Engineering



**Politecnico  
di Torino**

Master's Degree Thesis

## Digital performance monitoring of coherent optical fiber communication links

Supervisors

Prof. Gabriella BOSCO

Prof. Roberto GAUDINO

Dott. Dario PILORI

Candidate

Lorenzo ANDRENACCI

October 2022



# Summary

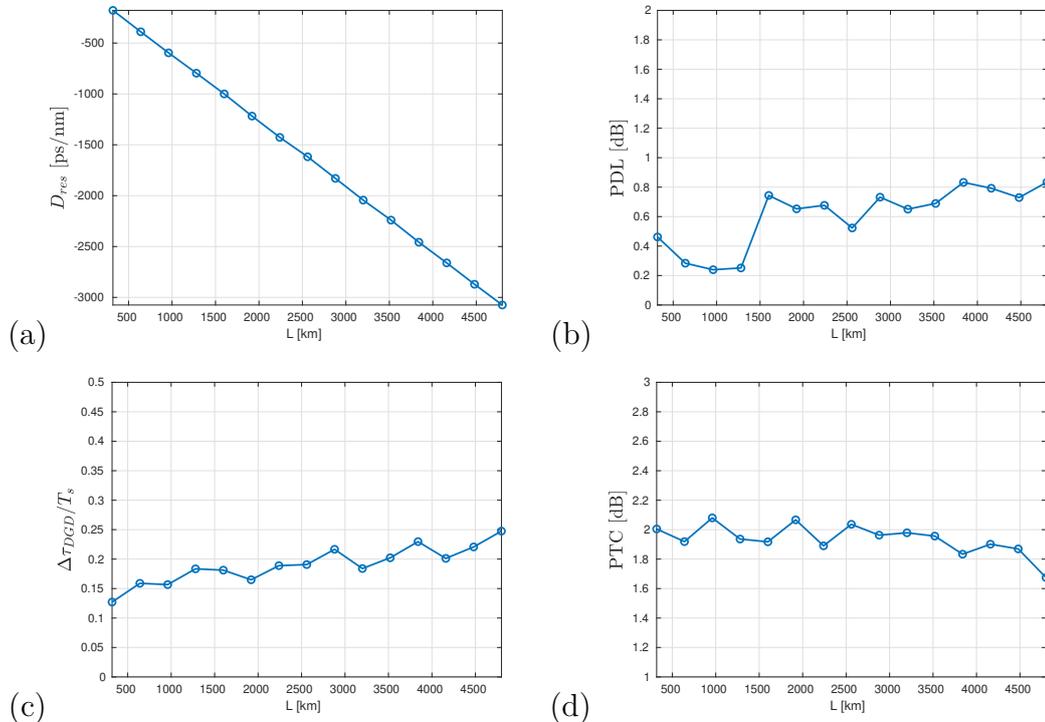
The main focus of this thesis is on the implementation and testing of digital Optical Performance Monitoring (OPM) algorithms applied to optical fiber communication links, in the scenario of dynamically reconfigurable all-optical networks. What characterizes this type of networks is a high flexibility with adaptive routing of optical paths and switching of optical wavelengths. Hence, flexibility and bandwidth efficiency are a core characteristic. In particular, they are obtained exploiting high-order modulation formats and Digital Signal Processing (DSP) techniques in digital coherent receivers.

In the first part of the thesis, the important role of the digital coherent receiver and DSP is explained, so as to highlight the fact that they are enablers to the OPM techniques that are the subject of this work. Particular attention is given to the digital filtering stage. More in detail, this stage is composed of a fixed Chromatic Dispersion (CD) compensator and an adaptive equalizer, both implemented as Finite Impulse Response (FIR) filters. As for the adaptive equalization, not only does it compensate for all the linear channel impairments, e.g., residual CD, Polarization Mode Dispersion (PMD), Polarization Dependent Loss (PDL) and potential bandwidth limitations, but also allows to get an estimation of them. Indeed, the first digital OPM technique presented manages to extract from the tap coefficients of the equalizer an estimate of the aforementioned impairments.

What makes this possible is the fact that the filter impulse response of the equalizer relates to the inverse channel impulse response. Thus, under the assumption of a weakly non-linear regime, it is possible to separate each individual contribution simply through the processing of such tap coefficients. The advantage of this approach lies in the fact that it does not require external devices, such as Optical Spectrum Analyzers (OSA) or other tapping devices, granting a cost efficient and reliable in-service estimation of impairments.

A few examples of results are presented, implementing the OPM algorithm in MATLAB. Therefore, at first results have been obtained in a simulation environment only, then, in order to have a more robust validation, also results obtained through post-processing of data collected in real long-haul transmission experiments

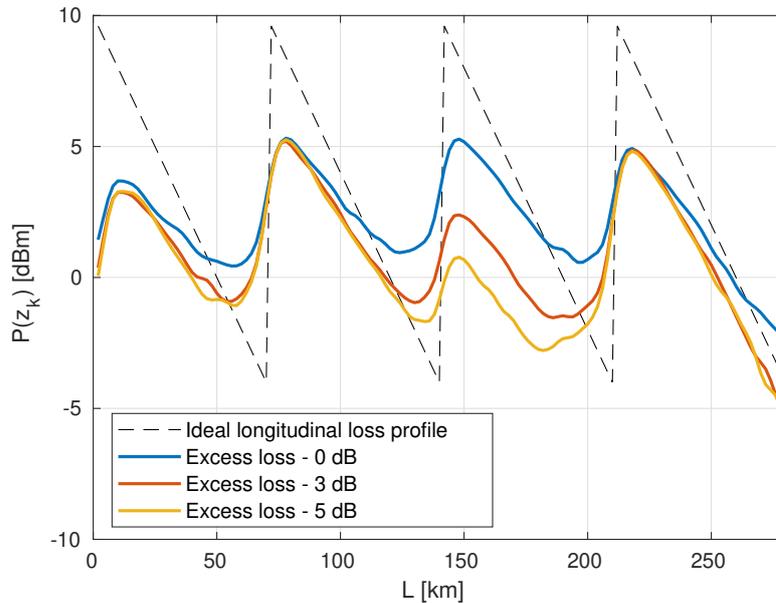
(performed in the Photonlab at Politecnico di Torino) are displayed. Figure 1 shows some of the results obtained post-processing the experimental data.



**Figure 1:** Monitoring result over a 4820-km transmission distance for (a) residual CD (b) PDL (c) DGD (d) bandwidth limitation or filtering detuning.

The second part of the thesis aims at taking a further step. The quantities estimated in the first part of the thesis are cumulative quantities. However, also the monitoring of fiber longitudinal parameters, such as loss profiles, CD maps and responses of individual link components (filters and amplifiers), is essential to guarantee the reliability and capacity maximization of dynamic optical networks. Being able to monitor those profiles solely relying on the data available in the digital receiver would lead to a “smarter” system, able to detect anomalies spatially. The objective of the second part of the thesis is then to design and implement an OPM algorithm able to perform a “distributed” estimation of the transmission parameters along the line, on the basis of the information available in the receiver. These techniques go under the definition of Digital Longitudinal Monitoring (DLM). In particular, what makes this possible is the “non-commutativity” of linear and non-linear operators and the intrinsic similarity between the structures of neural networks (NN) and the Digital Back-Propagation (DBP) algorithm. The DLM algorithm is based on the implementation of a Channel Reconstruction Method (CRM), which generates

a virtual copy of the actual non-linear transmission channel in the digital domain and tries to estimate the longitudinal loss profile of the optical link. The received signal is sent into a Channel Emulator which implements the DBP algorithm, a cost function is computed by comparing the obtained samples with the output of the CRM block and the parameters are updated according to the Stochastic Gradient Descent (SGD) algorithm in an iterative way. Then, the algorithm has been tested in a simulation environment. Figure 2 shows simulation results obtained for the longitudinal loss profile estimation and anomaly localization.



**Figure 2:** Estimated longitudinal loss profiles when an excess loss of 0 dB (solid blue line), 3 dB (solid orange line) and 5 dB (solid yellow line) is inserted at the beginning of the third span (at a distance of 140 km from the transmitter end). Also the ideal loss profile (dashed black line) is reported

In its most basic form, the CRM has proved to be a valid technique to move from a cumulative to a distributed OPM approach for optical fiber communication systems. Indeed, it managed to successfully extract some useful longitudinal quantities, i.e., the loss profiles and the localized excess losses. Of course, this represents just a starting point for these DLM applications, which have a lot of potential still to develop.

# Acknowledgements

*To everyone who supported me.*



# Table of Contents

<b>List of Tables</b>	IX
<b>List of Figures</b>	X
<b>Acronyms</b>	XV
<b>1 Introduction</b>	1
<b>2 Optical fiber communication</b>	3
2.1 Fundamentals . . . . .	3
2.1.1 LASER . . . . .	6
2.1.2 Photodiode . . . . .	10
2.1.3 Optical amplifier . . . . .	12
2.1.4 Optical fiber . . . . .	14
2.2 Advantages of fiber communication . . . . .	16
2.3 Fiber impairments . . . . .	18
2.3.1 Attenuation . . . . .	19
2.3.2 Optical noise . . . . .	19
2.3.3 Chromatic Dispersion . . . . .	20
2.3.4 Polarization Mode Dispersion . . . . .	22
2.3.5 Polarization Dependent Loss . . . . .	23
2.3.6 Non-linear effects . . . . .	23
2.3.7 Manakov Equation . . . . .	25
2.4 Evolution of optical fiber communication systems . . . . .	26
<b>3 Digital coherent receiver</b>	29
3.1 Principle of coherent optical detection . . . . .	30
3.1.1 Configuration . . . . .	30
3.1.2 Heterodyne detection . . . . .	31
3.1.3 Homodyne detection . . . . .	33

3.1.4	Phase and polarization diversity . . . . .	33
3.2	DSP in coherent receivers . . . . .	36
3.3	Linear impairments mitigation . . . . .	38
3.3.1	Chromatic dispersion compensation . . . . .	40
3.3.2	Adaptive equalization . . . . .	41
3.4	Non-linear impairments mitigation . . . . .	42
3.4.1	Digital back-propagation . . . . .	43
3.4.2	Simulation results . . . . .	44
3.4.3	Experimental results . . . . .	47
<b>4</b>	<b>Optical performance monitoring</b>	<b>51</b>
4.1	Introduction . . . . .	51
4.2	Digital Optical Performance Monitoring . . . . .	52
4.2.1	Channel Model . . . . .	53
4.2.2	PDL estimation . . . . .	55
4.2.3	Residual CD estimation . . . . .	56
4.2.4	System scalar transfer function estimation . . . . .	57
4.2.5	DGD estimation . . . . .	57
4.3	Simulation results . . . . .	57
4.4	Experimental results . . . . .	65
<b>5</b>	<b>Digital longitudinal monitoring</b>	<b>70</b>
5.1	Introduction . . . . .	70
5.2	Basic concepts . . . . .	71
5.2.1	Neural networks . . . . .	72
5.2.2	Gradient descent algorithm . . . . .	74
5.2.3	Stochastic gradient descent algorithm . . . . .	75
5.3	Distributed optical performance monitoring . . . . .	75
5.3.1	NN-DBP parallelism . . . . .	76
5.3.2	Problem formulation . . . . .	77
5.3.3	Longitudinal loss profile estimation . . . . .	79
5.3.4	Non-commutativity property . . . . .	83
5.4	Simulation results . . . . .	84
5.5	Conclusions . . . . .	87
<b>6</b>	<b>Future works</b>	<b>88</b>
6.1	Issues of CRM and possible improvements . . . . .	88
6.2	CD maps and multiple filters' responses extraction with CRM . . . . .	89
6.3	An alternative approach: the correlation method . . . . .	91
6.4	Conclusions . . . . .	91
	<b>Bibliography</b>	<b>92</b>

# List of Tables

3.1	Maximum reach when applying no DBP and DBP with different number of steps per span, along with the maximum channel power associated to it. Also the gain with respect to the case in which no DBP is performed is reported. . . . .	47
-----	---	----

# List of Figures

1	Monitoring result over a 4820-km transmission distance for (a) residual CD (b) PDL (c) DGD (d) bandwidth limitation or filtering detuning. . . . .	iii
2	Estimated longitudinal loss profiles when an excess loss of 0 dB (solid blue line), 3 dB (solid orange line) and 5 dB (solid yellow line) is inserted at the beginning of the third span (at a distance of 140 km from the transmitter end). Also the ideal loss profile (dashed black line) is reported . . . . .	iv
2.1	The electromagnetic spectrum showing the region used for optical fiber communications, as in [1]. . . . .	4
2.2	(a) The general communication system. (b) The optical fiber communication system, as in [1]. . . . .	5
2.3	Energy state diagram showing: (a) absorption; (b) spontaneous emission; (c) stimulated emission, as in [1]. . . . .	6
2.4	The basic laser structure incorporating plane mirrors, as in [1]. . . . .	7
2.5	The ideal light output against current characteristic for an injection laser, as in [1]. . . . .	8
2.6	Responsivity against wavelength characteristics: (a) an ideal silicon photodiode; (b) silicon, germanium and InGaAs photodiodes with quantum efficiencies also shown, as in [1]. . . . .	11
2.7	Gain–bandwidth characteristics of different optical amplifiers, as in [1].	12
2.8	Energy level structure of $Er^{3+}$ in glass fiber host. EDFA can be pumped either at 980 or at 1480 nm, as in [2]. . . . .	13
2.9	Optical fiber waveguide showing the core of refractive index $n_1$ , surrounded by the cladding of slightly lower refractive index $n_2$ , as in [1] . . . . .	14
2.10	Light rays incident on a high to low refractive index interface (e.g. glass–air): refraction, as in [1] . . . . .	15
2.11	The acceptance angle $\theta_a$ when launching light into an optical fiber, as in [1] . . . . .	16

2.12	Example of attenuation spectrum for a single-mode fiber (solid line) with the attenuation spectra for some of the loss mechanisms contributing to the overall fiber attenuation (dashed and dotted lines), as in [1] . . . . .	19
2.13	Single-mode fiber dispersion characteristics: comparing the profiles for nonzero-dispersion-shifted fiber (NZ-DSF) and negative-dispersion fiber (NDF) with standard single-mode fiber (SSMF), as in [1]. . . . .	22
2.14	Time domain effect of polarization mode dispersion in a short fiber length with a pulse being launched with equal power on the two birefringent axes, x and y, becoming two pulses at the output separated by the differential group delay, as in [1]. . . . .	23
2.15	Evolution of commercial optical transmission systems over the past 30 years and extrapolations for the coming 20 years, as in [4]. . . . .	26
3.1	Configuration of the coherent receiver that measures the beat between the signal and the LO. Balanced detection eliminates the dc component and maximizes the beat signal, as in [7]. . . . .	32
3.2	Spectra of (a) the optical signal and (b) the down-converted IF signal. By setting the LO frequency close to the signal frequency, we can obtain the IF spectrum shown in (b), as in [7]. . . . .	32
3.3	Comparison of the electrical spectra at the optical detector output for homodyne and heterodyne detection of a PSK signal, as in [1] . . . . .	33
3.4	Configuration of the phase-diversity homodyne receiver using a 90° optical hybrid. The in-phase component and the quadrature component of the signal complex amplitude are obtained simultaneously, as in [7]. . . . .	34
3.5	Configuration of the homodyne receiver employing phase and polarization diversities. Two phase-diversity homodyne receivers are placed in the polarization-diversity configuration, where a common LO is employed, as in [7]. . . . .	35
3.6	Basic configuration of the digital coherent receiver. Four outputs from the homodyne receiver consisting of phase and polarization diversities are sent to the ADC/DSP circuit. These outputs include all the information of the complex amplitude of the signal, and linear impairments are compensated for in the DSP core, as in [7]. . . . .	36
3.7	Block scheme of DSP chain in coherent receivers, as in [8] . . . . .	37
3.8	Block scheme of the digital filtering stage in a DSP chain. The two blocks can also be exchanged, thanks to system linearity. As in [8] . . . . .	39
3.9	Basic configuration of a FIR filter, as in [7] . . . . .	39

3.10	CD compensation block scheme. IDFT is the inverse DFT converting the signal from the frequency to the time domain. $H$ is the transfer function of the filter. As in [7]. . . . .	41
3.11	Maximum reach for a WDM system with 5 channels and 64 GBaud symbol rate, when using DBP algorithm at receiver side, for different values of $P_{ch}$ and $N_{step}$ . . . . .	46
3.12	Set-up of the experiment, as in [12]. Note that in this case the spans are $4 \times 80$ -km SSMF. . . . .	48
3.13	GMI versus distance when no DBP (blue) and DBP with 2 step/span (orange) is applied. . . . .	49
3.14	GMI versus channel power $P_{ch}$ after 5 recirculations. The optimal power in this case is -2 dBm. . . . .	50
4.1	PDL estimation with OPM algorithm for a nominal value of 3 dB. . . . .	59
4.2	Residual CD $\beta_{2,res}$ estimation with OPM algorithm for a nominal value of -212.25 ps <sup>2</sup> . . . . .	60
4.3	Electro-optical system transfer function estimation. . . . .	60
4.4	Power spectral density of the received signal (blue solid line) and the supergaussian filter (red dashed line), with a detuning of $\Delta f_{detun} = 10$ GHz. . . . .	61
4.5	Electro-optical system transfer function estimation when a detuning of $\Delta f_{detun} = 10$ GHz is applied. . . . .	62
4.6	Top: power spectral density of the received signal (blue solid line) and the supergaussian filter (red dashed line) without detuning. Bottom: Electro-optical system transfer function estimation without detuning. . . . .	63
4.7	DGD estimation normalized with respect to $T_s$ obtained with OPM algorithm, along with its mean value ( $0.55 \cdot T_s$ ) and the nominal value $1 \cdot T_s$ . The PMD parameter $D_p$ has been set to $0.7812 \text{ ps}/\sqrt{\text{km}}$ . . . . .	64
4.8	Residual CD $D_{res}$ estimated over 15 recirculations (4800 km) from experimental data. . . . .	65
4.9	Mean PDL estimated over 15 recirculations (4800 km) from experimental data with OPM algorithms. . . . .	66
4.10	Mean DGD estimated over 15 recirculations (4800 km) from experimental data with OPM algorithms. . . . .	67
4.11	PTC estimated over 15 recirculations (4800 km) from experimental data with OPM algorithms. . . . .	68
4.12	System scalar transfer function $ H^{-1}(f) ^2$ estimated over 15 recirculations (4800 km) from experimental data with OPM algorithms. . . . .	68
5.1	Perceptron structure, with inputs $x_i$ , weights $w_i$ and bias $b$ . . . . .	72

5.2	Generic structure of a neural network with multiple neuron layers, also said feed-forward neural network. . . . .	73
5.3	Similarity of DBP and NN structure. Cascaded blocks iteratively implementing a linear operation (CD compensation/linear combination of weights and biases) and a non-linear operation (non-linear phase rotation compensation (NLPR)/activation function). Note: Sigmoid and ReLu are just examples of activation function. Taken from [16]. . . . .	76
5.4	Behavior of $\alpha(z)$ and its integral along a multi-span optical system. The spikes in the curve representing $\alpha(z)$ are given by the amplifiers' gains $g_s$ at the end of each span. Taken from [15]. . . . .	78
5.5	(a) DLM algorithm scheme and the involved data structures for the estimation of $\gamma'(z_k)$ . In particular, in this example waveforms are composed of $N_{samp} = 1024$ samples, a batch of 60000 waveforms and a mini-batch of $N_{wf} = 100$ waveforms. (b) More detailed structure of DBP-based channel emulator. Note: the complex-valued FIR filter is not considered in this work. Taken from [15]. . . . .	80
5.6	Scheme of the learning process in the DLM algorithm, as in [15]. . .	83
5.7	Estimated longitudinal loss profile (solid blue line) and the ideal longitudinal loss profile (dashed black line) after 600 iterations with step $\Delta z = 2$ km and learning rate $\eta = 0.1$ . . . . .	85
5.8	Estimated longitudinal loss profiles when an excess loss of 0 dB (solid blue line), 3 dB (solid orange line) and 5 dB (solid yellow line) is inserted at the beginning of the third span (at a distance of 140 km from the transmitter end). Also the ideal loss profile (dashed black line) is reported. . . . .	86
5.9	Deviation of 3-dB loss profile (blue line) and 5-dB loss profile (orange line) from the 0-dB one. Ideal deviation for 3 dB case (dashed blue line) and 5 dB case (dashed red line) are also reported. . . . .	87
6.1	(a) Estimated profile over a $4 \times 70$ -km optical link for different step sizes, according to the rule $h(z) \propto P(z)^p$ . (b) Step size for each case. (c) Estimation error for each $p$ . Taken from [18]. . . . .	89
6.2	(a) Estimated $\beta_2(z_k)$ profile for a different number of SGD iterations when a DSF span is inserted between SSMF spans. (b) Maximum value reached by $\beta(z_k)$ versus number of iterations. Taken from [15]	90
6.3	Estimated FIR filters responses in a $3 \times 50$ -km system with an optical bandpass filter at each node when a detuning of +15 GHz is inserted (b) in the first filter and (c) in the second filter. Taken from [15] . .	90
6.4	In-situ PPE algorithm scheme, as in [17]. . . . .	91



# Acronyms

**LED**

Light Emitting Diode

**LASER**

Light Amplification by Stimulated Emission of Radiation

**MLM**

Multi Longitudinal Mode

**SLM**

Single Longitudinal Mode

**DFB**

Distributed FeedBack

**ECL**

External Cavity Laser

**EDFA**

Erbium Doped Fiber Amplifier

**ASE**

Amplified Spontaneous Emission

**PSD**

Power Spectral Density

**CD**

Chromatic Dispersion

**SSMF**

Standard Single Mode Fiber

**PSCF**

Pure Silica Core Fiber

**NZDSF**

Non-Zero Dispersion-Shifted Fiber

**DCF**

Dispersion Compensating Fiber

**NDF**

Negative Dispersion Fiber

**WDM**

Wavelength Division Multiplexing

**PMD**

Polarization Mode Dispersion

**PDL**

Polarization Dependent Loss

**ISI**

Inter-Symbol Interference

**DGD**

Differential Group Delay

**NL**

Non-Linear

**NLSE**

Non-Linear Schrödinger Equation

**SPM**

Self-Phase Modulation

**XPM**

Cross-Phase Modulation

**FWM**

Four-Wave Mixing

**IM-DD**

Intensity Modulation Direct Detection

**ADC**

Analog-To-Digital Converter

**DAC**

Digital-To-Analog Converter

**DSP**

Digital Signal Processing

**PSK**

Phase-Shift Keying

**QAM**

Quadrature-Amplitude Modulation

**DBP**

Digital Back-Propagation

**LO**

Local Oscillator

**PC**

Polarization Controller

**DC**

Direct Current

**PBS**

Polarization Beam Splitter

**FEC**

Forward Error Correction

**BICM**

Bit Interleaved Coded Modulation

**FIR**

Finite Impulse Response

**DFT**

Discrete Fourier Transform

**FFT**

Fast Fourier Transform

**ASIC**

Application Specific Integrated Circuit

**CMA**

Constant Modulus Algorithm

**LMS**

Least Mean-Squared

**SSFM**

Split-Step Fourier Method

**SRRC**

Square-Root Raised Cosine

**GMI**

Generalized Mutual Information

**BPS**

Blind Phase Search

**OPM**

Optical Performance Monitoring

**OSA**

Optical Spectrum Analyzer

**OTDR**

Optical Time-Domain Reflectometer

**PTC**

Peak-To-Center

**SI**

System Identification

**DLM**

Digital Longitudinal Monitoring

**NN**

Neural Network

**CRM**

Channel Reconstruction Method

**GD**

Gradient Descent

**SGD**

Stochastic Gradient Descent

**AI**

Artificial Intelligence

**MSE**

Mean Squared Error

**NLPR**

Non-Linear Phase Rotation

# Chapter 1

## Introduction

In the scenario of all-optical networks, increasing flexibility and bandwidth efficiency are key issues. Moreover, the possibility of automatizing network operations is highly appreciated, since it has a direct effect also on the operating expenditures related to the network infrastructure. In this sense, **Optical Performance Monitoring (OPM)** plays a role of utmost importance. Indeed, it provides functionalities to monitor the "health" status of the optics in the network and potentially detect anomalies. Of course, this monitoring needs to be performed on a large-scale optical network, whence the requirement for cost-effective and easily-deployable OPM techniques. Furthermore, if these techniques could be performed by solely leveraging the information which is present at the receiver side, it would be possible to avoid the employment of external monitoring devices, which should be scattered over the whole network, and improve system automatization. OPM algorithms and techniques are indeed expected to play an increasing role in next generation elastic optical networks, since they can help realize several key functionalities, such as the dynamic and adaptive adjustment of modulation formats for the maximization of transparent reach, and the optimization of routing and wavelength assignment algorithms.

OPM is exactly the subject of this work. In particular, its digital implementation, which is made possible by a fundamental system element, i.e., the **digital coherent receiver** and the **digital signal processing (DSP)** techniques implemented in it. Let us start, then, by giving an outline of the topics discussed in this work and their organization.

In Chapter 2, the concept of optical fiber communication and its main technological enablers are introduced. After that, the main characteristics of propagation over optical fibers are described, along with all the effects which impair the propagating signals. At the end, a brief overview of the evolution of optical fiber communication

systems is presented.

In Chapter 3, a description of the digital coherent receiver's working principles and structure is given. Besides, its important role as OPM enabler is also explained, together with the DSP chain of operations inside of it. Particular attention is given to the **digital filtering** stage, including Chromatic Dispersion (CD) compensation and adaptive equalization. Finally, the Digital Back-Propagation (DBP) algorithm is introduced, together with some simulation and experimental results related to its application.

In Chapter 4, the OPM concept is discussed more in detail. Several techniques to extract the major linear system impairments are introduced. Among these impairments, one may find CD, Polarization Mode Dispersion (PMD), Polarization Dependent Loss (PDL) and potential bandwidth limitations or frequency detuning. Also in this case, both simulation and experimental results are reported.

In Chapter 5, a step forward is taken. Indeed, up to this point the OPM techniques presented are able to extract only **cumulative** quantities. Here, **longitudinal** monitoring techniques are also discussed. In particular, a Channel Reconstruction Method (CRM) is presented, whose main enablers are the fiber non-linearity and the intrinsic similarity of the DBP structure with neural networks (NN). In particular, the target of the algorithm's estimation is the longitudinal loss profile of the optical link and potential anomaly detection. To conclude, simulation results are reported.

In Chapter 6, the future works over this topic are illustrated. More in detail, possible solutions to the issues which have come up during CRM application and more results which is possible to obtain are indicated. Furthermore, a proposed alternative method to CRM is also briefly described.

## Chapter 2

# Optical fiber communication

The process of transferring information between two or more points is generally referred to as **communication**. Information is modulated in order to be carried by an electromagnetic wave over the **communication system** and reach its destination. Then it is demodulated through specific techniques which attempt to retrieve the original message.

When the electromagnetic carrier is chosen from the optical range of frequencies and the communication system is composed of optical fibers, we refer to the whole communication process as **optical fiber communication**.

In this chapter an overview on the fundamentals of this type of communication systems is presented. This includes a brief description of the technological enablers, their principle of operation and the major transmission impairments in an optical fiber link. Finally, such systems and their evolution are discussed in a historical perspective.

The main reference for this chapter is [1].

## 2.1 Fundamentals

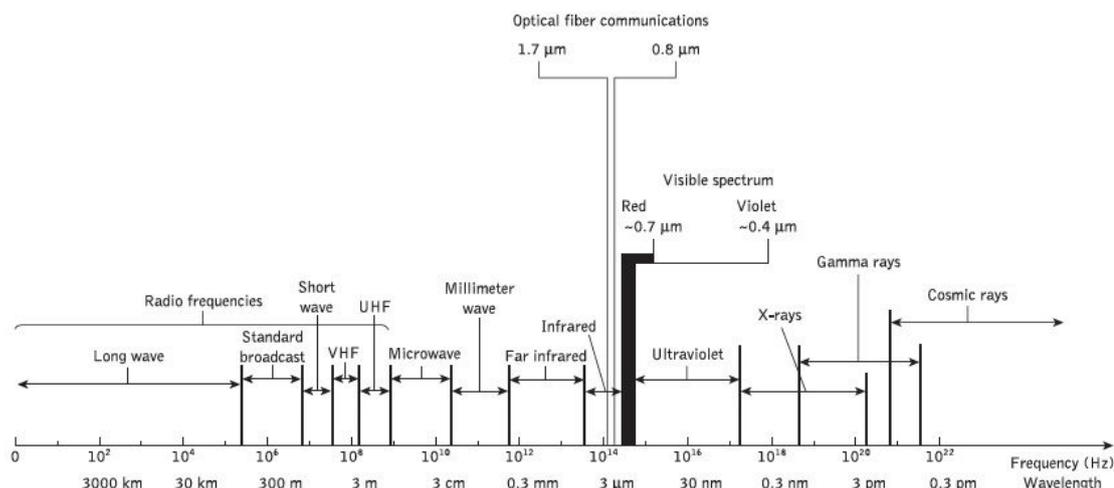
Optical systems use light as carrier waves. Similarly to radio and microwave frequency waves, light is also an electromagnetic wave whose frequency is relatively higher. Indeed, for fiber systems the central frequencies which are generally used are near the visible light spectrum.

However, in optical communications also wavelength is commonly used. The relation between the two is

$$\lambda = \frac{c}{f} \tag{2.1}$$

where  $\lambda$  is the wavelength,  $c$  is the speed of light in vacuum and  $f$  is the frequency.

Hence, the optical range can be defined as the wavelengths from  $0.8\ \mu\text{m}$  to  $1.7\ \mu\text{m}$ , which is fully contained within the **infrared region**. Figure 2.1 gives a graphical representation of the spectrum.

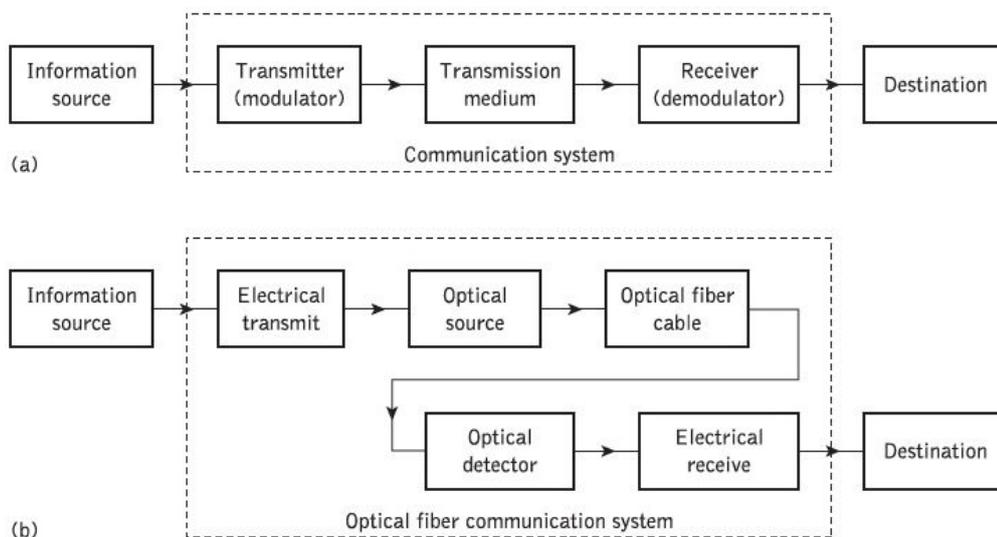


**Figure 2.1:** The electromagnetic spectrum showing the region used for optical fiber communications, as in [1].

Of course, the choice of the central frequency affects the system for what concerns both the technological solutions adopted and the performance of the system itself. These aspects will be highlighted when needed.

What is more, optical systems do not differ in principle from any of the other telecommunication systems. In a very simplified way, they are still composed of a transmitter, a transmission medium over which the signal propagates and a receiver. Figure 2.2 illustrates the schematic of a general communication system (a) and of an optical fiber communication system (b). The latter is basically the former, but with a greater level of detail. It is possible to notice that for the transmission part, a conversion from the electrical domain to the optical one is needed. The dual at the receiver side, since processing is generally performed in the electrical domain. What is more, the transmission medium has to be capable of transporting light, i.e., an optical fiber.

It is of utmost importance to have devices in the system which are capable of performing all of the aforementioned functions.



**Figure 2.2:** (a) The general communication system. (b) The optical fiber communication system, as in [1].

What makes this type of approach possible are the so-called technological enablers. The most important ones are:

- **semiconductor-based laser** technology;
- **semiconductor-based photodiodes**;
- **optical amplifier** technology;
- **optical fiber** technology;

The first three technologies are all strictly related to the energy levels of the materials used and the interaction of light with them.

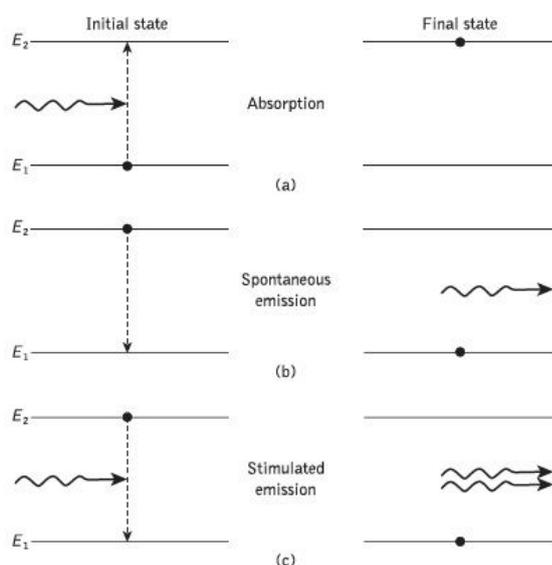
From quantum physics it is possible to learn that light interacts with matter by means of packets of energy, i.e., **photons**. Furthermore, electrons in an atom can only exist in specific discrete energy states. Thus, when light interacts with atoms, electrons (or carriers) can move from one energy state to the other.

Every time a carrier performs a transition from an energy level  $E_1$  to an energy level  $E_2$ , the frequency of the radiation emitted/absorbed is related to the difference between those energy states, i.e.,:

$$\Delta E = E_2 - E_1 = hf \tag{2.2}$$

where  $h = 6.626 \times 10^{-34} \text{ J} \cdot \text{s}$  is Planck's constant. The three main physical effects that are possible are:

- **absorption:** when a carrier moves from a lower energy level  $E_1$  to a higher one  $E_2$  due to the interaction with an incoming photon. The photon is then absorbed and an electrical current is generated;
- **spontaneous emission:** when a carrier decays from a higher energy level  $E_2$  to a lower one  $E_1$ , generating a photon whose frequency is related to Eq. (2.2);
- **stimulated emission:** when a carrier decays from a higher energy level  $E_2$  to a lower one  $E_1$  due to the interaction with an incoming photon. The generated photon has the same frequency and phase of the input photon, i.e., it is **coherent**;



**Figure 2.3:** Energy state diagram showing: (a) absorption; (b) spontaneous emission; (c) stimulated emission, as in [1].

All of the mechanisms showed in Figure 2.3 are at the basis of optical sources, photodiodes and optical amplification. As for optical fibers, their principle of operation is different and will be illustrated in the following.

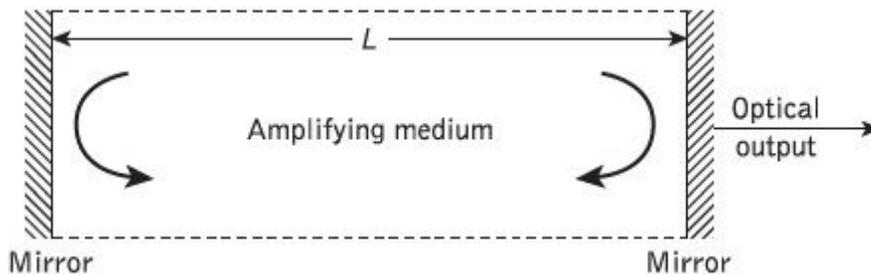
### 2.1.1 LASER

As mentioned at the beginning of Sec. 2.1, optical fiber communication systems use light as carrier waves. Therefore, the system needs devices which are capable of generating light and couple it into the optical fiber.

Several options are available, depending on the type of source needed. Among the most popular ones it is possible to find **LED**'s (Light Emitting Diode) for incoherent optical sources and **LASER**'s (Light Amplification by Stimulated Emission of Radiation) for coherent optical sources. Since the former fall outside the scope of this discussion, only the latter will be taken into consideration and commonly referred to as laser.

A laser is an opto-electronic device that converts an electrical current into an optical power. Its principle of operation is similar to that of electronic oscillators. A current is injected into the laser structure (a semiconductor P-I-N junction) in order to bring about a carrier inversion, i.e., to bring a large number of carriers into the conduction band. At this point, light generation takes place according to the mechanisms introduced at the end of Sec. 2.1. It is important to mention the fact that if no **optical feedback** is present in the structure, the dominant effect is spontaneous emission. However, the light wave needs to be coherent and monochromatic. This is the reason why an additional element has to be introduced in the structure, so as to control the spontaneous emission and have **stimulated emission** as dominant effect. Otherwise it would be a LED.

Figure 2.4 shows the most basic laser structure to achieve this feedback mechanism.



**Figure 2.4:** The basic laser structure incorporating plane mirrors, as in [1].

Indeed, the mirror elements create an optical cavity which traps the photons through partial reflection. While moving, the photons get amplified thanks to stimulated emission inside the amplifying medium. Apart from the amplification gain  $G_{active}$ , though, photons also experience passive losses due to attenuation and absorption effects, and wavelength selective losses due to the cavity filtering effect. Let  $L(\lambda)$  be the total passive wavelength-selective loss, if the total gain  $G(\lambda)$  seen by a photon with frequency  $f$  or, equivalently, wavelength  $\lambda$  is

$$G(\lambda) = G_{active}L(\lambda) = 1 \quad (2.3)$$

the laser is said to be **lasing** for that wavelength. On the other hand, if such condition is not satisfied, then only spontaneous emission is present and the laser

behaves as a LED.

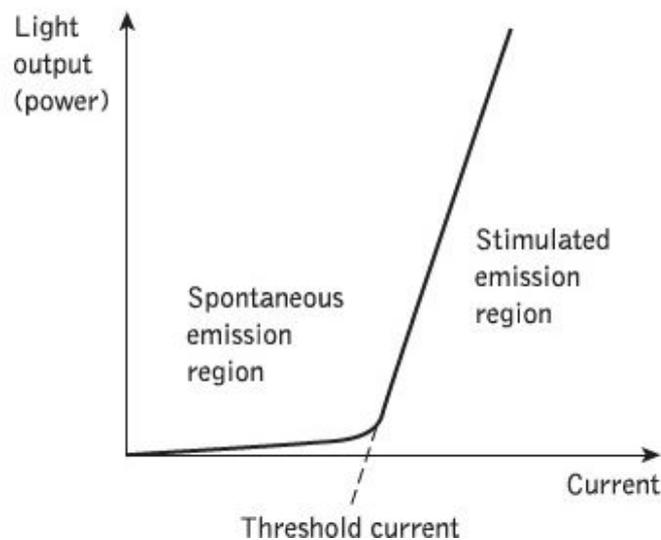
Let us recall Eq. (2.2). The wavelength of the photons generated through stimulated emission is related to the difference between the energy states of the material. Hence, the choice of the central frequency influences, as already anticipated in Sec. 2.1, the technological solution of the system itself. A few examples are:

- **GaAs** for 810 nm - 870 nm (first window);
- **InGaAsP** for 1000 nm - 1650 nm (second window);

Since one of the most important windows in today's optical communication systems is around 1550 nm, lasers are commonly made using InGaAsP.

To conclude this brief description about lasers, a set of parameters to characterize and classify different types of lasers are introduced.

First of all, the P-I characteristic gives information about the optical power generated for a defined value of the input current.



**Figure 2.5:** The ideal light output against current characteristic for an injection laser, as in [1].

Figure 2.5 pinpoints the different regions of operation of a laser. It operates in the spontaneous emission region, i.e., no lasing and LED-like behavior, when the condition expressed by Eq. (2.3) is not satisfied. Indeed, low power is generated and the output spectrum of the signal results to be very large, thus very far from being monochromatic.

When the input current is higher than a specific value, called **threshold current**  $I_{th}$ , then the lasing condition is satisfied and the laser operates in the stimulated emission region. It is necessary to say, then, that lasers always operate above threshold. In this condition, defining  $I_{in}(t)$  as the input current,  $P_{out}(t)$  as the output power and  $K$  as the slope of the stimulated emission region characteristic (which is ideally linear), it is possible to write the mathematical expression which relates current and power as

$$P_{out}(t) = K \cdot (I_{in}(t) - I_{th}) \quad (2.4)$$

Finally, lasers are also classified according to the optical spectrum they emit when driven by a continuous-wave current, i.e., fixed current. Let us recall that the bandwidth of an optical signal can be expressed both in terms of frequencies ( $\Delta f$ ) and wavelengths ( $\Delta\lambda$ ). The relation between the two can be approximated by means of the Taylor expansion of Eq. (2.1), truncated at the first order, as

$$\Delta f \approx \frac{c}{\lambda^2} \Delta\lambda \quad (2.5)$$

where  $\lambda$  is the considered central wavelength. Minus sign is ignored since bandwidth is always positive. In this work it is always equal to  $\lambda = 1550$  nm if not specified. It is possible to distinguish basically three classes of optical sources:

- **LED:** incoherent source and large bandwidth. Typical values are in the range  $\Delta\lambda = 30$  nm - 100 nm, or  $\Delta f = 3.75$  THz - 11 THz;
- **Multi Longitudinal Mode (MLM) Laser Diode:** several narrow coherent spectral lines. Bandwidth in the range  $\Delta\lambda = 10$  nm, or  $\Delta f = 1.25$  THz;
- **Single Longitudinal Mode (SLM) Laser Diode:**  $\Delta\lambda \leq 0.01$  nm, or  $\Delta f \leq 1.25$  GHz. Typical values for distributed feedback (DFB) lasers are  $\Delta f = 10$  MHz - 30 MHz;

Nowadays, External Cavity Lasers (ECL) are mostly used for high-performance coherent optical systems. They have an external optical filter which allows to have a very narrow spectrum at the output, in the range of 150 kHz.

## 2.1.2 Photodiode

If a laser is in charge of converting electrical current into optical power, a photodiode is its dual device. Indeed, it receives an optical power and turns it into an electrical current. The most common type is the P-I-N photodiode.

The working principle is simple. Photodiodes are semiconductor P-I-N junctions on which photons impact. A certain amount of those photons are absorbed and carriers manage to move from a lower to a higher energy state. The movement of carriers in the conduction band subjected to an electromagnetic field then results in an electrical current.

It is possible to write the relation between the received optical power  $P_{in}(t)$  and the output electrical current  $I_{out}(t)$  in the form

$$I_{out}(t) = R \cdot P_{in}(t) \quad (2.6)$$

Of course, ideally it would be nice if for each received photon, one electron is created. However, this is not true. The proportional term  $R$  is called **responsivity** and depends on a physical quantity called **quantum efficiency**  $\eta_{ph}$ . The latter is adimensional and defined as the number of electrons generated on average for each impacting photon. It is by definition lower than 1 for P-I-N photodiodes and greater than 1 for another type of photodiode called **avalanche photodiode**. Overall, the expression for  $R$  is

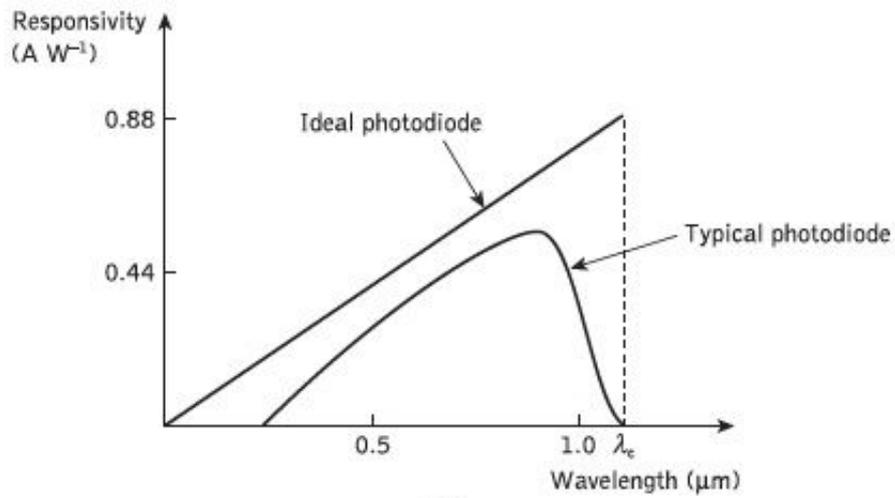
$$R = \frac{q}{hf} \eta_{ph} \quad (2.7)$$

where  $q = 1.6022 \cdot 10^{-19} C$  is the electron charge

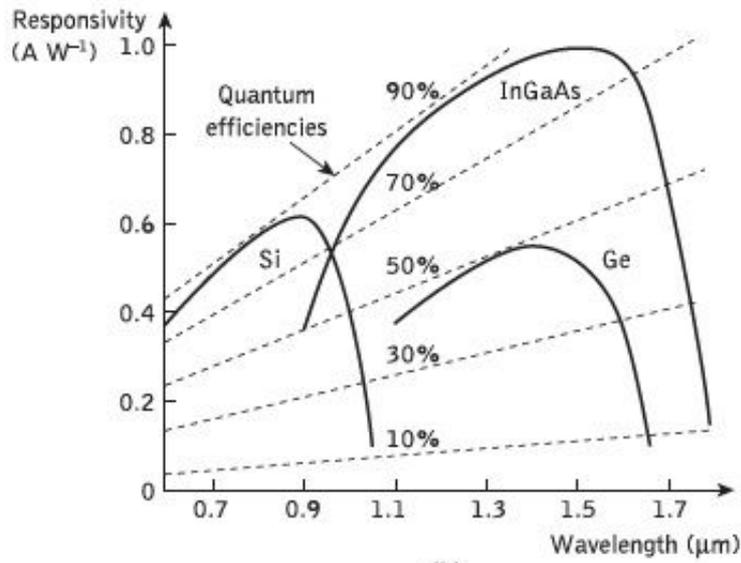
Photodiodes are another example of how the choice of the central frequency and the technological solution influence the performance. The responsivity is not at all a constant quantity. In fact, it depends on both the material used to make the photodiode itself and the central frequency at which it operates.

Figure 2.6 (a) shows both the ideal and realistic silicon photodiode responsivity for different values of the wavelength. On the other hand, Figure 2.6 (b) shows the responsivity values for different values of the wavelength, as well as of different materials. Quantum efficiency is also reported.

By taking an operating wavelength of  $\lambda = 1550$  nm, for instance, it is possible to notice that it is more convenient to use a photodiode made of InGaAs rather than one made of Ge.



(a)



(b)

**Figure 2.6:** Responsivity against wavelength characteristics: (a) an ideal silicon photodiode; (b) silicon, germanium and InGaAs photodiodes with quantum efficiencies also shown, as in [1].

To conclude, the role of the photodiode is extremely important in terms of performance. Hence, it is suggested that such devices be characterized by a high sensitivity at the operating frequency, low noise introduction and stable characteristics.

### 2.1.3 Optical amplifier

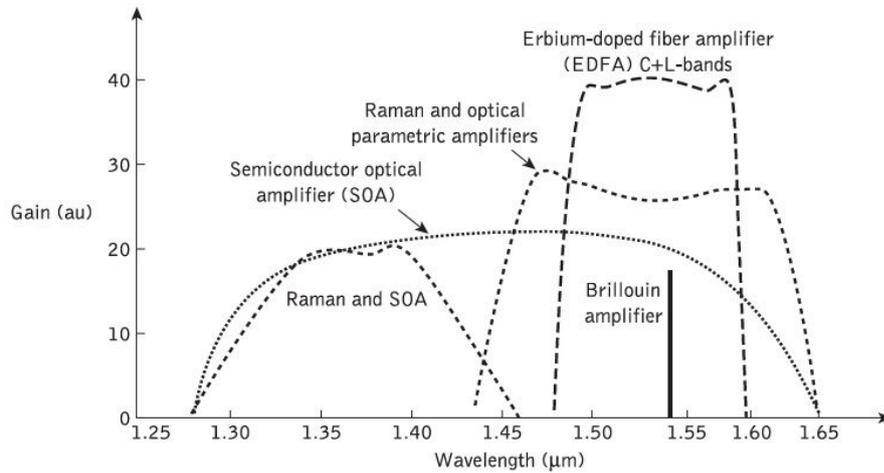
At this point, the important role of both lasers and photodiodes within an optical fiber communication system is clear. However, it has to be mentioned that those devices represent also a limiting factor. Indeed, the conversion of the signal from the electrical domain to the optical domain and vice versa may cause a reduction in the operating bandwidth, as well as a loss in signal quality.

This happens, for example, in intermediate nodes of the system where Optical/-Electrical (O/E) and Electrical/Optical (E/O) interfaces are present. In those points, the signal may be either processed or simply forwarded. Hence, to avoid all the limitations associated to those interfaces and reduce the costs, devices capable of operating on the signal in the optical domain (when it is not strictly mandatory to convert the signal into the electrical domain) are needed. This is the case of **optical amplifiers**, which allow to linearly amplify the optical signal directly in the optical domain, without converting it into the electrical one. Let  $E_{in}(t)$  and  $E_{out}(t)$  be, respectively, the input and output optical fields to/from the amplifier, it is possible to write the relation between them as

$$E_{out}(t) = \sqrt{G} \cdot E_{in}(t) \quad (2.8)$$

where  $G$  is the **optical power gain**.

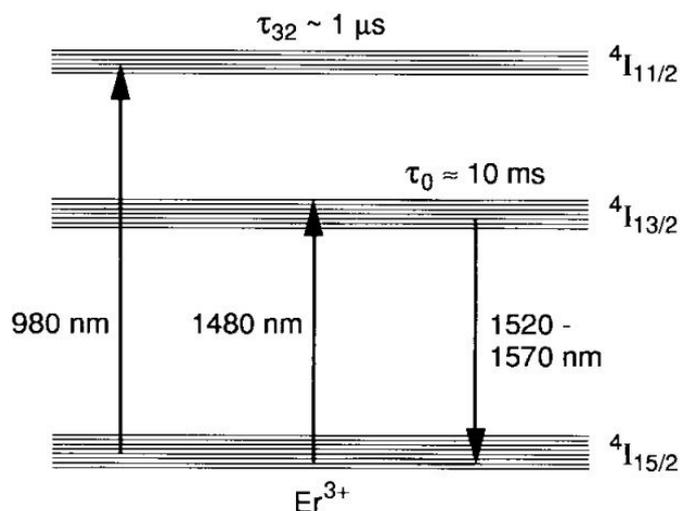
Optical amplifiers come in different types. Just to cite a couple of them, there are **semiconductor optical amplifiers**, i.e., P-N junctions, and **fiber amplifiers**, which use optical fibers previously doped with some material in order to exploit the stimulated emission mechanism and other scattering effects.



**Figure 2.7:** Gain–bandwidth characteristics of different optical amplifiers, as in [1].

Figure 2.7 supplies a few examples of gain spectra for different types of optical amplifiers.

For the usual 1550 nm window the **Erbium Doped Fiber Amplifier (EDFA)** is used. It consists of a silica fiber which is doped with erbium (Er) ions. Figure 2.8 shows the resulting energy levels of the structure.



**Figure 2.8:** Energy level structure of  $Er^{3+}$  in glass fiber host. EDFA can be pumped either at 980 or at 1480 nm, as in [2].

The working principle is the same as for the lasers. Photons are injected into the doped fiber so as to have carrier inversion and move the carriers to the excited energy levels. This operation is defined **pumping** and is generally performed with an unmodulated laser source. After pumping, the carriers will tend to go back to the lower energy levels and generate photons at a wavelength which is proportional to the energy gap. The first two excited levels of the structure allow for the absorption of photons at  $\lambda = 980$  nm and  $\lambda = 1480$  nm, which are also the possible pumping wavelengths.

The amplification mechanism takes place when, along with the pumping photons, also an input optical signal propagates within the fiber. If the stimulated emission mechanism is such that the generated photons are coherent with those of the input signal (e.g., at  $\lambda = 1550$  nm), then the overall effect is that of having an amplified optical signal. Stimulated emission, though, is not the only mechanism present. Hence, the overall working principle can be summed up as an exchange of energy between the pumping signal and the input signal.

Why are EDFA's used for the 1550 nm window, though? At first sight, it may

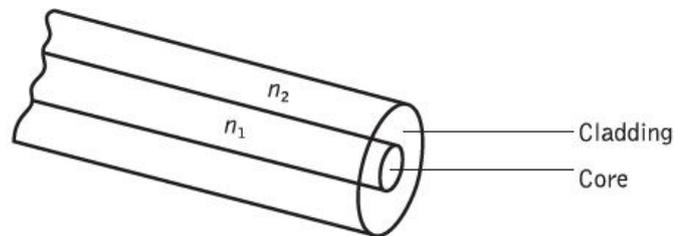
seem that the generated photons can only be at  $\lambda = 1480$  nm, but this is not entirely true. The reason lies in a physical effect known as **Stark effect** which causes a broadening of the energy levels and allows for amplification over a much larger range of wavelengths. This is also the reason why in Figure 2.8 the wavelength of the emitted photons includes the range 1520 - 1570 nm.

This concludes the discussion over the first three technological enablers, which more or less share the same working principles. It is now time to go on to describe the last key element of an optical fiber communication system, i.e., the **optical fiber**.

### 2.1.4 Optical fiber

The optical fiber is the system element which allows the signal (or light) to propagate between its source and destination. It is a cylindrical dielectric waveguide usually made of very pure glass ( $SiO_2$ ) and with very small dimensions (a diameter around  $125\ \mu\text{m}$ ). The fiber is composed by two sections. The inner section is called **core**, while the outer one is **cladding**. In particular, the core is characterized by a refractive index  $n_1$  which is slightly higher (usually less than 1%) than the refractive index  $n_2$  of the cladding and this is performed by doping the core with selected impurities, e.g, Ge.

Figure 2.9 highlights these elements in a piece of fiber.



**Figure 2.9:** Optical fiber waveguide showing the core of refractive index  $n_1$ , surrounded by the cladding of slightly lower refractive index  $n_2$ , as in [1]

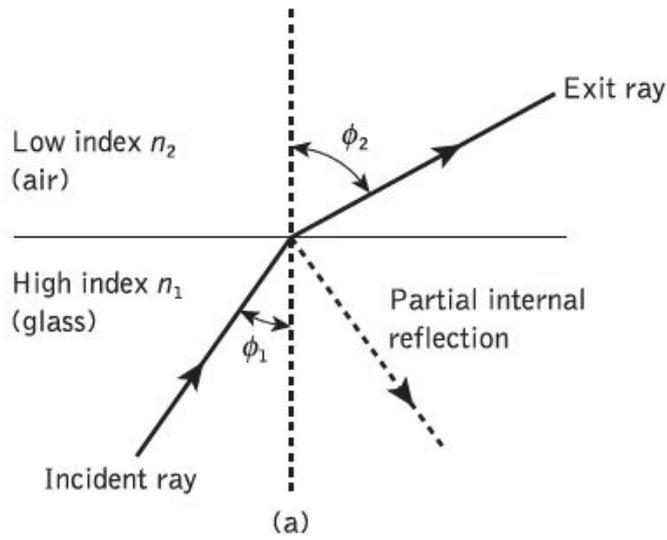
Indeed, refractive index is a crucial quantity in order to be able to understand the working principle of optical fibers. It basically defines the interaction of the electromagnetic field with a certain material. This interaction can be defined in terms of velocity, since a wave propagating in a medium with refractive index  $n$  moves at a velocity

$$v = \frac{c}{n}. \quad (2.9)$$

However, this interaction can also be expressed in terms of reflection and refraction. When a wave impacts on the interface between a medium with refractive index  $n_1$  and one with refractive index  $n_2$ , part of it is generally reflected back and the other part is refracted. Snell's law gives us the relation between the incident wave and the refracted one:

$$n_1 \cdot \sin \phi_1 = n_2 \cdot \sin \phi_2 \quad (2.10)$$

where  $\phi_1$  and  $\phi_2$  are, respectively, the incidence angle and the refraction angle. Figure 2.10 gives a graphical representation of the scenario in which the law is applied, taking as example a glass-air interface.



**Figure 2.10:** Light rays incident on a high to low refractive index interface (e.g. glass-air): refraction, as in [1]

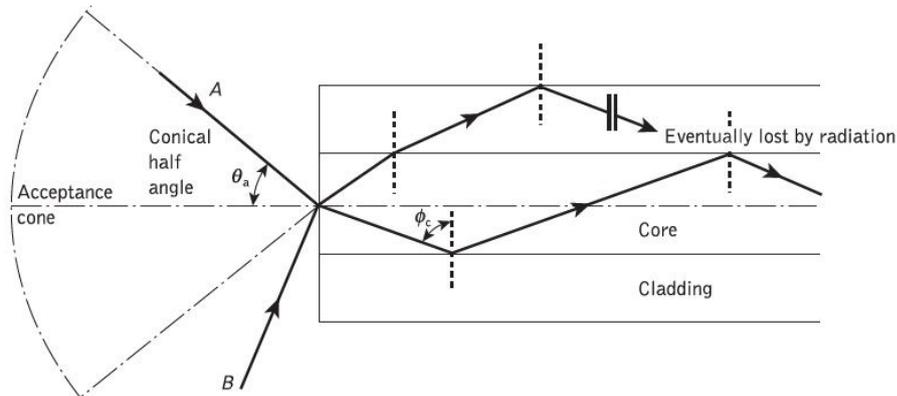
A special case is the case in which  $\phi_2 = \pi/2$ . This condition is called **total internal reflection** and is possible only in the case in which  $n_1 > n_2$ . The wave is completely reflected and no refraction occurs. By substituting  $\phi_2 = \pi/2$  into Eq. (2.10) it is possible to define the **critical angle**  $\phi_c$ , i.e., the minimum incident angle to have total internal reflection. In particular

$$\sin \phi_c = \frac{n_2}{n_1} \quad (2.11)$$

The working principle of optical fibers is based exactly on this phenomenon. The core and the cladding have refractive indices such that when light impacts on the core-cladding interface, it gets totally reflected and remains confined within the fiber while propagating.

Not all rays of light remain confined within the fiber, though. It depends on the **acceptance angle**  $\theta_a$ , defined as the maximum angle of entrance of the light ray into the fiber which grants an incidence angle at the core-cladding interface larger than  $\phi_c$ . Its derivation is of no interest for this discussion, hence only a graphical representation is given in Figure 2.11.

Of course, several types of fiber exist, depending on how they are manufactured. Each of them is characterized by different properties and most of them were mainly thought having in mind an impairment which affects fiber communications called **chromatic dispersion (CD)** and discussed in Sec. 2.3. In particular, such properties depend on the step-index profile, i.e., how the refractive index is distributed on a section of the fiber. For the sake of completeness, the main types are (ITU-T standard in brackets): Standard Single Mode Fiber (**SSMF**, G.652), Non-Zero Dispersion-Shifted Fiber (**NZDSF**, G.655), Dispersion Compensating Fiber (**DCF**, ) and Pure Silica Core (**PSC**, G.656) fibers.



**Figure 2.11:** The acceptance angle  $\theta_a$  when launching light into an optical fiber, as in [1]

## 2.2 Advantages of fiber communication

After going through the main elements of an optical fiber communication system and their functioning, it is important to highlight also the advantages deriving from this type of approach. Being guided propagation systems, a comparison with their equivalent in the electrical domain comes natural. Indeed, when fibers were first proposed in 1970 [3], their main competitors were copper-wired systems, such as twisted-pair-based and coaxial-cable-based systems.

The main advantages of optical fiber communications can be grouped in the following list:

- **Bandwidth:** the potential bandwidth offered by optical fibers is much higher than that of traditional cable systems. In numbers, simply considering the so-called C-band (191 - 196 THz) the available bandwidth is approximately 5 THz, while for copper cables it is in the range of MHz or GHz. In any case, fibers offer an available bandwidth from three up to six orders of magnitude higher. Besides, this also translates into much higher rates. For instance, in 2018 commercial long-haul C+L-band (185 THz - 196 THz) systems managed to carry an aggregate capacity up to around 48 Tb/s [4] by employing **Wavelength Division Multiplexing (WDM)** technologies.
- **Dimensions:** as already mentioned in Sec. 2.1.4, the typical diameter of a fiber is around one tenth of a millimeter, i.e., smaller than a human hair. Despite the fact that some more protections have to be added to the plain structure of the fiber, its dimensions are still very small compared to those of traditional copper cables.
- **Isolation:** fibers are usually made of glass, which is an electrical insulator. Therefore, no earth loop or short circuit can be generated from their usage. Furthermore, this also makes them very good for electromagnetically hazardous environment.
- **Interference and crosstalk:** fibers are dielectric waveguides, which makes them immune to any kind of electromagnetic interference. Plus, they do not irradiate because all photons are confined inside the fiber itself. No crosstalk is, then, present and a large number of fibers can be grouped together.
- **Transmission loss:** optical fibers are characterized by a very low loss. Typical values for fiber attenuation are around 0.2 dB/km. This is an extremely low attenuation which has some positive implication also on the system composition and cost. Indeed, such a low attenuation allows for fewer repeater/amplifiers on the line and, consequently, a lower cost. Finally, it is also important to highlight the fact that optical attenuation does not introduce any noise, which happens for coaxial cable, for instance.

These were some of the most important reasons why fiber optics had and continues to have an enormous success.

However, in order to have a better understanding on optical systems, also the transmission impairments that optical signals face during propagation need to be discussed, since they have a strong implication on how systems are designed.

## 2.3 Fiber impairments

Light is an electromagnetic wave and its behavior is dictated by Maxwell's laws. In a SSMF optical fiber, neglecting non-linearity for now, the field  $E$  obeys the following wave equation while propagating:

$$\frac{\partial}{\partial z}E(\omega, z) = [-j\beta(\omega) - \alpha]E(\omega, z) \quad (2.12)$$

where  $z$  is the distance on the central axis of the fiber,  $\omega = 2\pi f$  is the angular frequency,  $j$  is the imaginary unit,  $\beta$  is the **propagation constant** and  $\alpha$  is the **loss coefficient**. Let us recall the double polarization nature of light in the fiber, conventionally indicated as x-polarization and y-polarization. For simplicity, the field is considered to be on only one polarization for now, hence no subscript is indicated. Also, the field is expressed in the complex envelope notation. The solution of Eq. (2.12) is

$$E(\omega, z) = E(\omega, 0)e^{-j\beta(\omega)z}e^{-\alpha z} \quad (2.13)$$

However, the propagation constant is not actually constant. It depends on several factors which give rise to the different types of impairments. Such factors and the related main impairments are listed in the following:

- **Frequency:** whose main impairment is given by a second-order effect called **CD**. Higher-order effects exist, but their impact is generally smaller than CD;
- **Polarization:** generating **birefringence** and **Polarization Mode Dispersion (PMD)**;
- **Signal power:** generating non-linear distortions;

Each of them will be briefly discussed, in order to have an idea on what their effect is and how it is modelled.

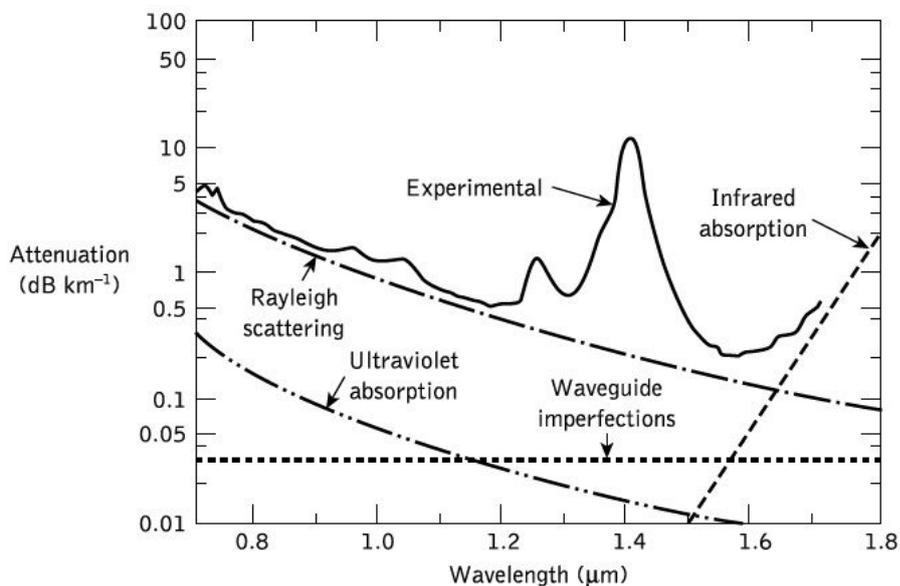
Also, other kind of impairments are introduced, not directly related to the fiber itself, but to components on the link: **Polarization Dependent Loss (PDL)** and **Amplified Spontaneous Emission (ASE) noise**.

In the end, the **Manakov equations** are presented, which supply a mathematical expression for the propagation of the field in the optical fiber, gathering all of the aforementioned impairments.

### 2.3.1 Attenuation

Attenuation is present in any transmission medium and represents a gradual loss in signal power. It is represented by the coefficient  $\alpha$  in Eq. (2.13). Hence, power loss in optical fibers is exponential.

What is interesting about it in this context has already been discussed in Sec. 2.2. Indeed, optical fibers are characterized by very low loss coefficients and in the C-band their value is approximately constant and around 0.2 dB/km. However, attenuation varies with wavelength. Figure 2.12 gives a representation of this, plus the attenuation related to other effects not discussed in this work.



**Figure 2.12:** Example of attenuation spectrum for a single-mode fiber (solid line) with the attenuation spectra for some of the loss mechanisms contributing to the overall fiber attenuation (dashed and dotted lines), as in [1]

A curious fact is that the range with the lowest attenuation in an optical fiber also corresponds to the range which allows the best amplification at the lowest cost, as explained in Sec. 2.1.3.

### 2.3.2 Optical noise

Attenuation does not introduce noise in optical fibers (see Sec. 2.2), but optical noise does exist. While describing the EDFA functioning (see Sec. 2.1.3), it has been mentioned the fact that amplification occurs due to the stimulated emission

mechanism. However, spontaneous emission is also there and is responsible for the **amplified spontaneous emission (ASE) noise**.

Considering also the noise which is introduced at the output of the EDFA in an additive way, Eq. (2.8) needs to be re-written as

$$E_{out}(t) = \sqrt{G} \cdot E_{in}(t) + n(t) \quad (2.14)$$

where  $n(t)$  is the ASE noise.

In the complex envelope notation, it is possible to write it as

$$n(t) = n_I(t) + jn_Q(t) \quad (2.15)$$

where I and Q subscripts indicate the in-phase and quadrature components. Both quadrature components are modelled as white Gaussian random processes with a power spectral density (PSD) equal to

$$\frac{N_0}{2} = \frac{hf}{2}(G - 1)n_{sp} \quad (2.16)$$

where  $h$  is the Planck's constant,  $f$  the central frequency,  $G$  the EDFA power gain and  $n_{sp}$  the **spontaneous emission factor**. The  $n_{sp}$  is always greater than 1 and acts as the noise indicator for EDFA's. A quantity of interest, which is generally used instead of  $n_{sp}$ , is the **noise figure**  $F$  defined as twice the spontaneous emission factor., i.e.,  $F = 2n_{sp}$ .

### 2.3.3 Chromatic Dispersion

Chromatic dispersion is a linear distortion effect. Going back to Eq. (2.12), the propagation constant is frequency dependent, hence it can be expanded according to its Taylor series around a generic central frequency  $f_0$ . Note that  $f_0$  is the central frequency of the considered optical band, e.g., C-band. This operation yields

$$\beta(\omega) = \beta_0 + \beta_1\Delta\omega + \frac{1}{2}\beta_2\Delta\omega^2 + \dots \quad (2.17)$$

and the field can be written as

$$E(\omega, z) = E(\omega, 0)e^{-j\beta_0 z} e^{-j\beta_1 \Delta\omega z} e^{-\frac{1}{2}\beta_2 \Delta\omega^2 z} e^{-j\alpha z} \quad (2.18)$$

by cutting the Taylor expansion at the second order.

The term  $\beta_0$  is frequency independent and is an irrelevant phase term for most applications.  $\beta_1$  is linearly dependent from frequency, hence it does not distort the

signal. It determines the group velocity and just introduces a fixed propagation delay. Finally,  $\beta_2$  is quadratically dependent from frequency and does distort the signal. It is responsible for CD.

Ignoring  $\beta_0$  and  $\beta_1$  and provided that the launched power is not too high, the system can be considered linear and represented by means of a transfer function. So can the effect of CD. Therefore, it is possible to write the transfer function of CD from Eq. (2.18), substituting  $\omega$  with  $2\pi f$ , as

$$H_{CD}(f) = e^{-j2\pi^2\beta_2(f-f_0)^2z} \quad (2.19)$$

Its effect is that of changing the phase of the optical field and, thus, the group delay of the signal in a frequency-selective way. Indeed, the **group delay**  $\tau_g$  is defined as

$$\tau_g(f) = -\frac{1}{2\pi} \frac{\partial}{\partial f} \phi(f) \quad (2.20)$$

being  $\phi(f)$  the phase of  $H_{CD}(f)$ . Substituting the expression of the phase  $\phi(f)$  obtained from Eq. (2.19), the group delay becomes

$$\tau_g(f) = 2\pi\beta_2(f - f_0)z \quad (2.21)$$

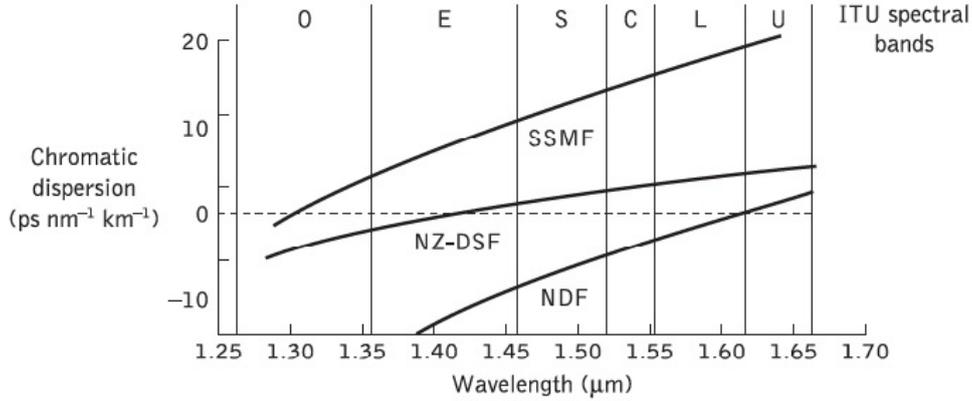
This mathematical expression allows us to better understand the effect of CD. Since the group delay changes linearly with frequency, each frequency component will "see" a different group delay and be delayed by a different amount of time. The final result is that of having a broadened pulse and a consequent increase in inter-symbol interference (ISI), which may make the communication impossible. Furthermore, the quantity  $\beta_2 \cdot z$  is called cumulated CD and represents the fact that the more the signal propagates, the more the effect of CD increases.

It is important to notice that CD is generally associated to the parameter  $\beta_2$ . However, also another parameter is usually used:  $D$ . The relation between the two is the following:

$$D = -\beta_2 \frac{2\pi c}{\lambda^2} \quad (2.22)$$

where  $D$  is in  $[\frac{ps}{nm \cdot km}]$ ,  $c$  in  $[\frac{km}{s}]$ ,  $\lambda$  in  $[nm]$  and  $\beta_2$  in  $[\frac{ps^2}{km}]$ .

To conclude, since CD is one of the major impairments for optical fiber communications, over the years different types of fibers have been manufactured, with the intent of modifying  $\beta_2$  (see Sec. 2.1.4). Figure 2.13 shows typical values for SSMF, NZDSF and DCF (sometimes referred to as Negative Dispersion Fiber, NDF) fibers.



**Figure 2.13:** Single-mode fiber dispersion characteristics: comparing the profiles for nonzero-dispersion-shifted fiber (NZ-DSF) and negative-dispersion fiber (NDF) with standard single-mode fiber (SSMF), as in [1].

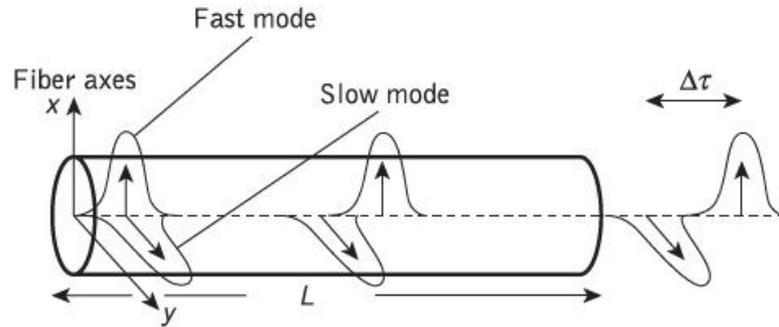
### 2.3.4 Polarization Mode Dispersion

As mentioned at the beginning of Sec. 2.3, the optical fiber has a dual polarization nature, which means that each signal is actually composed of two propagating modes associated to orthogonal polarization states.

Besides, the fiber is also a **birefringent** medium. This is due to the fact that its circular symmetry is broken by random imperfections introduced during the manufacturing stage or external stress. Both these aspects make the fiber an anisotropic medium, meaning that its refractive index varies according to the polarization state of the fields. Equivalently, it is possible to state that the two polarizations of the propagating fields are characterized by different propagation constants.

The birefringence of the optical fiber leads to the first distortion effect depending on the field's polarization state: the **polarization mode dispersion (PMD)**. It is a random effect which introduces a dependency of the group velocity on the polarization state. Hence, the two polarizations travel at different speed and tend to separate in time, leading to an overall pulse broadening. The distance in time of the two polarizations is called **differential group delay (DGD)**  $\Delta\tau_{DGD}$  and is generally expressed in [ps]. It is defined as the difference between the group velocity of the fast and slow mode in the fiber, i.e., the directions in which the group velocity is, respectively, maximum and minimum, over a certain length. Figure 2.14 gives a graphical representation of the effect of the PMD.

Sometimes, another parameter that is used to represent the PMD is  $D_p$ , whose units are [ps/ $\sqrt{km}$ ].



**Figure 2.14:** Time domain effect of polarization mode dispersion in a short fiber length with a pulse being launched with equal power on the two birefringent axes,  $x$  and  $y$ , becoming two pulses at the output separated by the differential group delay, as in [1].

### 2.3.5 Polarization Dependent Loss

It is convenient to discuss a further polarization-dependent impairment, which is not directly related to the optical fiber itself, but which will be taken into consideration in the next chapters: the **Polarization Dependent Loss (PDL)**.

It is caused by optical components on the communication link and defined as the difference between the maximum and minimum insertion loss introduced by such components, when stimulated by all of the possible polarization states. This means that, depending on the polarization state of the field while crossing the component, the signal power gets attenuated by a different amount.

This becomes an issue in long-haul optical communication systems, since the transmission distance is designed according to specific power levels. The random combination of all the PDL effects on the links makes this task much harder to accomplish and causes more margin to be considered, hence worse performances.

### 2.3.6 Non-linear effects

An assumption that has been made while discussing Eq. (2.12) is that non-linearities are neglected. However, when the power of the optical signal increases, some **non-linear (NL) effects** may arise and impair the communication process. Therefore, also those effects need to be introduced.

The most prominent one is the **Kerr effect**. It causes a variation in the refractive index  $n$  of the transmission medium depending on the propagating field power. Its expression can be written as

$$n(t, z) = n_L + \frac{P(t, z)}{A_{eff}} \quad (2.23)$$

where  $n_L$  is the nominal refractive index,  $P(t, z) = |E(t, z)|^2$  is the field power and  $A_{eff}$  is the effective area of the propagating mode (approximately corresponding to the fiber core area).

Eq. (2.23) shows how the refractive index of the medium varies in an instantaneous way. Consequently, this variation causes also a variation in the propagation constant which can be expressed as

$$\Delta\beta(t, z) = \gamma P(t, z) \quad (2.24)$$

where  $\gamma$  is defined as the **non-linear coefficient** and reported in [ $W^{-1}m^{-1}$ ].

For this reason Eq. (2.13) should be corrected so as to consider also this non-linear effect. In particular, it is necessary to convert the equation into the time domain, since the additional term is time dependent. Plus, the Taylor expansion to the second order for  $\beta(\omega)$  is also applied (once again neglecting  $\beta_0$  and  $\beta_1$ ). The equation, then, becomes

$$\frac{\partial}{\partial z} E(t, z) = \underbrace{j \frac{\beta_2}{2} \frac{\partial^2}{\partial t^2} E(t, z)}_{CD} - \underbrace{\alpha E(t, z)}_{Loss} - \underbrace{j \gamma |E(t, z)|^2 E(t, z)}_{NL} \quad (2.25)$$

Eq. (2.25) contains all the terms related to attenuation, CD and NL and is commonly known as **Non-Linear Schrödinger equation (NLSE)**.

In order to understand in a proper way the effects of NL, the  $i$ -th channel of a WDM system is considered, i.e., multiplexing of several channels (at different central frequencies) in the same fiber. The NL term is responsible for the coupling of the frequency components of the signal and its effect is generally cross-talking among channels and generation of new frequency lines. However, all the cross-talking effects it generates are not qualitatively equal. They can be grouped into three main categories:

- **Self-Phase Modulation (SPM)**: it does not create new frequency lines. It is generated by the  $i$ -th channel and acts on the  $i$ -th channel itself. The interference is in the form of a phase shift.
- **Cross-Phase Modulation (XPM)**: also this interference does not generate new frequency lines. It is generated by other channels and acts on the  $i$ -th channel. Again, the interference is in the form of a phase shift, but twice stronger than SPM.

- **Four-Wave Mixing (FWM):** this type of cross-talk generates new frequencies due to the beating of the different channels at different frequencies. To put it simple, the i-th, j-th and k-th channel frequencies interact to generate an l-th channel frequency which was not there before. Even worse, if the channels are equally and closely spaced, the new frequencies fall on the channels' frequencies, mixing coherently with them. Plus, in order to perform this, FWM robs the original signal of his power.

While SPM always exists within a fiber, XPM and FWM are peculiar of WDM, since more than one channel is needed. What is more, as these processes require a phase matching in order to be "efficient", one way to counteract them is using a larger channel spacing or introducing chromatic dispersion into the system. In this sense, NL and CD need to be balanced in order to have a proper system functioning.

### 2.3.7 Manakov Equation

To conclude this discussion, now that all the main impairments have been described, it is important to recall that at the beginning of Sec. 2.3, the dual polarization nature of light has been highlighted. But the signal has been considered on only one polarization (either x or y). Considering both polarizations, the optical field can be written as

$$\mathbf{E}(t, z) = [E_x(t, z) \ E_y(t, z)]^T \quad (2.26)$$

where  $[\ ]^T$  indicates the transposed operator.  $\mathbf{E}(t, z)$  is a **Jones vector**, according to Jones calculus [5], while optical elements are generally defined by means of a **Jones matrix**.

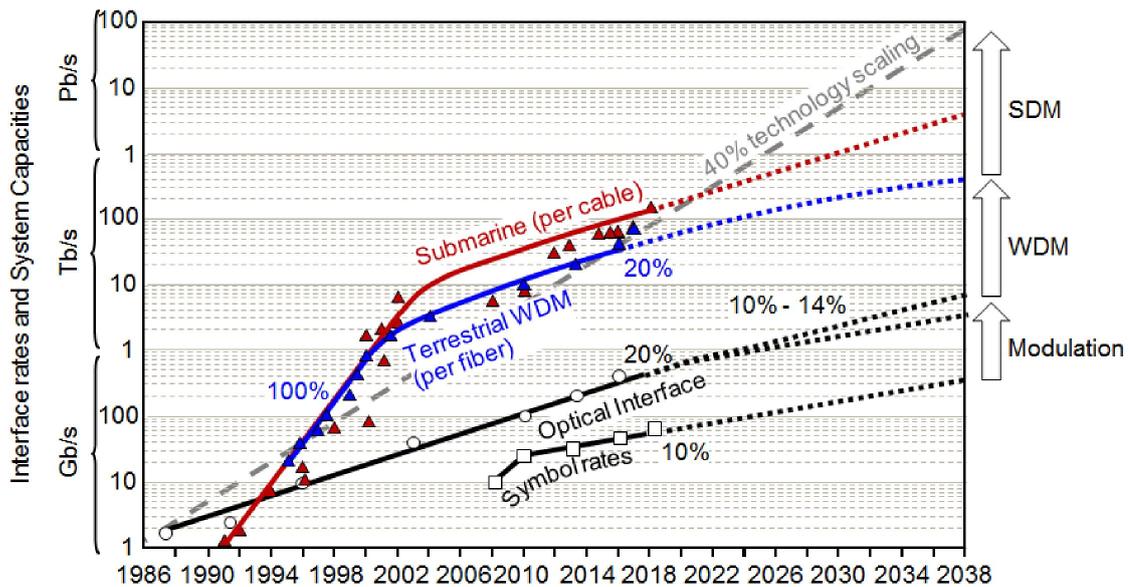
Let us then consider this notation, the stochastic birefringence of optical fibers and the fact that in this notation the power of the field becomes  $P(t, z) = |E_x(t, z)|^2 + |E_y(t, z)|^2$ . Plus, let us average out the stochastic contribution of birefringence. It is possible to re-write Eq. (2.25) in this notation as

$$\frac{\partial}{\partial z} \mathbf{E}(t, z) = j \frac{\beta_2}{2} \frac{\partial^2}{\partial t^2} \mathbf{E}(t, z) - \alpha \mathbf{E}(t, z) - j \gamma \frac{8}{9} \left[ |E_x(t, z)|^2 + |E_y(t, z)|^2 \right] \mathbf{E}(t, z) \quad (2.27)$$

This is a deterministic system of coupled NL differential equations, commonly known as **Manakov equations**. They supply mathematical expressions which describe how the fields propagate within an optical fiber, taking into account all of the impairments discussed so far.

## 2.4 Evolution of optical fiber communication systems

The first time signals were transmitted through optical fibers in a test system was back in 1977 [6]. Ever since, the performance of optical systems has been increasingly improving, also thanks to technological progress related to it. Just to have an idea, Figure 2.15 reports their over the past 30 years and provides a prospect on what the future trends might be.



**Figure 2.15:** Evolution of commercial optical transmission systems over the past 30 years and extrapolations for the coming 20 years, as in [4].

The achievable rates are extremely high and destined to increase even more, up to an aggregated capacity in the range of  $Pb/s$  per cable, i.e.,  $10^{15}$  bits transmitted every second within a cable of optical fibers. It is, then, interesting to have an overview on how such systems have come this far.

Over the years, three main eras can be identified:

- **Era of regeneration:** 1977 - 1995;
- **Era of amplified dispersion-managed systems:** 1995 - 2008;
- **Era of amplified coherent systems:** 2008 - present;

The first long-haul span-by-span optical fiber communication systems were based on opto-electronic regeneration. It has already been discussed how these interfaces

represent a major limitation in performance (see Sec. 2.1.3), since the overall transmission rate would depend on the rate transceivers were capable of supporting. Furthermore, the progress to increase these interface rates was very slow, making them unable to keep up with the increasing transmission rates achieved with techniques such as WDM.

This is the **era of regeneration (1977 - 1995)**, as it has been previously defined.

The game changer was the invention of EDFA's (replacing opto-electronic regeneration), together with the development of techniques which were able to deal with fiber non-linearities.

Indeed, during the first years of 1990s only two types of fibers were available. The first one was the SSMF, characterized by a CD of  $D = 17$  ps/nm/km around  $\lambda = 1550$  nm. CD was the major impairment for high-rate optical communications, limiting transmission distances in the range of few tens of kilometers. This leads to the second type of fiber: DSF. This fiber is characterized by a CD of 0 ps/nm/km around 1550 nm. Ideally, it would solve the problem of CD, but it turned out to be disastrous for what concerns WDM. NL effects were highly enhanced, especially FWM (see Sec. 2.3.6).

For this reason NZDSF fibers were invented. Their CD was not exactly 0 around 1550 nm, but slightly higher or lower.

These fibers also inspired a technique called **dispersion management**, which gives the name to the second era: the **era of amplified dispersion-managed systems (1995 - 2008)**. The idea was that of having a concatenation of fibers characterized by opposite CD signs, i.e., positive and negative. By doing so, the final CD results to be around 0, but in the link the CD is never really 0. The result is a mitigation of both CD and FWM.

This initial approach was successful, but not enough for higher-rate systems (e.g., 20 Gb/s). Plus, the addition of a NZDSF introduced extra losses during the signal propagation. A solution was the invention of DCF fibers, characterized by a much more negative CD than NZDSF's. Hence, DCF's could be inserted as a spool within the EDFA, without the need for two different fiber types on the same line. However, these additions also increased the complexity of the systems, particularly related to the emerging branch of optically-switched networking.

Note that also other impairments were still present. An example is the PMD. It was a problem for fibers produced before 1992, since each channel in a WDM signal required its own PMD compensation at the receiver. A method called **spinning**, though, allowed for the production of optical fibers characterized by a much lower level of PMD. It had the advantage of avoiding PMD compensation just by choosing a fiber with a low enough PMD. This was related to 40 Gb/s systems.

For the advent of 100 Gb/s systems, PMD was no longer an issue, since the introduction of coherent receivers allowed to compensate digitally both CD and PMD.

This is the beginning of the third era: the **era of amplified coherent systems (2008 - present)**. From the technological point of view, the coherent revolution was driven by the evolution of CMOS processing speed, which allowed analog-to-digital converters (ADC) and digital-to-analog (DAC) converters to catch up with 10 GBaud symbol rates.

Coherent receivers have the advantages of both homodyne and heterodyne detection, as will be discussed in Chapter 3. They allow to bring the signal to baseband and retrieve the full optical field, i.e., both its real and imaginary part and for both polarizations. Let us specify that up to this point **Intensity-Modulation Direct-Detection (IM-DD)** were the most popular systems. In these systems, the power of the optical field was modulated and demodulated, no information on the phase of the propagating field. Going coherent, instead, allows to have information also on the field phase and use higher-order modulation formats (e.g., QPSK), with the subsequent increase in spectral efficiency (four times higher, due to the double polarization and quadrature components). This was compliant with the CMOS technology at the time, since 40 Gb/s rates could be supported by 10 Gb/s receivers.

The impact on system designs was immense. Thanks to modern **digital signal processing (DSP)** techniques, coherent receivers are capable of compensating huge amounts of CD, as well as PMD, PDL and even NL. All of this with basically no additional loss. What is more, dispersion-compensated systems were eliminated, going back to fibers characterized by a high CD. In this way, NL effects are reduced, the system is much simpler and all of the impairments are electronically compensated for at the receiver side.

This concludes the overview over what optical fiber communication is and what makes it possible.

In addition, the enabling element of this work has been introduced: the **digital coherent receiver**, discussed more in depth in Chapter 3.

## Chapter 3

# Digital coherent receiver

In this chapter, the **digital coherent receiver** is analyzed more in detail. Its huge impact is by now clear, given by the combination of **coherent detection** and **digital signal processing (DSP)** techniques. Just to cite the main advantages, it allows to move towards higher-order modulation formats, e.g., phase-shift keying (PSK) and quadrature-amplitude modulation (QAM). This is made possible thanks to the ability of efficiently estimating the carrier and phase of the incoming signal in the digital domain. In addition, because the information on the phase of the field is preserved, it is possible to mitigate the impairments the signal suffers from by means of proper DSP techniques.

In order to have a better understanding of all these aspects, the principles of coherent detection will be looked further into, so as to arrive to the description of the phase-diverse and polarization-diverse homodyne receiver, which allows to introduce also the **intradyne receiver**. This type of receiver is the one commonly used today and also the enabler of the techniques which are the subject of this work.

After that, a brief description of the DSP chain implemented within the receiver is supplied. Particular attention is given to the description of the digital filtering techniques, which allow for the compensation of linear impairments, e.g. CD, PMD, PDL and filtering effects along the line.

To conclude, the **digital back-propagation (DBP)** algorithm is presented, which manages to compensate NL effects, too. Together with its theoretical description, some examples of results are presented, both from a simulation and experimental environment.

The main references for this chapter are [7] and [8].

## 3.1 Principle of coherent optical detection

The working principle of coherent detection is that of extracting the complex envelope of the field. This can be performed by combining the signal itself and another continuous-wave signal, called **local oscillator (LO)**.

Two of the main types of coherent detection available are **heterodyne detection** and **homodyne detection**, to which the concepts of **phase-diversity** and **polarization-diversity** will be added at the end, together with the third type of coherent detection: **intradyne detection**.

### 3.1.1 Configuration

In order to understand the basic configuration of the coherent receiver, it is fundamental to go back to the operation performed by photodiodes, i.e., a conversion of the input optical power into an output electrical current.

Let us consider the optical field in its analytic signal notation and on one polarization for simplicity:

$$\hat{E}(t) = E(t)e^{j\omega_s t} \quad (3.1)$$

where  $E(t)$  is the complex envelope and  $\omega_s$  the angular frequency.

When using a photodiode to detect the optical field, its functioning is described by Eq. (2.6). Because the power of the signal is  $P(t) = |\hat{E}(t)|^2$ , all of the information on the phase gets lost. So, in order to detect the whole complex envelope, some modifications are required.

Let us also define the LO in the same notation as

$$\hat{E}_{LO}(t) = E_{LO}e^{j\omega_{LO}t} \quad (3.2)$$

and consider it aligned to the optical field for now. If these two fields are summed together before impinging on the photodiode, then the generated current can be written as

$$I(t) \propto |\hat{E}(t) + \hat{E}_{LO}(t)|^2 \quad (3.3)$$

which can be further expanded as

$$I(t) \propto |\hat{E}(t)|^2 + |\hat{E}_{LO}(t)|^2 + 2\Re(\hat{E}(t)\hat{E}_{LO}^*(t)) \quad (3.4)$$

where  $[\ ]^*$  and  $\Re$  denote the complex conjugate and real part operators. Note that the proportionality coefficient is given by the photodiode responsivity  $R$ , as defined in Eq. (2.7).

Let  $P(t) = |E(t)|^2$  and  $P_{LO} = |E_{LO}|^2$  be the power of the optical signal and of the LO, respectively. The expression for the current becomes

$$I(t) \propto P(t) + P_{LO} + 2\sqrt{P(t)P_{LO}} \cos(\Delta\omega t + \theta_{sig}(t) - \theta_{LO}(t)) \quad (3.5)$$

where  $\Delta\omega = |\omega_s - \omega_{LO}|$ ,  $\theta_{sig}(t)$  is the phase of the optical field and  $\theta_{LO}(t)$  is the phase of the LO, basically its phase noise.

It has been proven that by means of this procedure, it is possible to extract the beat of the optical field and the LO. However, the first two terms from Eq. (3.5), called direct-current (DC) components, are not interesting from the point of view of coherent detection. Hence, two ways are possible to eliminate them.

The first one consists of using the  $P_{LO}$  term as an amplifying term, which permits to have the beat term much higher than the DC component.

The second one, instead, consists in a technique called **balanced detection**. The idea is that of having two different photodiodes: on the first one, the impinging field is  $\hat{E}(t) + \hat{E}_{LO}(t)$ , whereas on the second one  $\hat{E}(t) - \hat{E}_{LO}(t)$ . The resulting currents are:

$$\begin{aligned} I_1(t) &\propto P(t) + P_{LO} + 2\sqrt{P(t)P_{LO}} \cos(\Delta\omega t + \theta_{sig}(t) - \theta_{LO}(t)) \\ I_2(t) &\propto P(t) + P_{LO} - 2\sqrt{P(t)P_{LO}} \cos(\Delta\omega t + \theta_{sig}(t) - \theta_{LO}(t)) \end{aligned} \quad (3.6)$$

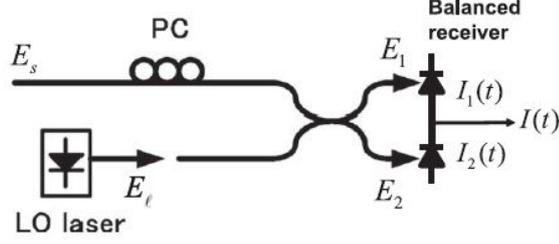
Therefore, by subtracting them, the final output current is

$$I(t) \propto 4\sqrt{P(t)P_{LO}} \cos(\Delta\omega t + \theta_{sig}(t) - \theta_{LO}(t)) \quad (3.7)$$

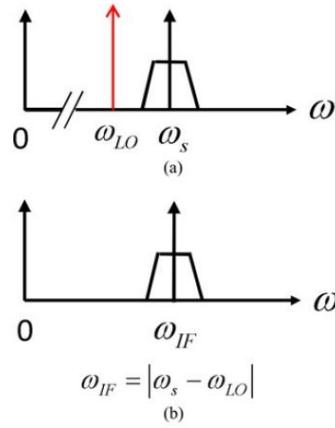
The final result is that of having only the beat between the optical field and the LO. In particular, Eq. (3.7) can be used to implement the two types of coherent detection introduced at the beginning of this section, i.e., **heterodyne** and **homodyne** detection. Figure 3.1 shows the basic configuration for coherent detection.

### 3.1.2 Heterodyne detection

The first technique described is **heterodyne detection**. This type of coherent detection refers to the case in which  $|\Delta\omega| \gg \frac{2\pi}{T}$ , where  $T$  is the symbol time of the signal. This means that the field is down-converted to an **intermediate frequency**  $\omega_{IF}$  and the complex envelope information is preserved, without spectral folding. Figure 3.2 shows this process. Let us recall that the phase of the optical field has been defined as  $\theta_{sig}(t)$ . It is necessary to remember that this phase is composed of two terms, namely the phase modulation  $\theta_s(t)$  and the phase noise of the optical source  $\theta_{sn}(t)$ . Hence, the total signal phase is given by  $\theta_{sig}(t) = \theta_s(t) + \theta_{sn}(t)$ . By



**Figure 3.1:** Configuration of the coherent receiver that measures the beat between the signal and the LO. Balanced detection eliminates the dc component and maximizes the beat signal, as in [7].



**Figure 3.2:** Spectra of (a) the optical signal and (b) the down-converted IF signal. By setting the LO frequency close to the signal frequency, we can obtain the IF spectrum shown in (b), as in [7].

combining this phase noise with the one of the LO, it is possible to separate the contributions of the phase modulation (in which we are interested) and the total phase noise, i.e.,  $\theta_n(t) = \theta_{sn}(t) - \theta_{LO}(t)$ . In this way, the output current from the balanced detector, according to Eq. (3.7), can be re-written as

$$I(t) \propto 4\sqrt{P(t)P_{LO}} \cos(\omega_{IF}t + \theta_s(t) + \theta_n(t)) \quad (3.8)$$

What is more, the complex envelope of the current  $I(t)$  is

$$I_{c.e.}(t) \propto 4\sqrt{P(t)P_{LO}} e^{j(\theta_s(t) + \theta_n(t))} \quad (3.9)$$

which is basically the complex envelope of the optical field, with the exception of the phase noise term. For proper heterodyne detection, then, a phase-locked loop (PLL) is required at the intermediate frequency stage, so as to compensate for the phase noise.

### 3.1.3 Homodyne detection

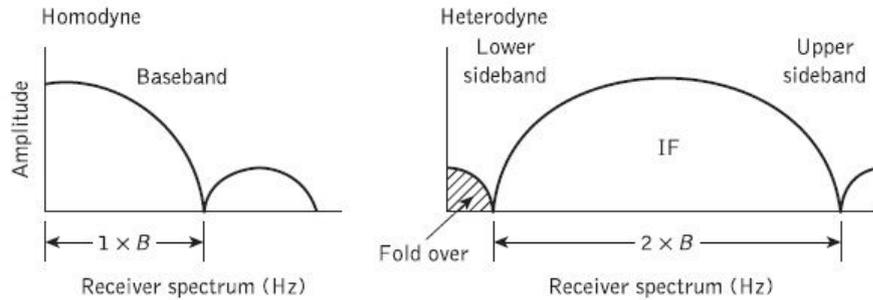
If heterodyne detection brings down the signal to an intermediate frequency  $\omega_{IF}$ , **homodyne detection** refers to the case in which  $\Delta\omega = 0$ , or equivalently in which such intermediate frequency is  $\omega_{IF} = 0$ .

By doing so, the current which is generated from the balanced detector results to be

$$I(t) \propto 4\sqrt{P(t)P_{LO}} \cos(\theta_{sig}(t) - \theta_{LO}(t)) \quad (3.10)$$

Eq. (3.10) basically performs the scalar product between the optical field and the LO. The outcome is that only one quadrature component is measured and the full information on the complex envelope is not available at the receiver at the same time. In addition, for the whole system to work, it is required that the LO phase noise and the optical field phase noise are equal. This is achieved through an optical phase-locked loop (OPLL), which increases by a lot the complexity of the system.

It needs to be specified, though, that the advantage of homodyne detection with respect to heterodyne detection lies in the fact that the signal generated by the former is a base-band signal. Thus the receiver has to deal with a much smaller bandwidth than that in the former case (which is at least twice the size). An example with a PSK signal is presented in Figure 3.3.



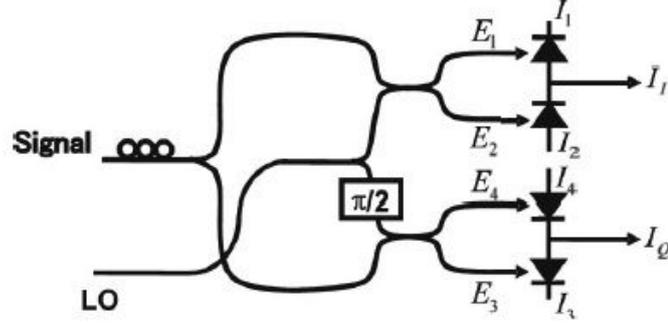
**Figure 3.3:** Comparison of the electrical spectra at the optical detector output for homodyne and heterodyne detection of a PSK signal, as in [1]

### 3.1.4 Phase and polarization diversity

Given that one of the main limitations associated to the homodyne detection is that only one quadrature component can be detected at a time, a solution is supplied by the concept of **phase-diversity**.

Up to now, only one LO has been considered, with the addition of the 3-dB coupler,

to introduce a rotation of  $\pi$  between the two copies of the LO itself. However, in the case of homodyne detection, this allows to detect only one of the two quadrature components. Phase-diversity consists in adding one more LO, which is rotated by  $\pi/2$  with respect to the other one. Such rotation can be introduced by making use of another optical device called **90° optical hybrid**. Furthermore, two balanced detectors are used, i.e., one for each LO. The overall configuration is reported in Figure 3.4.



**Figure 3.4:** Configuration of the phase-diversity homodyne receiver using a 90° optical hybrid. The in-phase component and the quadrature component of the signal complex amplitude are obtained simultaneously, as in [7].

In this case, the fields impinging on the photodiodes are:

$$\begin{aligned}\hat{E}_1(t) &= \hat{E}(t) + \hat{E}_{LO}(t) \\ \hat{E}_2(t) &= \hat{E}(t) - \hat{E}_{LO}(t)\end{aligned}\quad (3.11)$$

for the first balanced detector, and:

$$\begin{aligned}\hat{E}_3(t) &= \hat{E}(t) + j\hat{E}_{LO}(t) \\ \hat{E}_4(t) &= \hat{E}(t) - j\hat{E}_{LO}(t)\end{aligned}\quad (3.12)$$

for the second one.

After that, naming the two currents at the output of the balanced detectors  $I_I(t)$  and  $I_Q(t)$ , it is possible to write

$$\begin{aligned}I_I(t) &\propto 2\sqrt{P(t)P_{LO}} \cos(\theta_{sig}(t) - \theta_{LO}(t)) \\ I_Q(t) &\propto 2\sqrt{P(t)P_{LO}} \sin(\theta_{sig}(t) - \theta_{LO}(t))\end{aligned}\quad (3.13)$$

Finally, the complex envelope can be restored:

$$I_{c.e.}(t) = I_I(t) + jI_Q(t)\quad (3.14)$$

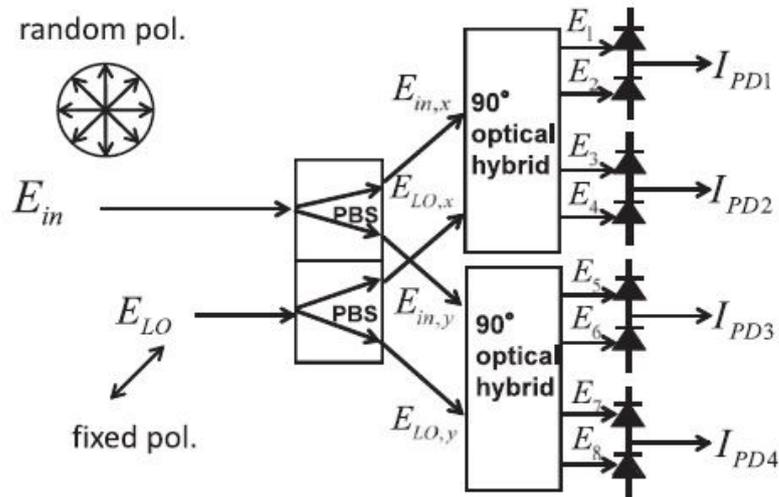
Eq. (3.14) demonstrates how phase-diversity still allows to down-convert the signal to base-band, but avoids the folding of the spectrum, which brings to the detection of only one quadrature component.

This was the definition of **phase-diverse homodyne receiver**. However also the **intradyn receiver** is possible and similar to the homodyne one. The term "intradyn" refers to the fact that, as long as the intermediate angular frequency  $\Delta\omega$  is below  $\frac{2\pi}{T}$ , the whole configuration still works out. Plus, both heterodyne and intradyne detection manage to extract the full information on the signal complex envelope. However, intradyne detection makes the signal base-band (or brings it around a small intermediate frequency). Hence it is preferable with respect to heterodyne detection, since the receiver has to deal with a much smaller bandwidth.

All of these aspects are true, but one element is still missing. The optical field has so far been considered on only one polarization. Moving towards the dual polarization nature of the signal and also considering the fact that the optical field and the LO are not necessarily aligned, due to birefringence, the last concept to introduce is **polarization-diversity**.

The idea is to use two phase-diverse homodyne receivers, after dividing the incoming signal into two orthogonal polarizations with a **polarization beam splitter (PBS)**. Then, from each of them the quadrature components are extracted.

Figure 3.5 shows the configuration of the polarization-diverse and phase-diverse homodyne receiver.

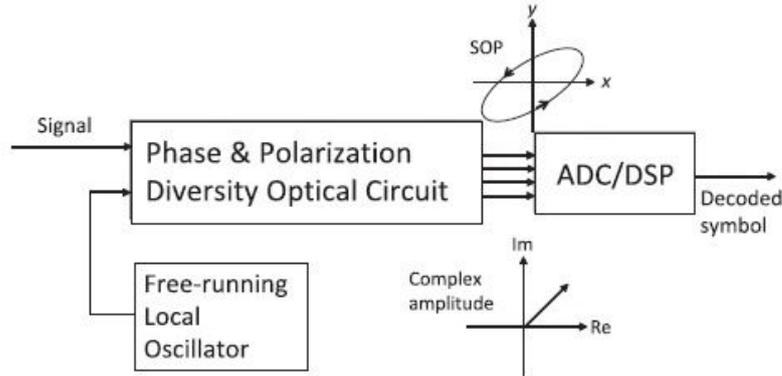


**Figure 3.5:** Configuration of the homodyne receiver employing phase and polarization diversities. Two phase-diversity homodyne receivers are placed in the polarization-diversity configuration, where a common LO is employed, as in [7].

## 3.2 DSP in coherent receivers

The four extracted components are then the input to the DSP chain which follows the coherent detection stage. Since all of the signal processing techniques are implemented in the digital domain, the first element of the chain is the ADC, to convert the four components from analog to digital.

The configuration of the system is presented in Figure 3.6.



**Figure 3.6:** Basic configuration of the digital coherent receiver. Four outputs from the homodyne receiver consisting of phase and polarization diversities are sent to the ADC/DSP circuit. These outputs include all the information of the complex amplitude of the signal, and linear impairments are compensated for in the DSP core, as in [7].

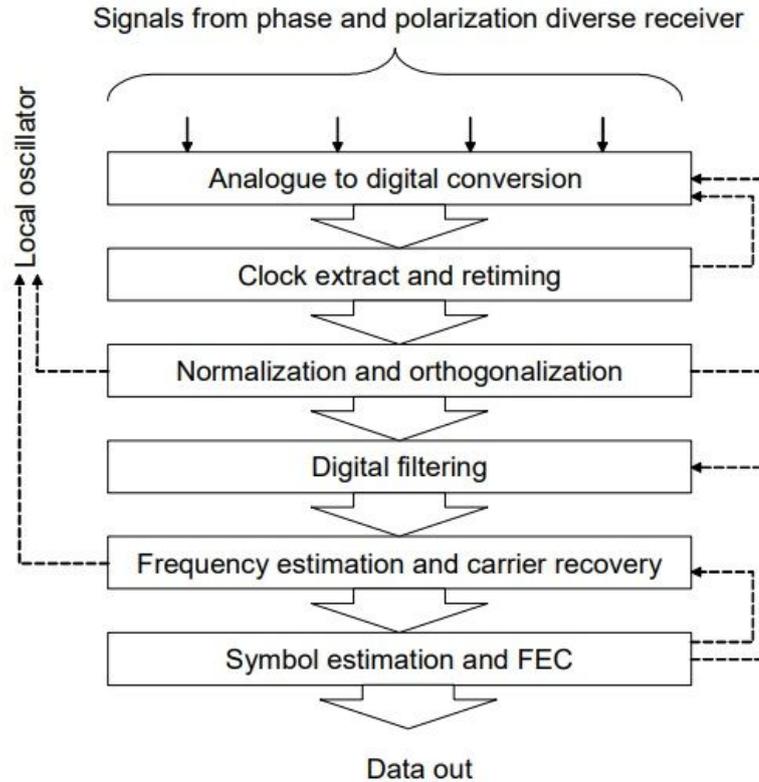
The task of the DSP chain is that of restoring the complex envelope of the input signal in order to be able to decode the originally transmitted data.

Of course, all of these operations need to be performed taking into account the phase noise introduced by the transmitter laser and LO, as well as the polarization multiplexing induced by the fiber birefringence and all the other impairments introduced so far. Hence, it is fundamental to have proper control of such time-varying parameters in the digital domain.

Besides, DSP algorithms improve system stability with respect to other optical control methods (e.g., OPLL), allowing to take them out and eliminate part of the complexity which made coherent detection systems hard to implement.

Figure 3.7 reports the block scheme of the DSP chain in the coherent receiver. The first three blocks are necessary to have the four electrical signals, i.e., the I/Q components on both polarizations, characterized by an integer number of samples per symbol and synchronized with one another.

What is more, the normalization and orthogonalization stage is in charge of compensating for the possible variation in the responsivity values of the two balanced



**Figure 3.7:** Block scheme of DSP chain in coherent receivers, as in [8]

detectors, as well as the imperfections which might be present in the  $90^\circ$  optical hybrids.

Then, digital filtering is applied. Its task is that of compensating for all the impairments affecting the signal during its propagation. As usual, CD, PDL, PMD and all those filtering effects originating from devices along the line and inside the receiver itself can be mentioned.

Finally, going back to Eq. (3.13), the eventual difference between the transmitting laser frequency and LO frequency, together with the phase noise of the two lasers need to be compensated for. For this reason, a block implementing frequency offset and carrier phase recovery is required.

The last block in the DSP chain, reported as "symbol estimation and FEC" refers to the stage at which forward error correction (FEC) decoding is applied, if present, and a decision is made over the received symbols. FEC refers to signal processing techniques which add some redundancy in the transmitted data in order to be able

to detect and correct potential transmission errors at the receiver side. Several techniques exist and with the advent of soft-decision FEC (SD-FEC) (e.g., low parity density codes (LDPC), turbo codes and polar codes) most modern optical fiber communication transceivers make use of binary SD-FEC's based on a bit-interleaved coded modulation (BICM) structure.

Even though this falls out of the scope of this work, it is still mentioned since a couple of experiments will be presented in the following chapters. Indeed, obtaining post-FEC metrics, such as bit error rate (BER), becomes prohibitive. The reason lies in the fact that post-FEC BER's are generally very low (e.g.,  $10^{-15}$ ) and a lot of samples would be required.

Hence, pre-FEC metrics are required. A good metric is the **generalized mutual information (GMI)**, expressed in bit/symbol, representing the number of information bits per symbol that can be reliably transmitted over the channel. Furthermore, it proves to be a good metric especially for codes based on BICM [9]. This aspect justifies the use of GMI as main metric to evaluate the system performance during the experiments which have been carried out.

For the purpose of this work, **digital filtering** is the most interesting aspect. Therefore, the details related to it are presented in Sec. 3.3 and Sec. 3.4.

### 3.3 Linear impairments mitigation

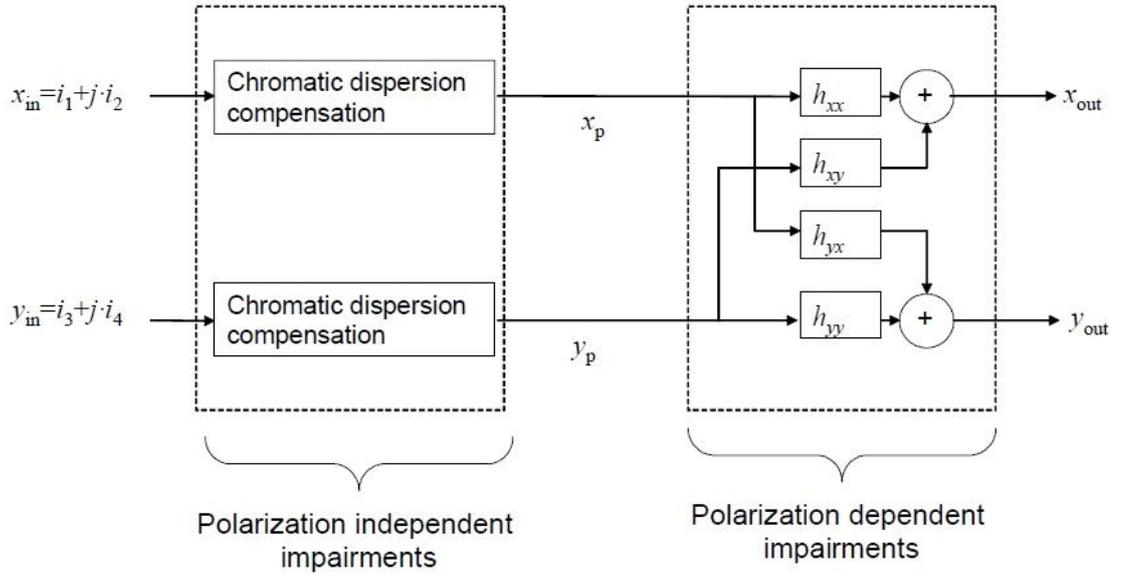
In this section, the equalization of linear transmission impairments is discussed. It is convenient to separate the contribution of polarization-independent (e.g., CD) and polarization-dependent impairments (e.g., PMD, PDL, birefringence).

In this way, it is possible to have two separate blocks whose adaptation rate can be even very different. Indeed, for a fixed optical path the value of cumulated CD is basically fixed, whereas polarization varies on a much faster time scale.

Hence, the block in charge of CD compensation can implement a fixed equalizer in order to compensate for a large amount of CD. This allows to have an equalizer whose tap coefficients do not need frequent updating and relieves the following block of large CD compensation, since only the residual CD needs to be accounted for. What is more, this also allows to have the equalizer characterized by a significantly lower number of taps with respect to the fixed CD compensation stage.

Figure 3.8 shows the general block scheme of the digital filtering stage in a DSP chain.

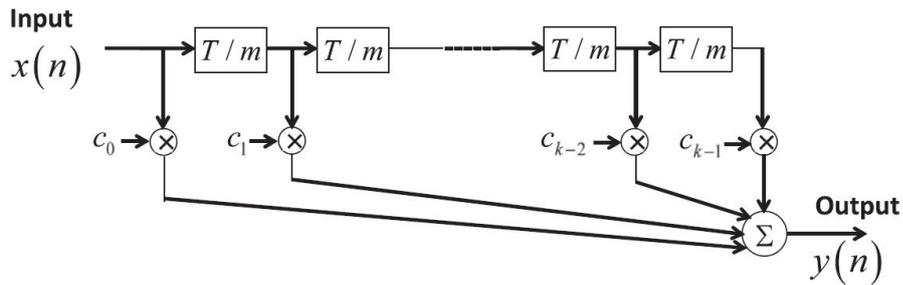
On the other hand, the polarization-dependent block is implemented as  $2 \times 2$  butterfly-structured **finite impulse response (FIR) filters**. This type of structure takes into account also the interaction of the two polarizations and tries to



**Figure 3.8:** Block scheme of the digital filtering stage in a DSP chain. The two blocks can also be exchanged, thanks to system linearity. As in [8]

compensate for it, i.e., polarization de-multiplexing. Note that if the first CD-compensating stage is accurate, the FIR filters can be characterized by a few number of taps, generally in the order of a few tens.

Before going through the techniques used to equalize linear transmission impairments, it is useful to have a general understanding on FIR filters functioning. Figure 3.9 illustrates the basic configuration of a FIR. The filter is composed of  $k$



**Figure 3.9:** Basic configuration of a FIR filter, as in [7]

memory elements, which delay the input signal  $x(n)$  by multiples of  $T/m$ , being  $m$  the oversampling factor.

Let  $\mathbf{x}(n)$  be the vector of these delayed inputs and  $c_i$  the tap coefficients by which

each input gets multiplied ( $i = 0, 1, \dots, k - 1$ ), i.e.,

$$\mathbf{x}(n) = [x(n) \ x(n - 1), \dots, x(n - k + 1)]^T \quad (3.15)$$

$$\mathbf{c} = [c_0 \ c_1, \dots, c_{k-1}]^T \quad (3.16)$$

According to this notation, the output  $y(n)$  can be written as

$$y(n) = \mathbf{c}^T \mathbf{x}(n) \quad (3.17)$$

The tap coefficients  $\mathbf{c}$  thus define the impulse response of the filter itself and its discrete Fourier transform (DFT) represents the transfer function.

Hence, the proper design and adaptive control of such coefficients allow to implement the filters needed to compensate for the signal transmission impairments.

### 3.3.1 Chromatic dispersion compensation

As for CD compensation, FIR filters can be efficiently implemented both in the time domain [8] and frequency domain. The latter are easily implemented thanks to efficient algorithms to compute the DFT of a signal, e.g., the fast Fourier transform (FFT) algorithm. Plus, Eq. (2.19) gives the expression of the CD transfer function  $H_{CD}(f)$ . Therefore, the implementation of a FIR filter to compensate CD in the frequency domain is simply achieved by making sure that the tap coefficients are characterized by values which represent the inverse of the transfer function, i.e.,  $H_{CD}(f)^{-1}$ . Its expression is:

$$H_{CD,eq}(f) = H_{CD}(f)^{-1} = e^{j2\pi^2\beta_2 f^2 z} \quad (3.18)$$

Note that the term  $f_0$  is not present in this expression since the complex envelope of the signal (or WDM channel of interest) is considered.

Finally, the minimum number of taps  $k_{min}$  in the FIR filter can be computed. A good estimation of the channel memory in symbols results to be

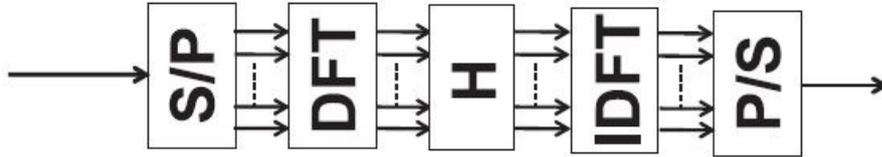
$$\mu = \lambda^2 c^{-1} \frac{R_b}{M^2} DL \quad (3.19)$$

where  $R_b$  is the bit rate,  $M$  is the number of bits per symbol carried and  $L$  is the total distance over which the signal propagated. Considering the number of samples per symbol  $n_{sps}$  at the receiver side, the minimum number of tap coefficients required is given by

$$k_{min} = n_{sps} \mu \quad (3.20)$$

Note that the complexity of a FIR filter scales with the number of taps used. For this reason, considering practical implementations (e.g., application specific integrated

circuit, ASIC), if too many taps are required, a block-by-block processing is possible. A simplified block scheme of the whole compensation process is presented in Figure 3.10, making use of parallel-to-serial (P/S) and serial-to-parallel (S/P) converters and DFT/IDFT operations for block-by-block processing.



**Figure 3.10:** CD compensation block scheme. IDFT is the inverse DFT converting the signal from the frequency to the time domain.  $H$  is the transfer function of the filter. As in [7].

### 3.3.2 Adaptive equalization

Moving to the polarization-dependent impairments, they can be modeled according to a time-varying and frequency-dependent Jones matrix. Thus, the goal is to have the equalizer implementing the inverse of the Jones matrix. Since the polarization variation time scale is very fast, the tap coefficients of the equalizer need to be adaptive. The main focus of this section is on polarization-dependent effects, however it is important to recall the fact that the adaptive equalizer also compensates for eventual filtering effects along the line and residual CD, just to cite a couple of them.

In order to update the tap coefficients and make them adaptive to the channel conditions, some algorithms are needed. A couple of examples are the constant modulus algorithm (**CMA**) and least mean-squared (**LMS**) algorithm.

Let the notation be the same as in Figure 3.8. In this way, the two output signals can be written as

$$x_{out}(n) = \mathbf{h}_{xx}^T \cdot \mathbf{x}_p(n) + \mathbf{h}_{xy}^T \cdot \mathbf{y}_p(n) \quad (3.21)$$

$$y_{out}(n) = \mathbf{h}_{yx}^T \cdot \mathbf{x}_p(n) + \mathbf{h}_{yy}^T \cdot \mathbf{y}_p(n) \quad (3.22)$$

where  $\mathbf{h}_{xx}$ ,  $\mathbf{h}_{xy}$ ,  $\mathbf{h}_{yx}$  and  $\mathbf{h}_{yy}$  are the  $k \times 1$  tap coefficient vectors of the four FIR filters composing the adaptive equalizer.

**CMA** is a possible choice to adapt these coefficients. It works well for modulation formats whose symbols have magnitude equal to one (e.g., QPSK) and attempts to optimize the tap coefficients so as to minimize the errors in MSE sense, defined as

$$\epsilon_x = 1 - |x_{out}(n)|^2 \quad \epsilon_y = 1 - |y_{out}(n)|^2 \quad (3.23)$$

by making use of a stochastic gradient algorithm with learning rate  $\eta$ . The updating equations then become

$$\mathbf{h}_{xx} \rightarrow \mathbf{h}_{xx} + \eta \epsilon_x x_{out}(n) \cdot \mathbf{x}_p^*(n) \quad (3.24)$$

$$\mathbf{h}_{xy} \rightarrow \mathbf{h}_{xy} + \eta \epsilon_x x_{out}(n) \cdot \mathbf{y}_p^*(n) \quad (3.25)$$

$$\mathbf{h}_{yx} \rightarrow \mathbf{h}_{yx} + \eta \epsilon_y y_{out}(n) \cdot \mathbf{x}_p^*(n) \quad (3.26)$$

$$\mathbf{h}_{yy} \rightarrow \mathbf{h}_{yy} + \eta \epsilon_y y_{out}(n) \cdot \mathbf{y}_p^*(n) \quad (3.27)$$

The other solution is **LMS**. The process is the same as CMA, but the equations to update the tap coefficients are

$$\mathbf{h}_{xx} \rightarrow \mathbf{h}_{xx} + \eta \epsilon_x \cdot \mathbf{x}_p^*(n) \quad (3.28)$$

$$\mathbf{h}_{xy} \rightarrow \mathbf{h}_{xy} + \eta \epsilon_x \cdot \mathbf{y}_p^*(n) \quad (3.29)$$

$$\mathbf{h}_{yx} \rightarrow \mathbf{h}_{yx} + \eta \epsilon_y \cdot \mathbf{x}_p^*(n) \quad (3.30)$$

$$\mathbf{h}_{yy} \rightarrow \mathbf{h}_{yy} + \eta \epsilon_y \cdot \mathbf{y}_p^*(n) \quad (3.31)$$

where the errors are defined as

$$\epsilon_x = d_x(n) - x_{out}(n) \quad \epsilon_y = d_y(n) - y_{out}(n) \quad (3.32)$$

and  $d_x(n)$  and  $d_y(n)$  can be either pilot symbols, if a training phase is present during the equalization (**data-aided LMS** or **DA-LMS**), or the closest symbols to  $x_{out}(n)$  and  $y_{out}(n)$ . In this last case, the algorithm is said to be **decision-directed (DD)** LMS, since  $d_x(n)$  and  $d_y(n)$  are not known symbols, but the symbols decided according to the previously equalized ones.

### 3.4 Non-linear impairments mitigation

Up to this point, only linear impairments have been considered. However, for applications involving WDM systems and higher power levels per channel (e.g., a few milliwatts), NL effects arise. SPM, XPM and FWM are some examples and have already been discussed in Chapter 2. It is thus necessary to have some mechanism which is able to compensate also for those effects.

The solution is the **digital back-propagation (DBP)** algorithm [10].

The idea is relatively simple. DBP tries to retrieve the originally transmitted signal starting from the received one. This is achieved by back-propagating the received signal into a digitally reconstructed optical fiber, with all of its main parameters (i.e.,  $\alpha$ ,  $\beta_2$  and  $\gamma$ ) with opposite sign.

The result is that of having the signal propagating into a fiber which compensates for the impairments introduced during the original propagation. What is more, both CD and NL effects are simultaneously compensated. Hence, this algorithm can be implemented in place of the CD compensation stage, composed of a fixed frequency-domain equalizer (see Sec. 3.3.1).

In this section, a theoretical presentation of the DBP algorithm is given. Then, a few examples of result are reported, both in a simulation and experimental environment. This is done especially because DBP is one of the key elements for the techniques presented in Chapter 5. Therefore, it is useful to understand how it works.

### 3.4.1 Digital back-propagation

The propagation of a signal inside an optical fiber is described by the Manakov equations, as in Eq. (2.27). Such equations have two different contributions, one is linear and the other is NL. Therefore, it is possible to re-write them as

$$\frac{\partial}{\partial z} \mathbf{E}(t, z) = \hat{\mathcal{D}}\{\mathbf{E}(t, z)\} + \hat{\mathcal{N}}\{\mathbf{E}(t, z)\} \quad (3.33)$$

where  $\hat{\mathcal{D}} = (j\frac{\beta_2}{2} \frac{\partial^2}{\partial t^2} - \alpha)$  and  $\hat{\mathcal{N}} = -j\frac{8}{9}\gamma(|E_x(t, z)|^2 + |E_y(t, z)|^2)$  are the linear and non-linear operators, respectively.

Since the whole equation is not completely linear, superposition of effects is not valid and such contributions cannot be separated. In addition, no analytical solution exists. However, if a sufficiently short section of fiber is considered, then such separation is possible with a small error and a numerical approach can be implemented. The **split-step Fourier method (SSFM)** is the name of the algorithm used.

Let us consider a single fiber span of length  $L_{span}$ . The method consists of dividing such span into smaller sections of length  $h$ . For each step, the linear part is integrated in the frequency domain, which is made fast and efficient thanks to the FFT algorithm. The field at frequency  $f$  and step  $i$  after this operation can be written as

$$\mathbf{E}(f, ih) = \mathcal{F}\{\mathbf{E}(t, (i-1)h)\} \cdot e^{\hat{\mathcal{D}}h} \quad (3.34)$$

where  $\mathcal{F}$  is the Fourier transform operator.

Then, also the NL contribution is integrated, but in time domain. The final expression for the field after the  $i$ -th step becomes

$$\mathbf{E}(t, ih) = \mathcal{F}^{-1}\{\mathbf{E}(f, ih)\} \cdot e^{\xi \hat{\mathcal{N}}h} \quad (3.35)$$

where  $\mathcal{F}^{-1}$  is the inverse Fourier transform operator, while  $\xi$  is a parameter between 0 and 1 which needs to be optimized in an empirical way and depends on dispersion map, launched power and oversampling factor [10].

This is exactly what is done during DBP. In this case, the Manakov equations become

$$\frac{\partial}{\partial z} \mathbf{E}(t, z) = \hat{\mathcal{D}}^{-1} \{ \mathbf{E}(t, z) \} + \hat{\mathcal{N}}^{-1} \{ \mathbf{E}(t, z) \} \quad (3.36)$$

where  $\hat{\mathcal{D}}^{-1} = (-j \frac{\beta_2}{2} \frac{\partial^2}{\partial t^2} + \alpha)$  and  $\hat{\mathcal{N}}^{-1} = j \frac{8}{9} \gamma (|E_x(t, z)|^2 + |E_y(t, z)|^2)$  are the inverse linear and non-linear operators, respectively. DBP simply implements an asymmetric SSMF backwards. Note, the name **asymmetric** derives from the fact that first the linear part and then the non-linear one are integrated at each step. If the integration of the linear part occurs half before and half after the integration of the non-linear one, then the SSMF is said to be **symmetric**.

Finally, the step size is also an important parameter to consider. In the simplest case, an equi-spaced approach can be adopted, i.e.  $h = L_{span}/N_{step}$ , being  $N_{step}$  the number of steps to perform. Nevertheless, something more elaborate can be done taking into account the characteristics of the effects considered. An example is the **logarithmic step size** [11], which considers the fact that the power level in a fiber span has a logarithmic distribution, so smaller steps are used when the power is higher (i.e., higher NL), and larger steps when power is lower. Its expression can be written as

$$h_i = -\frac{1}{2\alpha} \ln \left[ \frac{1 - (N_{step} - i + 1)\delta}{1 - (N_{step} - i)\delta} \right], \quad i = 1, \dots, N_{step} \quad (3.37)$$

where  $\delta = (1 - e^{-2\alpha L_{span}})/N_{step}$  and  $\ln(\cdot)$  is the natural logarithm.

### 3.4.2 Simulation results

To prove the correct functioning of DBP, the algorithm is tested in a simulation environment, using **MATLAB**.

In particular, the simulated system is a WDM system with 5 channels. Each of them carries a PM-64QAM modulation format at a symbol rate of  $R_s = 64$  GBaud. In addition, the channel spacing is  $\Delta f = 1.2R_s = 76.8$  GHz.

The pulse is shaped by means of a Square-Root Raised Cosine (SRRC) filter characterized by a roll-off factor  $\rho = 0.2$ .

Moving to the optical fiber link, it is composed of 10 spans of SSMF. The main parameters are  $\alpha_{dB} = 0.21$  dB/km,  $D = 16.7$  ps/nm/km and  $\gamma = 1.3$  1/W/km. Each span is  $L_{span} = 100$  km long and EDFA's are placed between spans. The gain

of such optical amplifiers perfectly compensates for the losses along the span and they introduce ASE noise with a noise figure  $F = 6$  dB.

The goal of the simulation is to compute the maximum reach of the system when using DBP at the receiver side. Besides, the maximum reach is computed considering a target GMI of 5.22 bit/symbol.

Hence, the propagation of the signal along the optical fiber is simulated and the output is sent to the receiver DSP. Here, the signal is resampled at 2 samples/symbol and DBP is applied. Note that only the central WDM channel gets back-propagated. This means that it is generally not possible to compensate for all the NL impairments. Then, the signal goes through a matched SRRC filter and the training sequence gets realigned. Adaptive equalization is performed according to a DA-LMS algorithm for the training stage and DD-LMS for the tracking stage. No frequency offset compensation is present, since no phase noise is present. However, even if it was, the following phase recovery stage manages to track it as long as it is characterized by a low enough value. In the end, phase recovery is applied, according to a blind phase search (BPS) scheme, and the symbols are decoded. At this point GMI is computed.

Several power levels have been tested, to check how high the launched power can be, before NL effects kick in and start impairing the system way too much. Indeed, the power level for each channel  $P_{ch}$  has been swept in the range -2 dBm to 10 dBm. What is more, the parameter  $\xi$  has been optimized for each DBP application and a different number of steps per span has been considered.  $N_{step}$  has been set equal to 1, 2, 4, 8 and 16, with logarithmic spacing as defined in Eq. (3.37).

Figure 3.11 shows the results obtained. In order to be able to make comparisons and draw conclusions, also the case in which no DBP is applied is reported. In this way, it is possible to appreciate how DBP allows for a higher maximum reach.

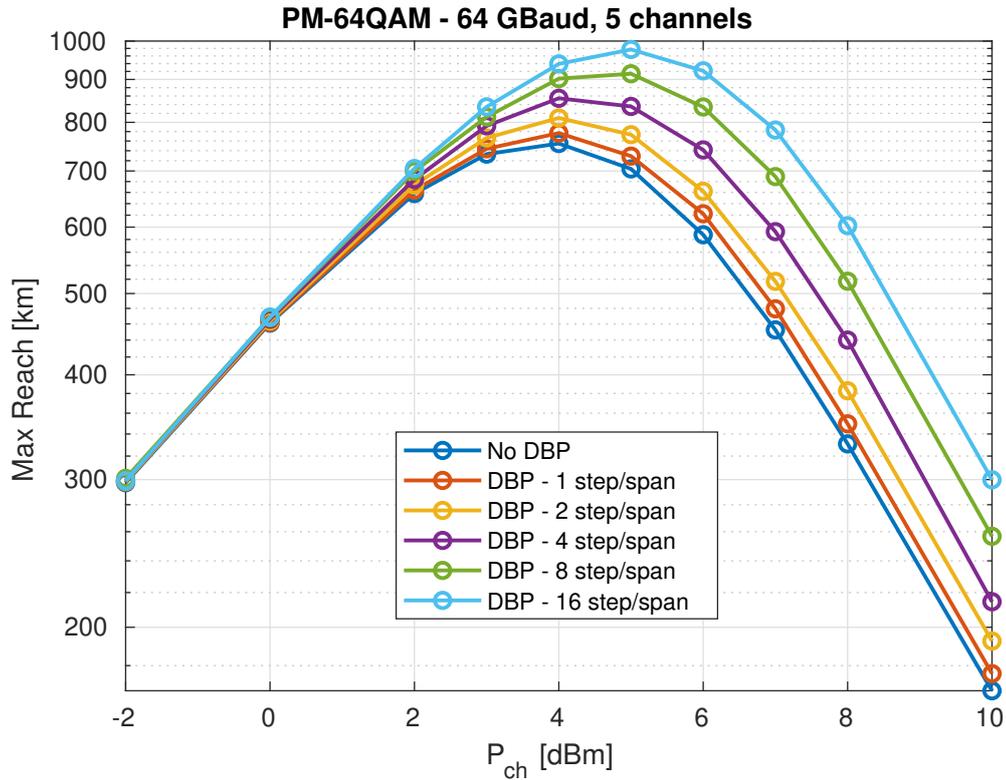
Of course, no DBP means that only fixed CD compensation is performed.

Let us recall that theoretically, if no NL effects were present, the maximum reach of an optical fiber communication system would increase with power.

In particular, for optically amplified multi-span systems, such as the one considered in this simulation, the increase is linearly proportional to the launched power.

This is coherent with what happens for the channel powers from -2 dBm up to around 1 dBm. Furthermore, no difference can be noticed between no-DBP and DBP cases. The reason which lies behind this fact is that up to that point the system can be assumed to be linear.

Hence, only CD compensation is enough, as no strong NL effects occur. Both fixed CD compensation and DBP work.



**Figure 3.11:** Maximum reach for a WDM system with 5 channels and 64 GBaud symbol rate, when using DBP algorithm at receiver side, for different values of  $P_{ch}$  and  $N_{step}$ .

However, as power increases, NL effects become stronger and stronger. The linear increase in maximum reach versus power stops and the curves begin to bend. Indeed, after a specific channel power level, the maximum reach starts decreasing. This is a clear example of how fiber non-linearities represent one of the major impairments for fiber communications when power levels are high. Just to have numerical values, in the case in which only CD is compensated for, the maximum reach for  $P_{ch} = -2$  dBm is around 300 km. For  $P_{ch} = 10$  dBm, instead, it goes down to 168 km, i.e., a loss of 2.5 dB. An optimal power, though, exists in the middle. For no-DBP case, it is 4 dBm and corresponds to a reach of 754 km.

It is evident how DBP manages to compensate for non-linearities and leads to higher launched powers and system reach. Although the bending of the curves is always present, since for increasingly higher power levels DBP does not manage to keep up, the peak moves towards higher  $P_{ch}$  values. This improvement is even more

pronounced when a larger number of steps per span in the algorithm is performed. Of course, the trade off is the complexity of the algorithm itself.

Table 3.1 reports all the results of interest. In particular, the gain refers to the gain in maximum reach with respect to the case in which no DBP is performed.

	Maximum reach [km]	$P_{ch}$ [dBm]	Gain [dB]
<b>No DBP</b>	754	4	0
<b>DBP - 1 step/span</b>	776	4	0.12
<b>DBP - 2 step/span</b>	809	4	0.31
<b>DBP - 4 step/span</b>	854	4	0.54
<b>DBP - 8 step/span</b>	914	5	0.84
<b>DBP - 16 step/span</b>	977	5	1.13

**Table 3.1:** Maximum reach when applying no DBP and DBP with different number of steps per span, along with the maximum channel power associated to it. Also the gain with respect to the case in which no DBP is performed is reported.

Note that when using DBP a gain up to 1.13 dB can be achieved. This translates into a maximum reach up to 220 km higher with respect to the no-DBP case. To sum up, DBP algorithm works well when NL effects need to be compensated. For now, this has been verified for simulations only, though.

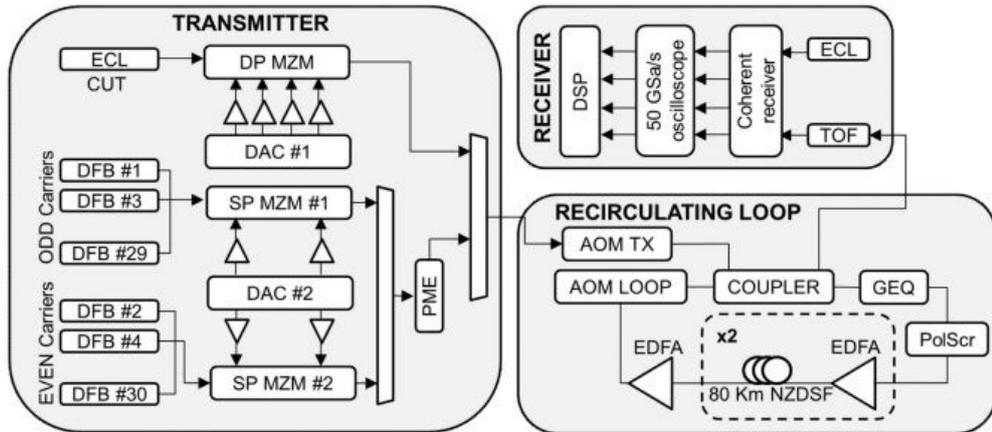
### 3.4.3 Experimental results

In this last section, DBP is applied in the post-processing of data collected in a real long-haul transmission experiment, performed in the Photonlab at Politecnico di Torino [12]. The experimental set-up is reported in Figure 3.12.

At the transmitter side, a WDM signal with 31 channels is generated. Each channel is characterized by a symbol rate  $R_s = 16$  GBaud and a spacing  $\Delta f = 25$  GHz. The central channel, which is also the channel under test, is generated through an ECL laser with a linewidth lower than 100 kHz. The other 30 interfering channels, instead, through DFB lasers. The shape of the pulses is given by a SRRC filter with roll-off  $\rho = 0.15$ .

The WDM signal is then transmitted onto a recirculating loop. Such loop structure is composed of four 80 km spans of G.652 SMF.

The main parameters of the fiber are  $\alpha_{dB} = 0.2$  dB/km,  $\beta_2 = -21.27$  ps<sup>2</sup>/km and  $\gamma = 1.3$  1/W/m. Each span is then terminated by an EDFA with noise figure  $F = 5.2$  dB, whose gain fully recovers span loss.

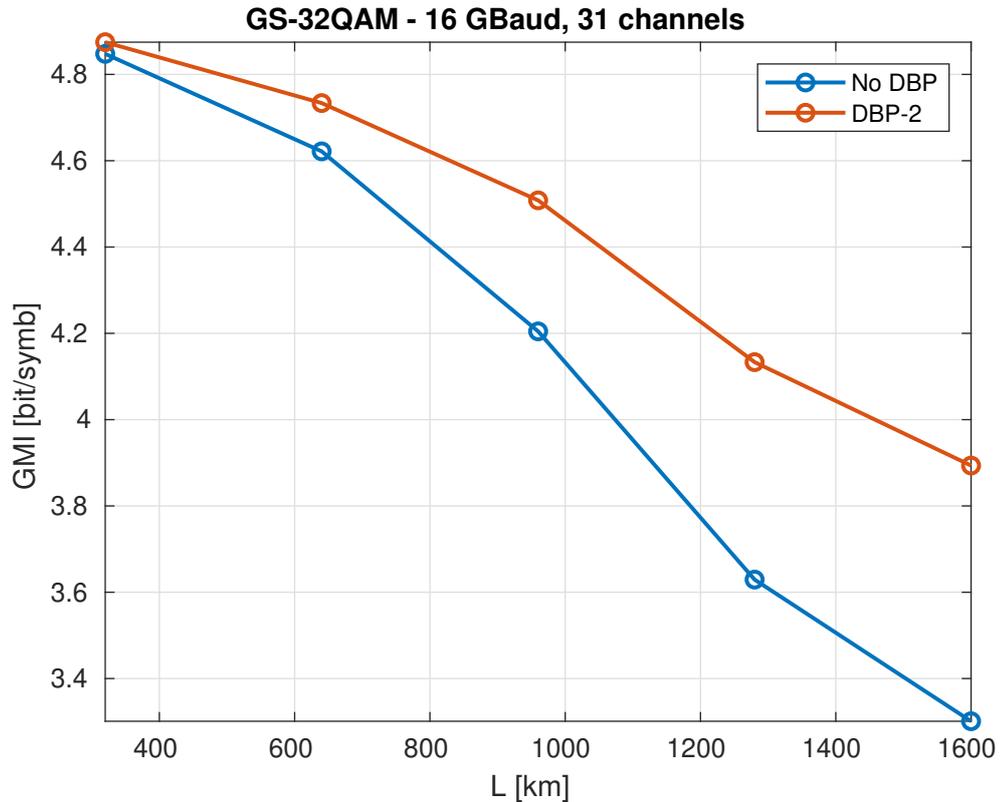


**Figure 3.12:** Set-up of the experiment, as in [12]. Note that in this case the spans are  $4 \times 80$ -km SSMF.

At the receiver side, a tunable optical filter extracts the channel under test, which gets mixed with another ECL laser. Its linewidth is once again lower than 100 kHz. Finally, the four outputs of the coherent receiver are sampled and converted into the digital domain by means of an oscilloscope, working at 50 GSa/s.

The collected data are processed offline. They are down-sampled to 2 samples/symbol and DBP is applied to the channel under test. The frequency offset gets coarsely recovered and the signal gets equalized with a LMS-based adaptive equalizer. Finally, phase recovery is implemented according to a BPS-ML (BPS followed by maximum likelihood) algorithm.

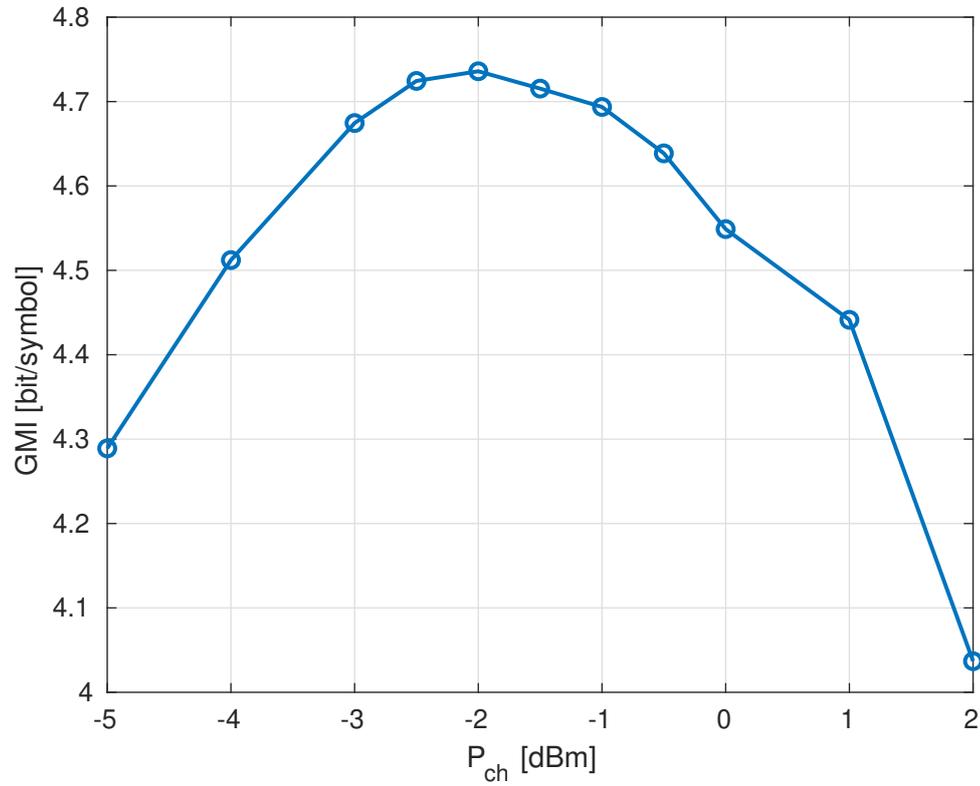
The goal of this experiment is to show how DBP helps improve the resulting GMI of the signal in such scenario. In particular, two cases are considered: no DBP case, in which only fixed CD compensation is applied before equalization and DBP-2 case, in which DBP with 2 step/span is applied. The parameter  $N_{step}$  has been chosen equal to 2 since it is a commonly used value and represents a good trade-off between performance and complexity, given the long link to back-propagate. Indeed, the signal has been propagated over 5 recirculations, corresponding to a total of 20 80-km spans and a resulting transmission distance of 1600 km. As for the channel power  $P_{ch}$ , it has been swept in the range  $P_{ch} \in [-5, -4, -3, -2, -1.5, -1, -0.5, 0, 1, 2]$  dBm. What is more, given that 5 waveforms are collected at the end of each loop, the median GMI is considered so as to cut outliers out. Figure 3.13 reports the results obtained.



**Figure 3.13:** GMI versus distance when no DBP (blue) and DBP with 2 step/span (orange) is applied.

Note that the represented points are chosen considering only the optimal power among the tested ones, namely the power associated to the highest GMI for a defined recirculation. Indeed, the GMI varies depending on the launched power and with the distance. Figure 3.14 gives a representation of this fact, considering the GMI after 5 recirculations.

It is evident that also in a real experimental scenario DBP allows for a huge improvement in terms of performance. After the first recirculations, the GMI gain is low, while it starts increasing after successive loops. It is coherent with what was obtained during simulations. For short distances, NL effects do not impair the system too much and the two cases are almost equivalent. For instance, at the end of the first loop, the gain after using DBP is just 0.02 dB. However, moving towards longer distances, the gain keeps increasing, reaching a value of 0.7 dB after 1600 km. Furthermore, depending on the target GMI, the algorithm allows to gain at least 100 km in transmission distance.



**Figure 3.14:** GMI versus channel power  $P_{ch}$  after 5 recirculations. The optimal power in this case is -2 dBm.

# Chapter 4

## Optical performance monitoring

### 4.1 Introduction

In this chapter, coherent optical fiber communication systems are considered within a dynamically reconfigurable all-optical networks scenario.

Such networks introduce **high flexibility** by means of adaptive routing of optical paths and switching of optical wavelengths. In addition, **bandwidth efficiency** is required to meet the increasingly high bandwidth demand, which can be obtained resorting to multilevel modulation formats. The enabler of both these aspects is the digital coherent receiver, which has already been discussed in Chapter 3.

It performs intradyne coherent detection and allows for the processing of all the data in the electrical domain, making use of DSP techniques.

Not only does it manage to compensate for all the linear impairments affecting the signal while propagating in the fiber (see Chapter 2, Sec. 2.3), but it also enables another important functionality: **optical performance monitoring (OPM)**.

OPM entails evaluating the quality of a data channel by measuring its optical properties without directly inspecting the original bit sequence.

Basically, it represents a "health" monitoring of the optical system and a potential approach to enhance control and automatization of transmission and physical layer fault management in optical transmission systems.

The adaptive equalization within the digital coherent receiver allows to compensate for Chromatic Dispersion (CD), Polarization-Mode Dispersion (PMD), Polarization-Dependent Loss (PDL) and bandwidth limitations, but it also allows to get an estimation of such impairments. What makes this possible is the fact that the filter impulse response of the equalizer relates to the inverse channel impulse response.

Thus, in a weakly non-linear regime, it is possible to separate each individual contribution simply through the processing of the **equalizer tap coefficients**. The advantage of this approach lies in the fact that it does not decrease the signal power and does not require external devices, such as Optical Spectrum Analyzers (OSA) and tapping devices, granting a **cost efficient** and **reliable in-service estimation** of impairments.

This leads to **digital OPM**, i.e., OPM techniques fully implemented in the digital domain. It is the subject of this work and will be discussed in the following chapters. In particular, the main references for this chapter are [13] and [14].

## 4.2 Digital Optical Performance Monitoring

Let us consider a coherent receiver as the one described by the schematic reported in Figure 3.6. The four electrical signals, representing the four quadrature components, are converted into the electrical domain by means of an ADC and sent into the DSP block. Here, they go through clock extraction and retiming, a normalization and orthogonalization stage and finally digital filtering is applied.

Digital filtering is composed of two parts, as described in Figure 3.8. At first, fixed CD compensation is performed and then also adaptive equalization, according to a four-complex-FIR-filter structure in a butterfly configuration.

This is the key point in the DSP chain for what concerns OPM. Indeed, the complex tap coefficients of the adaptive equalizer are constantly updated according to specific algorithms, e.g., CMA and LMS (see Chapter 3 Sec. 3.3.2). After some time, such coefficients converge to an optimal value for the equalization process.

It is possible to represent these complex FIR filters with a  $2 \times 2$  matrix  $\mathbf{W}(n)$  with  $n = 1, 2, \dots, N_{tap}$ , where  $N_{tap}$  is the number of taps in the FIR filters. Each element represents one of the four FIR filters constituting the adaptive equalizer, i.e.,

$$\mathbf{W}(n) = \begin{bmatrix} \mathbf{w}_{xx}(n) & \mathbf{w}_{yx}(n) \\ \mathbf{w}_{xy}(n) & \mathbf{w}_{yy}(n) \end{bmatrix} \quad (4.1)$$

The matrix  $\mathbf{W}(n)$  is the filter impulse response and converges to the inverse of the channel impulse response, so as to compensate for it. It is exactly thanks to this relation between the two impulse responses that it is possible to extract an estimate of the main parameters of interest.

It is, therefore, convenient to first give a description of how the channel is modeled according to this matrix notation.

### 4.2.1 Channel Model

In coherent optical fiber communication systems, the transmitted signal in the frequency domain is in the form

$$\mathbf{s}(f) = [s_x(f) \ s_y(f)]^T \quad (4.2)$$

where the subscripts x and y indicate the x- and y-polarization in the optical fiber. Let us define the  $2 \times 2$  Jones matrix of the optical channel as  $\mathbf{J}(f)$  and the AWGN noise vector as  $\mathbf{n}(f)$ . Neglecting all NL propagation effects, at the receiver side the received signal can be written as

$$\begin{aligned} \mathbf{r}(f) &= \mathbf{J}(f)\mathbf{s}(f) + \mathbf{n}(f) \\ &= \begin{bmatrix} J_{xx}(f) & J_{yx}(f) \\ J_{xy}(f) & J_{yy}(f) \end{bmatrix} \begin{bmatrix} s_x(f) \\ s_y(f) \end{bmatrix} + \begin{bmatrix} n_x(f) \\ n_y(f) \end{bmatrix} \end{aligned} \quad (4.3)$$

It is well-known by now that, during its propagation along the optical communication channel, the signal is affected by CD, PMD, PDL and filtering effects of different nature. Hence, it is necessary to model  $\mathbf{J}(f)$  accordingly.

Let us define the the electro-optical system scalar transfer function as  $H(f)$ , the PMD matrix as  $\mathbf{U}(f)$ , the PDL matrix as  $\mathbf{P}(f)$  and the residual CD scalar factor  $\exp(j2\pi^2\beta_{2,res}f^2)$ . The channel matrix can be written as

$$\mathbf{J}(f) = H(f)e^{-j2\pi^2\beta_{2,res}f^2}\mathbf{P}(f)\mathbf{U}(f) \quad (4.4)$$

Note that in this context  $\beta_{2,res}$  is the amount of cumulated CD which has not been compensated for during the first fixed CD compensation stage. Thus, its units are simply  $\text{ps}^2$ .

In particular, the PMD matrix can be modeled as

$$\mathbf{U}(f) = \begin{bmatrix} e^{j(\phi+2\pi f\Delta\tau_{DGD})/2} & 0 \\ 0 & e^{-j(\phi+2\pi f\Delta\tau_{DGD})/2} \end{bmatrix} \begin{bmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{bmatrix} \quad (4.5)$$

where  $\phi$  is a phase shift,  $\Delta\tau_{DGD}$  is the DGD between the fast and slow axes of the fiber and  $\alpha$  is the angle between the state of polarization of the signal  $s_x$  and the slow axis. Indeed, the second matrix in Eq. (4.5) is the Jones matrix of a polarization rotation by an angle  $\alpha$ .

Owing to its structure,  $\mathbf{U}(f)$  is a **unitary** matrix. This means that

$$\begin{aligned} \mathbf{U}^\dagger(f) &= \mathbf{U}^{-1}(f) \\ |\det \mathbf{U}(f)| &= 1 \end{aligned} \quad (4.6)$$

where  $[\ ]^\dagger$  denotes the conjugate transpose operator and  $\det$  the matrix determinant operator.

On the other hand, the PDL matrix  $\mathbf{P}(f)$  can be modeled as

$$\mathbf{P}(f) = \begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix} \quad (4.7)$$

where  $k$  indicates the attenuation of one polarization with respect to the other introduced by PDL. Hence,  $k$  is always a value between 0 and 1 and the PDL in dB can be computed as  $PDL_{dB} = |20 \log_{10}(k)|$ . In addition, the PDL matrix  $\mathbf{P}(f)$  is **Hermitian**, i.e.,

$$\mathbf{P}^\dagger(f) = \mathbf{P}(f) \quad (4.8)$$

Then, the signal  $\mathbf{r}(f)$  goes through all the previously described DSP stages, up to adaptive equalization. The transfer function of the equalizer is given by the DFT of its impulse response, that is

$$\begin{aligned} \mathbf{W}(f) &= \mathcal{F}\{\mathbf{W}(n)\} \\ &= \begin{bmatrix} W_{xx}(f) & W_{yx}(f) \\ W_{xy}(f) & W_{yy}(f) \end{bmatrix} \end{aligned} \quad (4.9)$$

where each element is the DFT of the single elements from  $\mathbf{W}(n)$ , i.e.,  $\mathcal{F}\{\mathbf{w}_i(n)\}$  with  $i = xx, xy, yx, yy$ . Besides, given that the impulse response of the adaptive equalizer approaches the inverse impulse response of the channel, the same applies for their transfer function. It is, thus, possible to write

$$\mathbf{W}(f) = \mathbf{J}^{-1}(f) \quad (4.10)$$

which can be further expanded into

$$\mathbf{W}(f) = H^{-1}(f) e^{2\pi^2 \beta_{2,res} f^2} \mathbf{U}^{-1}(f) \mathbf{P}^{-1}(f) \quad (4.11)$$

Even more interesting is the fact that it is easily proved that the inverse transfer function of the channel, represented by the matrix  $\mathbf{W}(f)$ , is characterized by the same DGD, PDL and residual CD (but with opposite sign) as the original channel matrix  $\mathbf{J}(f)$ . Plus, the inverse matrices share the same properties as the original ones.

This aspect is of fundamental importance, because it means that, even though the characteristics of the optical channel are not known at the receiver side, they can be extracted by simply acting on the equalizer transfer function.

The only operations to perform are: extract the tap coefficients from the four FIR

filters in the adaptive equalizer, take the DFT to obtain the transfer function and perform some matrix operations in order to isolate the quantities related to the impairments one is interested in monitoring.

Particularly, this last operation is made easier if one exploits the characteristics of the matrices involved, such as their being unitary or Hermitian.

In the next sections, a few techniques to extract PDL, residual CD  $\beta_{2,res}$ , electro-optical system transfer function  $H(f)$  and DGD are presented.

### 4.2.2 PDL estimation

The first estimated parameter is the PDL. This computation is made easier considering the fact that the matrix  $\mathbf{P}(f)$  is Hermitian and  $\mathbf{U}(f)$  is unitary. Indeed, this allows to isolate the contribution of the PDL from the other ones. For this purpose, the matrix  $\mathbf{T}(f)$  is introduced and defined as

$$\mathbf{T}(f) = \mathbf{W}^\dagger(f) \cdot \mathbf{W}(f) \quad (4.12)$$

Therefore, substituting in Eq. (4.12) the expression of matrix  $\mathbf{W}(f)$ , as reported in Eq. (4.11), we obtain

$$\mathbf{T}(f) = (H^{-1}(f))^* e^{-2\pi^2 \beta_{2,res} f^2} (\mathbf{P}^{-1}(f))^\dagger (\mathbf{U}^{-1}(f))^\dagger \mathbf{U}^{-1}(f) \mathbf{P}^{-1}(f) e^{2\pi^2 \beta_{2,res} f^2} H^{-1}(f) \quad (4.13)$$

which can be further simplified applying the properties described in Eq. (4.6) and Eq. (4.8) as

$$\mathbf{T}(f) = |H^{-1}(f)|^2 \mathbf{P}^{-2}(f) \quad (4.14)$$

which completely eliminates the PMD and residual CD contributions.

To conclude, given the eigenvalues of the matrix  $\mathbf{T}(f)$ , namely  $\lambda_1(f)$  and  $\lambda_2(f)$ , the PDL can be estimated for each frequency  $f$  as

$$PDL(f) = \left| \log_{10} \left( \frac{\lambda_1(f)}{\lambda_2(f)} \right) \right| \quad (4.15)$$

It is important to notice that in Eq. (4.14) there is still a dependence on the system scalar transfer function. However, such dependence simply appears as a multiplicative factor of the computed eigenvalues and gets simplified when the ratio between the two is taken in Eq. (4.15).

### 4.2.3 Residual CD estimation

The second step consists in isolating the residual CD contribution. In order to do this, the inverse matrix of  $\mathbf{P}^{-1}(f)$ , i.e.,  $\mathbf{P}(f)$ , needs to be computed. This operation can be carried out resorting to the similarity relation to diagonalize a matrix. Recall that, given the previously computed eigenvalues  $\lambda_1(f)$  and  $\lambda_2(f)$  of  $\mathbf{P}^{-1}(f)$ , its inverse's eigenvalues are  $\lambda_1^{-1}(f)$  and  $\lambda_2^{-1}(f)$ . Hence, defining the diagonal matrix  $\mathbf{D}(f)$  as

$$\mathbf{D}(f) = \begin{bmatrix} \lambda_1^{-1}(f) & 0 \\ 0 & \lambda_2^{-1}(f) \end{bmatrix} \quad (4.16)$$

and the matrix  $\mathbf{V}(f)$  as the matrix whose columns are the eigenvectors of  $\mathbf{P}^{-1}(f)$ , then  $\mathbf{P}(f)$  can be computed as

$$\mathbf{P}(f) = \mathbf{V}(f)\mathbf{D}(f)\mathbf{V}^{-1}(f) \quad (4.17)$$

At this point, the residual CD contribution can be isolated just by considering

$$\det(\mathbf{W}(f) \cdot \mathbf{P}(f)) = \left[ H^{-1}(f) e^{2\pi^2 \beta_{2,res} f^2} \right]^2 \det(\mathbf{U}^{-1}(f)) \quad (4.18)$$

and the second expression in Eq. (4.6). Indeed, by computing the phase of the determinant of such matrix product, one obtains

$$\arg(\det(\mathbf{W}(f) \cdot \mathbf{P}(f))) = 2 \arg(H^{-1}(f)) + 4\pi^2 \beta_{2,res} f^2 \quad (4.19)$$

where  $\arg$  is the argument operator to compute the phase of a complex number. The result is then the sum of the phases of the system scalar transfer function and the residual CD contribution. The last step is to eliminate the first term.

So as to do that, it is possible to take into account the fact that the phase of the system scalar transfer function is an odd function around the origin. Thus, to cut its contribution out, just a flipping and sum operation is needed, i.e.,

$$\frac{\arg(\det(\mathbf{W}(f) \cdot \mathbf{P}(f))) + \mathit{flip}\{\arg(\det(\mathbf{W}(f) \cdot \mathbf{P}(f)))\}}{4} = 2\pi^2 \beta_{2,res} f^2 \quad (4.20)$$

where  $\mathit{flip}\{ \}$  is the operator inverting the order of the elements in a vector.

Note that Eq. (4.20) represents a parabolic function versus the frequency. For this reason, the quantity  $\beta_{2,res}$  can be obtained performing a parabolic interpolation of the unwrapped phase and a proper normalization by  $2\pi^2$ .

#### 4.2.4 System scalar transfer function estimation

Starting once again from Eq. (4.18), the system scalar transfer function can also be estimated or, to put it better, its squared magnitude.

It is enough to consider that the residual CD contribution is just a phase factor, meaning that it is a complex number with magnitude equal to one and introducing only a phase rotation. For this reason, if the magnitude of such quantity is computed, the factor is taken out. This operation yields

$$|\det(\mathbf{W}(f) \cdot \mathbf{P}(f))| = |H^{-1}(f)|^2 \quad (4.21)$$

For what concerns this transfer function, it arises from any band-pass filtering that occurs during the transmission on the communication link or low-pass filtering during the demodulation process. However, it is necessary to precise that if one really wants to estimate the real inverse system transfer function, the contribution related to the shape of the transmitted pulses should be removed from it.

Finally, if the complex expression of  $H^{-1}(f)$  is needed, its phase can be retrieved from Eq. (4.19), by subtracting the estimated residual CD phase. Thus, the whole transfer function can be written as  $H^{-1}(f) = |H^{-1}(f)|e^{j\arg(H^{-1}(f))}$ .

#### 4.2.5 DGD estimation

The last parameter to be estimated is then the DGD, related to PMD effects.

Given that all the previous quantities have been estimated, it is possible to fully invert Eq. (4.11) and obtain the inverse PMD matrix  $\mathbf{U}^{-1}(f)$ . At this point, the DGD  $\Delta\tau_{DGD}(f)$  is estimated as

$$\Delta\tau_{DGD}(f) = \left| 2\sqrt{\det\left(\frac{1}{2\pi} \frac{d}{df} \mathbf{U}^{-1}(f)\right)} \right| \quad (4.22)$$

### 4.3 Simulation results

After having introduced all the main techniques to extract and estimate the parameters related to the major system impairments, some examples of result are presented and discussed. In particular, as already done at the end of Chapter 3, the algorithms are first tested within a simulation environment (**MATLAB**) and then on experimental data.

Let us begin reporting a few simulation results. Of course simulations make everything much easier, since there is complete control over what happens and allow to insert all of the discussed impairments with the desired values.

Therefore, the goal of this section is that of introducing the previously discussed

impairments with well-aimed values, so as to be able to check if the algorithms manage to properly estimate them.

The system considered in this section is again a WDM system, but with 3 channels. Each channel carries a PM-16QAM constellation at a symbol rate  $R_s = 64$  GBaud and is spaced from the others by  $\Delta f = 76$  GHz. The pulses are shaped according to a SRRC filter characterized by a roll-off factor  $\rho = 0.2$  and the channel power is  $P_{ch} = 0$  dBm.

The optical link is composed of 4 spans of SMF with a span length  $L_{span} = 100$  km. The main parameters of the optical fiber are  $\alpha_{dB} = 0.2$  dB/km and  $\beta_2 = -21.285$  ps<sup>2</sup>/nm. Note that the parameter  $\gamma$  has been set to 0, since one of the requirements for the algorithm to properly work is to operate in weakly non-linear conditions. Hence, to make it easier, NL effects have been ignored for now.

At the end of each span is an EDFA, characterized by a noise figure  $F = 5$  dB and a gain  $G$  which perfectly compensates for the total span loss.

The signal is then resampled at 2 samples/symbol and goes into the DSP chain. Here, fixed CD compensation and matched filtering are performed. For the same reason as before, no DBP is performed on the signal.

Afterwards, the signal gets realigned and adaptive equalization is applied with  $N_{tap} = 26$ . In particular, the tap coefficients update strategy consists of a DA-LMS based training stage and a DD-LMS based tracking stage (see Sec. 3.3.2 for details). Finally, phase recovery is carried out with a BPS-ML strategy and the symbols are decoded.

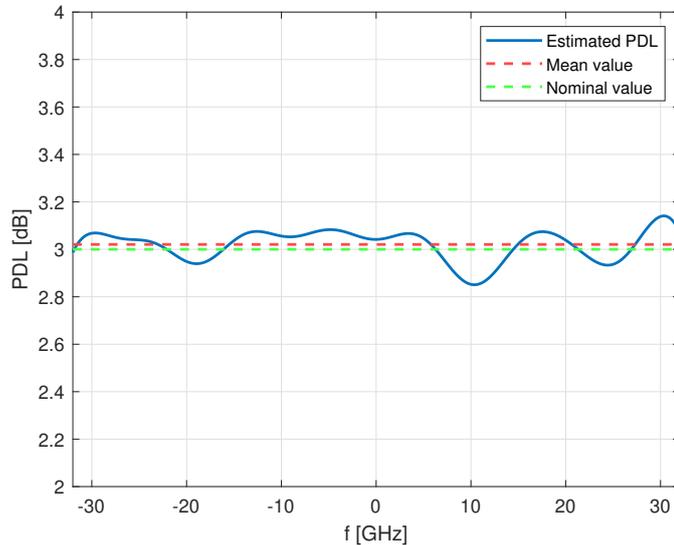
It is important to mention that this is the overall basic system structure. However, since the OPM algorithm is applied immediately after the adaptive equalization, everything that comes after that is not really interesting from our perspective.

The first test consists in inserting a PDL element in the considered system. For this purpose, a value of 3 dB for the PDL has been chosen.

Figure 4.1 shows the estimation result. In particular, only the values in the receiver filter band are shown, i.e.,  $f \in [-32 \text{ GHz}, 32 \text{ GHz}]$ . The reason is that the values which belong in the frequency range beyond the receiver filter cut-off frequency are not significant. Besides, this also allows to not consider the high-frequency values, which get distorted by the MMSE solution, yielding unreliable estimates.

In any case, the estimation results to be very close to the nominal value. Indeed, the estimate oscillates between the values 2.85 dB and 3.14 dB, with a mean value of  $\langle PDL(f) \rangle = 3.02$  dB. The error of the mean value with respect to the nominal one is just 0.02 dB. Even more interesting is the fact that the algorithm is capable of estimating the correct PDL, independently on the polarization state of the signal.

Indeed, the PDL value represents the maximum length of the loss interval the signal may experience, but the real value is somewhere in the middle.



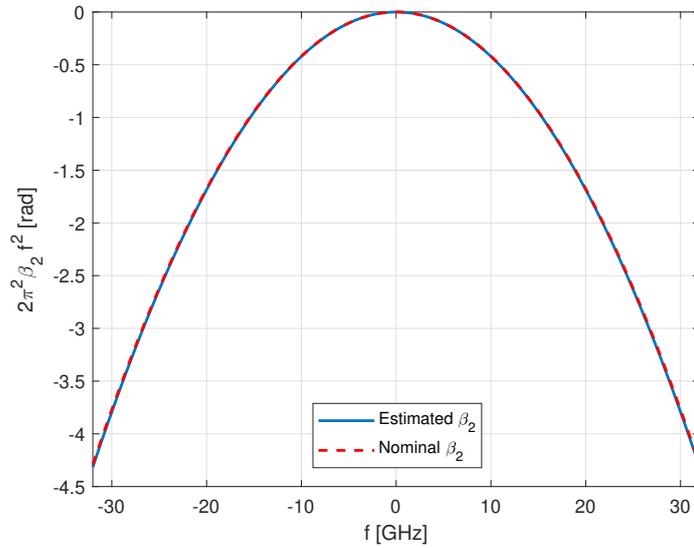
**Figure 4.1:** PDL estimation with OPM algorithm for a nominal value of 3 dB.

In order to test the residual CD estimation, instead, a CD quantizer has been inserted before the DSP. In this way, a controlled error is introduced in the nominal cumulated CD, generating a residual to be estimated.

For the considered case, the quantization grid has been built up considering a maximum error of 25% of the nominal cumulated CD over a single span, i.e.,  $-532 \text{ ps}^2$ , which translates into a step of  $1064 \text{ ps}^2$ . Moreover, given the length of the communication link ( $4 \times 100\text{-km}$  spans), the grid range has been chosen from a minimum of  $-16177 \text{ ps}^2$  up to a maximum of  $-212.85 \text{ ps}^2$ .

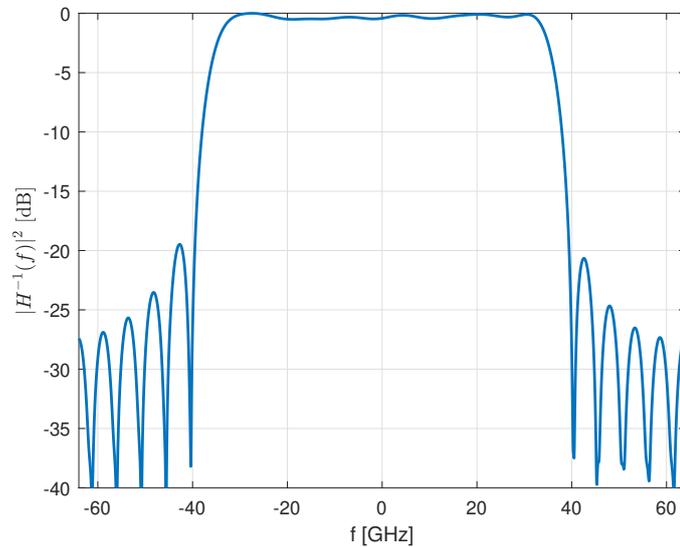
After quantizing, the residual CD which is introduced in the system is equal to  $-212.85 \text{ ps}^2$  and the quantized CD, which is also the value used in the fixed CD compensation stage of the DSP, is  $-8726 \text{ ps}^2$ .

From Figure 4.2 it is possible to learn that the estimate is in line with its nominal value. Indeed, the nominal curve (dashed line) and the estimated one (solid line) are basically overlapping. Furthermore, performing a quadratic fit to extract the parameter  $\beta_{2,res}$ , one gets an estimate equal to  $-213.5 \text{ ps}^2$ , which corresponds to an error of  $0.65 \text{ ps}^2$ .



**Figure 4.2:** Residual CD  $\beta_{2,res}$  estimation with OPM algorithm for a nominal value of  $-212.25 \text{ ps}^2$ .

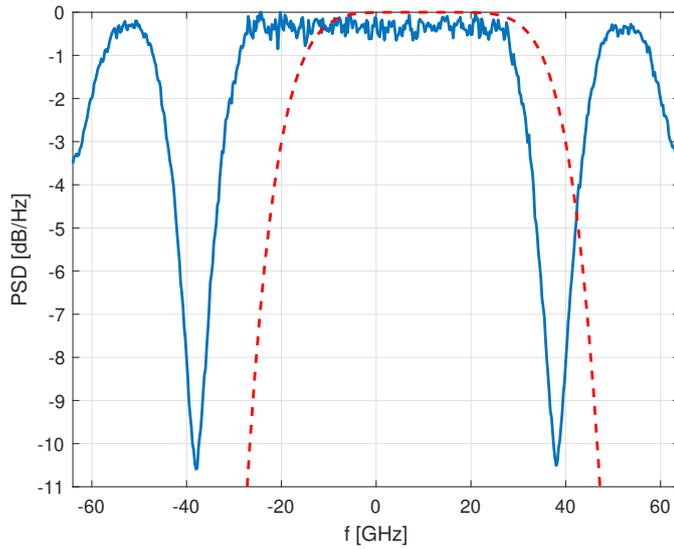
As for the electro-optical system transfer function, if no strong or particular filtering effect is present, the result is basically a lowpass filter, centered at the CUT frequency and with a flat response over the receiver filter bandwidth. An example is shown in Figure 4.3.



**Figure 4.3:** Electro-optical system transfer function estimation.

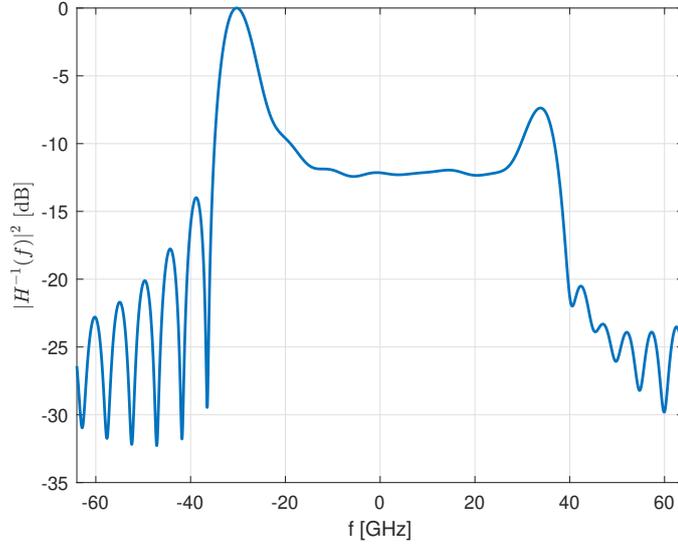
However, a more interesting aspect regarding this type of estimation is that it can be used to deduce if the filtering is tuned with the transmitter laser frequency or if the signal has gone through a strong filtering effect. A few examples of these cases will be presented.

Let us consider the case in which the signal gets filtered by a supergaussian filter of order  $m = 3$  and 3-dB bandwidth  $B_{3dB} = 60$  GHz, as reported in Figure 4.4. In particular, the filter is detuned by  $\Delta f_{detun} = 10$  GHz, meaning that its central frequency differs from the signal's central frequency by that amount.



**Figure 4.4:** Power spectral density of the received signal (blue solid line) and the supergaussian filter (red dashed line), with a detuning of  $\Delta f_{detun} = 10$  GHz.

Its effect on the estimation is that of creating two peaks at the borders of the receiver filter bandwidth, as can be noticed from Figure 4.5. The reason is that the equalizer tries to compensate for the bandwidth limitation that the filter introduces in the signal, localized at the borders of its spectrum. Besides, the two peaks are asymmetrical. This gives a hint on how the filter's detuning is characterized. Indeed, since the left peak is higher than the right one, it suggests that the filter's central frequency is located towards higher central frequencies, or easily speaking, it is simply moved "to the right". Of course, the same applies for a negative detuning. In that case, the two peaks are still present, but the right one is the highest, as one would expect.



**Figure 4.5:** Electro-optical system transfer function estimation when a detuning of  $\Delta f_{detun} = 10$  GHz is applied.

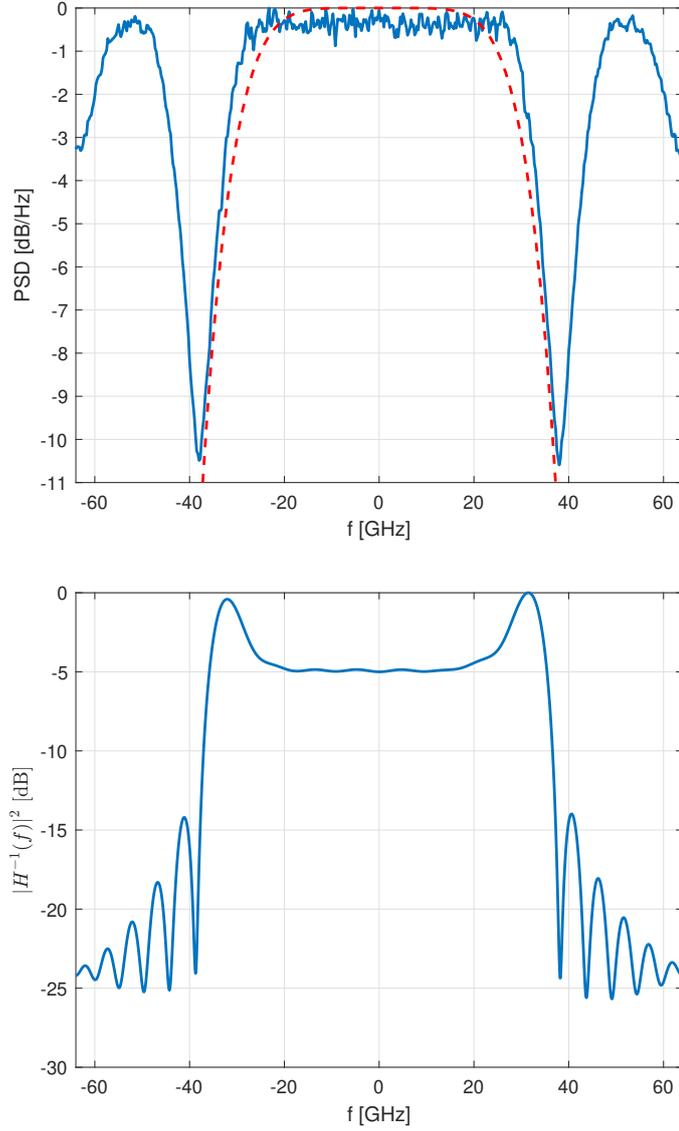
The same can be noticed in the case with no detuning, but with just a bandwidth limitation. Indeed, if the supergaussian filter is centered on the central frequency of the signal, the estimate of the system transfer function is more symmetrical and the two peaks have almost the same value.

An example is reported in Figure 4.6. Indeed, from the top figure, it is possible to learn that the applied filtering is narrower with respect to the signal's bandwidth. This introduces a bandwidth limitation which is basically equal for both band edges and results in two peaks, which are almost equal, in the estimated system transfer function, as shown in the bottom figure.

A more synthetic way to monitor bandwidth limitations is to consider the **peak-to-center (PTC) ratio**. It is defined as

$$PTC = 20 \log_{10} \left( \frac{\max(|H^{-1}(f)|)}{|H^{-1}(0)|} \right) \quad (4.23)$$

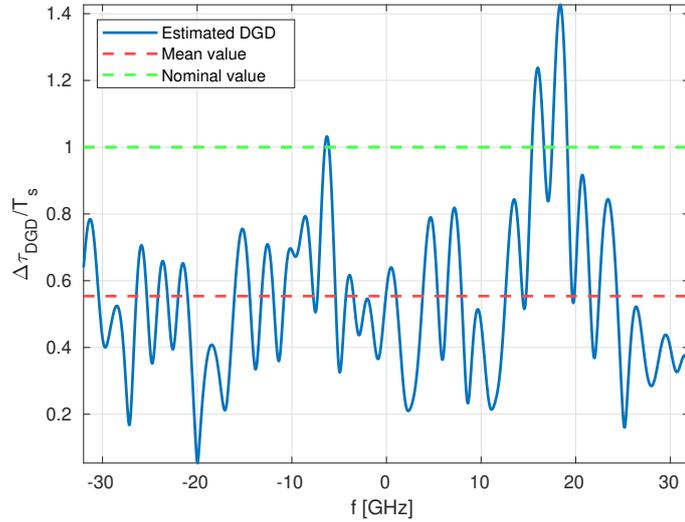
The stronger the bandwidth limitation, the higher the PTC. In the example, for a 3-dB bandwidth of the supergaussian filter equal to approximately  $0.94 \cdot R_s$ , the PTC results to be around 5 dB.



**Figure 4.6:** Top: power spectral density of the received signal (blue solid line) and the supergaussian filter (red dashed line) without detuning. Bottom: Electro-optical system transfer function estimation without detuning.

The last parameter to estimate is the DGD. It is important to mention the fact that the PMD is a stochastic effect, which depends on the polarization state of the signal in the fiber. Hence its estimate is in general between 0 and the nominal value. In order to have a more reliable PMD estimation, a continuous monitoring over time should be performed. Here, just a single example is proposed.

In this case, the parameter  $D_p$  has been set to  $0.7812 \text{ ps}/\sqrt{\text{km}}$ , which corresponds to a maximum DGD of one symbol time, i.e.,  $T_s = 15.625 \text{ ps}$ . Figure 4.7 reports the obtained result.



**Figure 4.7:** DGD estimation normalized with respect to  $T_s$  obtained with OPM algorithm, along with its mean value ( $0.55 \cdot T_s$ ) and the nominal value  $1 \cdot T_s$ . The PMD parameter  $D_p$  has been set to  $0.7812 \text{ ps}/\sqrt{\text{km}}$

As already mentioned, the mean value, which is equal to  $0.55 \cdot T_s$ , is around half the nominal value. Even if this estimation is not really significant, it has been presented for the sake of completeness.

This completes the section about simulation results. All of the discussed impairments have been estimated and the results have proven to be in good agreement with the expected ones. Now the OPM algorithm will be tested also in an experimental environment, so as to have a more robust validation of the results obtained.

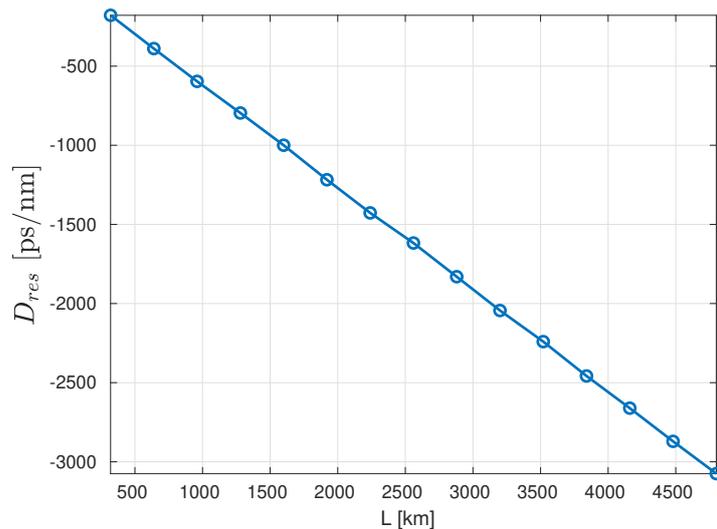
## 4.4 Experimental results

The experimental results have been obtained applying the OPM algorithm to the data collected from the experimental set-up already described in Sec. 3.4.3 and shown in Figure 3.12.

In particular, the signal was propagated over 15 recirculations (i.e., 4800 km) and the data have been collected for each of them. Plus, the channel power has been varied in the range  $P_{ch} \in [-5, -4, -3, -2.5, -2, -1.5, -1, -0.5, 0, 1, 2]$  dBm.

In order to have the best estimate possible, the estimate at the optimal power is chosen for each recirculation, corresponding to the  $P_{ch}$  value which is associated to the highest GMI. What is more, since in each case (i.e., for each recirculation and  $P_{ch}$ ) 5 waveforms have been collected, the median of the results is considered, so as to have a unique value and cut out possible outliers.

Figure 4.8 reports the estimation result for the residual CD.

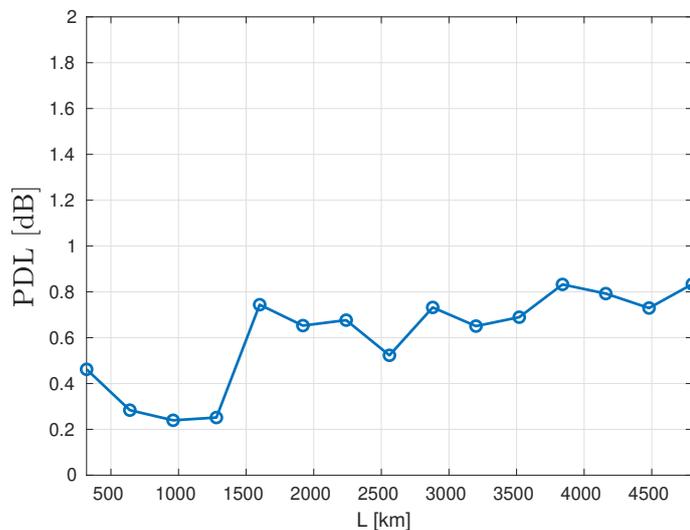


**Figure 4.8:** Residual CD  $D_{res}$  estimated over 15 recirculations (4800 km) from experimental data.

The residual CD turns out to cumulate linearly over successive recirculations. This is coherent with what one would expect, since the value of  $\beta_2$  or, equivalently,  $D$  inserted in the fixed CD compensator is generally a reasonable one, but it is not optimized. In the considered case, the nominal value for G.652 SMF fibers is used during the fixed CD compensation stage of the DSP, corresponding to  $\beta_2 = -21.27$  ps<sup>2</sup>/km or  $D = 16.7$  ps/nm/km.

This leads to a compensation which leaves out a certain amount of cumulated CD, generating the residual. In particular, such residual is always the same and keeps adding up at each recirculation, which explains the linear increase over the distance. Just to give some numbers, the residual CD at the end of each loop is around -200 ps/nm, leading to a total uncompensated CD which goes from -200 ps/nm up to around -3000 ps/nm over the whole propagation distance.

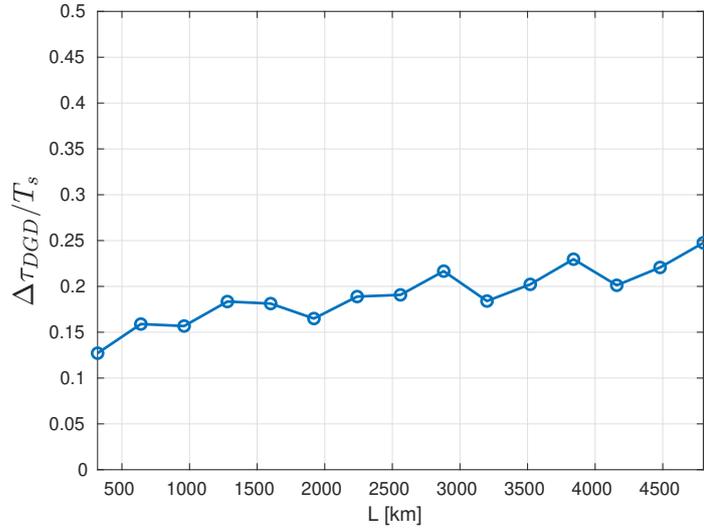
Then, also the PDL has been estimated and the result is shown in Figure 4.9. Since the OPM algorithm estimates the whole PDL spectrum, the reported values in the graph are the mean values over the frequency, i.e.,  $\langle PDL(f) \rangle$ .



**Figure 4.9:** Mean PDL estimated over 15 recirculations (4800 km) from experimental data with OPM algorithms.

The PDL value results to be generally low, also owing to the fact that no specific experimental set-up for PDL testing has been used. Furthermore, the presence of the polarization scrambler in the recirculating loop randomly modifies the polarization state of the signal during the propagation, leading to an averaging effect on the PDL itself. For the first loops, it is around 0.2-0.4 dB and then keeps growing up to around 0.8 dB. Overall, it is characterized by an increasing behavior, which is reasonable considering that the PDL effects combine as the signal propagates.

The same procedure has been applied for the DGD estimation. Once again, the mean value  $\langle \Delta\tau_{DGD}(f) \rangle$  has been computed and reported with respect to the transmission distance, as in Figure 4.10. As already done for the simulation results, the DGD value is normalized with respect to the symbol time  $T_s = 62.5$  ps.



**Figure 4.10:** Mean DGD estimated over 15 recirculations (4800 km) from experimental data with OPM algorithms.

Similarly to PDL, DGD estimation presents the same increasing behavior. In particular, the estimated distance in time between the two polarizations doubles when moving from the first loop to the 15<sup>th</sup> one.

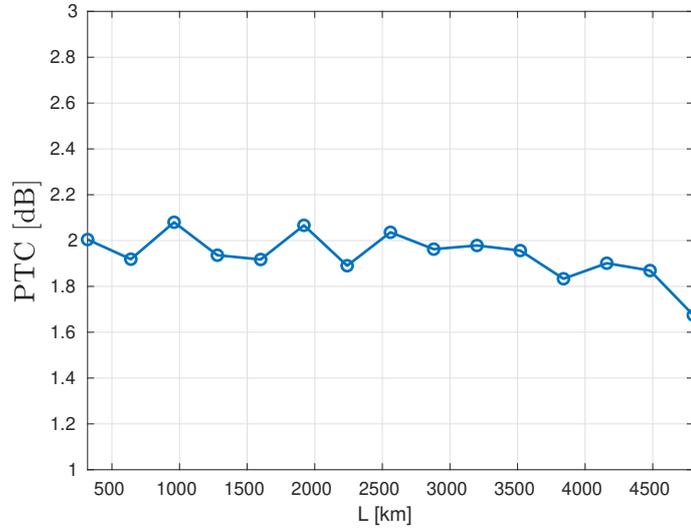
Indeed, the mean  $\Delta\tau_{DGD}$  goes from a minimum of 12.5% up to a maximum of 25% of a symbol time over the successive recirculations. Also this result is in line with what one would expect.

Finally, also the system scalar transfer function is estimated. For simplicity, as discussed in Sec. 4.3, the potential bandwidth limitation is monitored according to the PTC ratio, expressed in Eq. (4.23).

This operation makes it easier to monitor the filtering effects. From Figure 4.11 it is possible to learn that the PTC for each recirculation is always around 2 dB, despite a few oscillations.

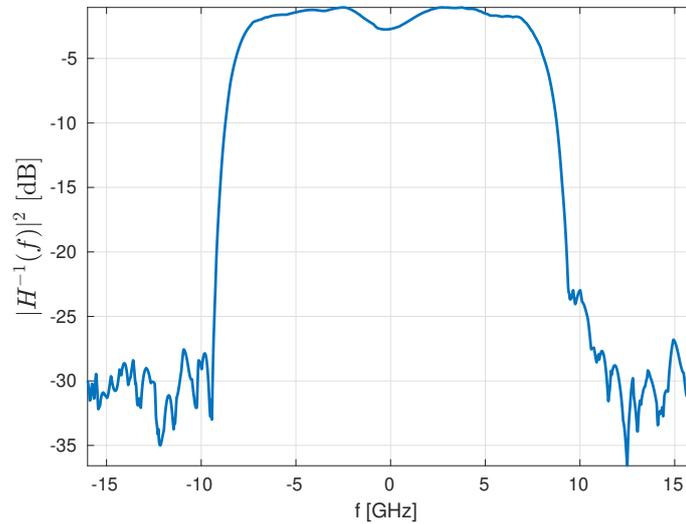
This translates into the fact that no strong bandwidth limitations are introduced when the signal gets filtered. This is coherent with the experimental set-up.

Indeed, the optical bandpass filter in the recirculating loop has a relatively large bandwidth because the WDM channels are widely spaced. This means that the signal's bandwidth is smaller (in the range of 16 GHz) when compared to the transmitter or receiver bandwidth.



**Figure 4.11:** PTC estimated over 15 recirculations (4800 km) from experimental data with OPM algorithms.

To conclude this section about the experimental results, also an example of system scalar transfer function estimation is reported. In particular, it is taken from the last recirculation and shown in Figure 4.12.



**Figure 4.12:** System scalar transfer function  $|H^{-1}(f)|^2$  estimated over 15 recirculations (4800 km) from experimental data with OPM algorithms.

Of course, real filters are not ideal, as the ones used during simulations. This explains the lack of symmetry. Besides, it is also necessary to consider the fact that in general a frequency detuning is present between the lasers at the transmitter and at the receiver side, which also contributes to the asymmetry of the estimated spectrum. In any case, this result confirms what has been said about the system while discussing the estimated PTC and is also very similar to the ones obtained by means of simulations (e.g., Figure 4.3).

Generally speaking, OPM algorithms have proven to be capable of supplying a good estimation of the parameters related to the major system linear impairments, as well as draw some conclusions about the system itself (e.g., strong filtering effect and detuning). In Chapter 5, a further step forward is taken, in order to implement more advanced OPM techniques which are capable of exploiting both optical fiber non-linearities and some optimization techniques related to machine learning.

# Chapter 5

## Digital longitudinal monitoring

### 5.1 Introduction

One of the concepts that has mainly contributed to the recent progress in optical fiber communications is **system identification (SI)**. Practically speaking, SI attempts to supply a simplified model of the system itself or extract its internal parameters based on boundary conditions, i.e., input and output data. This allows for two main types of applications:

1. Possibility to **predict** the outputs of a system given an arbitrary set of input data and vice versa. Generally with a lower complexity.
2. Extraction of system parameters, which enables its **monitoring**

Examples of the first application have already been presented in the previous chapters. Let us recall the SSFM algorithm to numerically solve the NLSE or the DBP algorithm to compensate for both CD and NL effects (Chapter 3 Sec. 3.4.1). However, the main subject of this work is related to the second application. It involves the identification of the system parameters starting from the received signals, in order to monitor the system status. Also in this case, some techniques have been presented in Chapter 4, which allowed to estimate residual CD, PDL, DGD and filtering effects. The aspect that all these parameters have in common, though, is that they refer to **cumulative** quantities. Indeed, the assumption which has been made when discussing the aforementioned techniques is that the system works in a **weakly non-linear regime**. This means that in a mostly linear system, **superposition of effects** holds and the single effects add up along the propagation over the communication link. For this reason, at the receiver side, it is not possible

to separate those effects and localize them along the line. Only the cumulated effect can be estimated.

The step forward which is taken in this chapter regards exactly this aspect. Fiber **longitudinal** parameters, e.g., longitudinal loss profile and CD maps, are also necessary to increase the reliability of the system and maximize its capacity. Hence, all the elements of a system get tested with respect to these aspects both before and during their operation. Up to now, these tests have been performed by means of analog devices, such as OTDR's and OSA's. However, this is a costly and time-consuming operation, since these devices should be placed on-site and in a span-by-span fashion. Thus, if this operation could be performed through a **digital SI** approach, that would lead to a strong reduction of both capital and operational expenditures (CAPEX and OPEX). Optical systems would become "smarter" and more automatization could be achieved for what concerns measurements and fault detection. In general, these techniques to estimate fiber parameters in a **distributed** way and relying only on **DSP** go under the definition of **digital longitudinal monitoring (DLM)**.

In this chapter, at first a brief description of the concept of neural network (NN) is given, together with a couple of optimization algorithms related to it, namely gradient descent and stochastic gradient descent algorithms. Then, the enablers of DLM are presented, including the important role of fiber non-linearity and the intrinsic similarity of NN's and SSFM which inspired this technique. Finally, a **channel reconstruction method (CRM)** to implement DLM and estimate the longitudinal loss profile in a link is described together with some simulation results.

The main references for this chapter are [15] and [16].

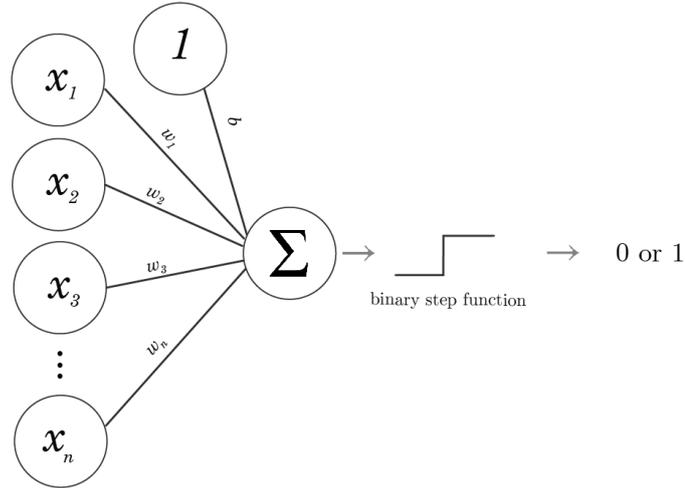
## 5.2 Basic concepts

Before going into the details of the DLM algorithm, it is convenient to introduce the concept of neural network and its learning process, since it will be exploited in the following sections. Hence its basic elements are illustrated, along with their functioning and the overall NN structure. Then, the learning process is discussed, according to two optimization techniques: **gradient descent (GD) algorithm** and **stochastic gradient descent (SGD) algorithm**, with particular attention to the latter. Of course, an in-depth discussion over these topics falls out of the scope of this work. Only the very basic concepts and general descriptions will be given.

### 5.2.1 Neural networks

Neural networks are computing systems belonging to the field of **machine learning**. The latter refers to a branch of artificial intelligence (AI) and closely belongs to the first application of SI described in the introduction of this chapter. In particular, machine learning involves finding models, training such models with training data to fit them and make some decision rules. The problem is that in general it is hard to define a good model and write computer programs to "describe" the data. Whence, the introduction of NN's. Indeed, NN's **learn** both the model and the decision rules. Mathematically speaking, they simply implement a multidimensional **non-linear** function, which attempts to predict the output of a system given a set of arbitrary data.

Its most basic element is the **neuron**, whose structure is reported in Figure 5.1.



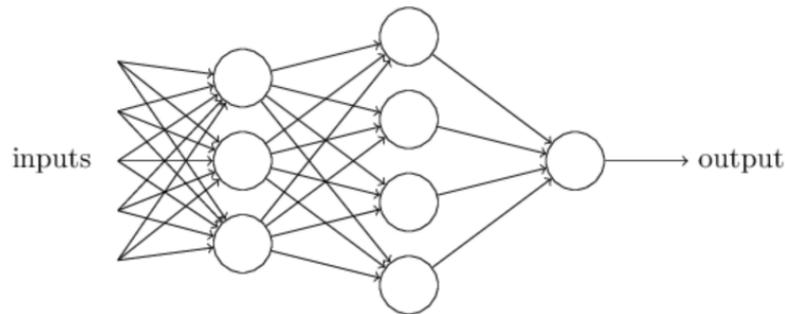
**Figure 5.1:** Perceptron structure, with inputs  $x_i$ , weights  $w_i$  and bias  $b$ .

Its functioning consists in weighing different inputs in order to make a binary decision. Basically, the inputs  $\mathbf{x} = [x_1, \dots, x_N]^T$  get weighed by a set of weights  $\mathbf{w} = [w_1, \dots, w_N]^T$ , summed and compared to a threshold  $b$ , said bias. The output is then either 0 or 1, i.e., a step function. In this particular case, the neuron is also said **perceptron**. However, to improve the learning process, more elaborated functions are applied to the linear combination of inputs and weights. So, generally the output of a neuron can be written as

$$y = f(\mathbf{w}^T \cdot \mathbf{x} + b) \tag{5.1}$$

where  $\cdot$  denotes the scalar product and  $f$  is a non-linear function, called **activation function**, with a smoother behavior and better differentiability properties than the step function.

The neuron is just the basic element in a NN. However, in order to make more complex decisions, neurons are organized in layers and multiple layers compose the whole NN. Then, neurons from one layer are connected to those of the successive layer. In this way, the output of one neuron becomes the input to the next one. Figure 5.2 gives a representation of the generic structure of a **feed-forward NN**. Other types exist, but they are of no interest for this discussion.



**Figure 5.2:** Generic structure of a neural network with multiple neuron layers, also said feed-forward neural network.

It is clear, then, that the correct functioning of a NN depends on the value of its weights  $\mathbf{w}$  and biases  $\mathbf{b}$ . Hence, they need to be optimized according to some optimization techniques. Plus, an indicator to monitor how the optimization process is going needs to be introduced. In particular, the latter is generally referred to as **cost function**, i.e., a function which evaluates the error between the inferred output  $\hat{\mathbf{y}}$  of the NN and the "correct answer"  $\mathbf{y}$  given an input  $\mathbf{x}$ . Among the most common ones is the **mean squared error (MSE)** cost function, defined as

$$C(\mathbf{w}, \mathbf{b}) = \frac{1}{N} \sum_{n=1}^N \|\hat{\mathbf{y}}_n(\mathbf{x}_n) - \mathbf{y}_n(\mathbf{x}_n)\|^2 \quad (5.2)$$

Note that MSE is also the cost function used to optimize the tap coefficients in adaptive equalizers (Chapter 3 Sec. 3.3.2).

For this purpose, the process to have the NN learn the optimal parameters is composed of several steps: pick a set of training inputs, propagate them along the NN, compute the cost function and apply an optimization algorithm to update all the weights and biases so as to decrease the cost function. These optimization algorithms are the subject of the next sections and belong to the **gradient-based learning** approach.

## 5.2.2 Gradient descent algorithm

Let us start by defining the problem. Given a function  $C(\mathbf{v})$ , depending on several parameters  $\mathbf{v} = [v_1, \dots, v_N]^T$ , the objective is that of finding the values for  $\mathbf{v}$  which minimize  $C(\mathbf{v})$ . The discussion applies in general for a high number of parameters to optimize, as in NN's. However, for simplicity, two dimensions are considered to explain the algorithm functioning, i.e.,  $\mathbf{v} = [v_1, v_2]^T$ .

When the values of the parameters in  $\mathbf{v}$  are changed by a small enough amount  $\Delta\mathbf{v} = [\Delta v_1, \Delta v_2]^T$ , the variation of  $C$  can be written according to its Taylor expansion as

$$\Delta C(\mathbf{v}) \approx \frac{\partial}{\partial v_1} \Delta v_1 + \frac{\partial}{\partial v_2} \Delta v_2 \quad (5.3)$$

Since the objective of the optimization algorithm is that of minimizing  $C(\mathbf{v})$ , then it is necessary to find  $\Delta\mathbf{v}$  such that  $\Delta C(\mathbf{v})$  is negative. In a more compact notation, given the gradient of  $C(\mathbf{v})$ , defined as

$$\nabla C(\mathbf{v}) = \left[ \frac{\partial}{\partial v_1} C(\mathbf{v}), \frac{\partial}{\partial v_2} C(\mathbf{v}) \right]^T \quad (5.4)$$

it is possible to write

$$\Delta C(\mathbf{v}) \approx \nabla C(\mathbf{v})^T \cdot \Delta\mathbf{v} \quad (5.5)$$

Now, let us suppose that the chosen variation of the parameters  $\mathbf{v}$  is  $\Delta\mathbf{v} = -\eta \nabla C(\mathbf{v})$ , then Eq. (5.5) becomes

$$\Delta C(\mathbf{v}) \approx -\eta \|\nabla C(\mathbf{v})\|^2 \quad (5.6)$$

which is an always negative quantity. In particular,  $\eta$  is the **learning rate** and needs to be small enough in order for the Taylor approximation to hold.

In conclusion, the update of the parameters  $\mathbf{v}$  can be written as

$$\mathbf{v} \leftarrow \mathbf{v} - \eta \nabla C(\mathbf{v}) \quad (5.7)$$

and the whole algorithm goes under the name of **gradient descent (GD)** algorithm. It can be proved that GD algorithm converges to a **local minimum** of the function  $C$  and does not require the computation of second derivatives, i.e., Hessian matrices. As already mentioned, this type of algorithm is used for the updating process of the tap coefficients for adaptive equalization, but it can also be used for the update of weights and biases in a NN considering the cost function as the aforementioned function  $C(\mathbf{v})$ . However, the problem becomes computationally complex, since the number of parameters to optimize in a NN is huge.

### 5.2.3 Stochastic gradient descent algorithm

One issue related to the GD algorithm is that it becomes slow when the training set of data becomes large. The reason is that all the inputs need to be propagated and the gradient is computed only at the end, averaging over the whole training set. This leads to the second optimization algorithm, which is a modification of the GD, i.e., the **stochastic gradient descent (SGD)** algorithm.

The idea is simple. Instead of estimating the cost function gradient from all the training inputs, such estimation is carried out considering only a subset of inputs, called **mini-batch**. Therefore, if only a mini-batch  $\mathbf{x}^{MB} = [x_1, \dots, x_M]^T$  is considered, with  $M < N$  and  $N$  being the training set dimension, one expects that

$$\frac{1}{M} \sum_{m=1}^M \nabla C(x_m^{MB}) \approx \frac{1}{N} \sum_{n=1}^N \nabla C(x_n) = \nabla C(\mathbf{x}) \quad (5.8)$$

In other words, avering over the mini-batch yields a gradient estimate which approximates the gradient estimate obtained considering the whole training set  $\mathbf{x}$  (GD algorithm). This algorithm proves to be faster from a computational point of view and takes out the need to download the whole training set. This is repeated for each mini-batch, paying attention to the fact that a mini-batch cannot contain the same elements which have already been used for another one. No overlapping is allowed.

For the particular case of NN, the SGD algorithm is generally used. Of course, in order to apply it, the computation of the gradient for all the weights and biases is required. This is achieved by means of the **back-propagation** algorithm which supplies simple expressions to compute all of these derivatives. In particular, the idea is that of starting from the end of the NN and "back-propagate" the gradient so as to compute all the others. A further discussion of this algorithm, though, is not reported, since the system presented in Sec. 5.3 is slightly different from traditional NN's and a specific back-propagation algorithm is presented, starting from DBP.

## 5.3 Distributed optical performance monitoring

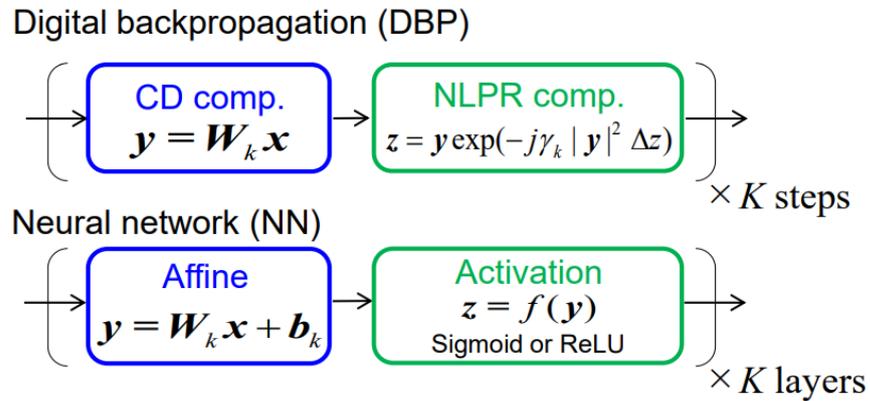
After introducing all the basic concepts related to NN and gradient-based optimization techniques, a **channel reconstruction method (CRM)** is presented, so as to perform DLM. Hence, in this section the similarity between NN's and the structure of the DBP algorithm is highlighted, to justify the CRM method. Then, an analytical description of such method is provided and a description of the

important role of fiber non-linearity is given, in relation to the non-commutativity property of the linear and non-linear operators.

### 5.3.1 NN-DBP parallelism

Two approaches are possible to implement SI techniques. The first one is said **black-box** and an example is machine learning together with NN's. It has already been discussed how NN's can be trained in order to implement a multidimensional non-linear function which is capable of predicting the system outputs from the system inputs. However, this type of approach is not really suitable for performance monitoring purposes. Indeed, the term "black-box" refers to the fact that everything one knows is related to the input/output of the system. It is in general very difficult, if not impossible, to give an explanation of the learned parameters. Therefore, though these methods are generally powerful, they are not useful for what concerns useful system parameters extraction.

Nonetheless, optical fiber communications supply an interesting way to move from a purely data-driven approach to a more structure-aware one. This is exactly the second approach, i.e., a **behavioral approach**. It is simply necessary to notice the similarity between the structure of a NN and that of the SSFM or, equivalently, DBP algorithm. In particular, both structures involve a sequence of linear and non-linear operations which is repeated in an iterative way and in a cascaded fashion, as Figure 5.3 shows.



**Figure 5.3:** Similarity of DBP and NN structure. Cascaded blocks iteratively implementing a linear operation (CD compensation/linear combination of weights and biases) and a non-linear operation (non-linear phase rotation compensation (NLPR)/activation function). Note: Sigmoid and ReLU are just examples of activation function. Taken from [16].

Therefore, this similarity allows for an easier physical interpretation of the learned parameters (e.g.,  $\beta_2$ ,  $\gamma$  etc) and the possibility to exploit the optimization techniques which have been developed in the field of machine learning. Even more importantly, it allows for a distributed performance monitoring of those parameters, e.g., the longitudinal loss profile.

### 5.3.2 Problem formulation

Let us first recall the model for the propagation of a polarization-multiplexed signal in an optical fiber. It is given by the Manakov equations, as in Eq. (2.27) and here reported with some modifications:

$$\frac{\partial}{\partial z} \mathbf{E}'(t, z) = j \frac{\beta_2(z)}{2} \frac{\partial^2}{\partial t^2} \mathbf{E}'(t, z) - \alpha(z) \mathbf{E}'(t, z) - j\gamma(z) \frac{8}{9} \|\mathbf{E}'(t, z)\|^2 \mathbf{E}'(t, z) \quad (5.9)$$

where  $\mathbf{E}'(t, z) = [E'_x(t, z), E'_y(t, z)]^T$  is the Jones vector of the optical field characterized by an instantaneous power of  $\|\mathbf{E}'(t, z)\|^2 = |E'_x(t, z)|^2 + |E'_y(t, z)|^2$ . The modification consists in introducing a spacial dependency on the main fiber parameters, namely  $\alpha(z)$ ,  $\beta_2(z)$  and  $\gamma(z)$ . What is more, if the field with normalized amplitude

$$\mathbf{E}(t, z) = \mathbf{E}'(t, z) e^{\int_0^z \alpha(z') dz'} \quad (5.10)$$

is substituted in Eq. (5.9), then it is possible to merge together  $\alpha(z)$  and  $\gamma(z)$ . In this way, all the contributions to the variation of the signal power can be imputed to a single parameter  $\gamma'(z)$ . Eq. (5.9) then becomes

$$\frac{\partial}{\partial z} \mathbf{E}(t, z) = j \frac{\beta_2(z)}{2} \frac{\partial^2}{\partial t^2} \mathbf{E}(t, z) - j\gamma'(z) \|\mathbf{E}(t, z)\|^2 \mathbf{E}(t, z) \quad (5.11)$$

where

$$\gamma'(z) = \frac{8}{9} \gamma(z) e^{\int_0^z \alpha(z') dz'} \quad (5.12)$$

This equation can be numerically solved making use of the SSFM, iteratively computing first the linear part and then the non-linear one independently. To retrieve the original transmitted signal, instead, DBP is available. In particular, defining the linear and non-linear operators in Eq. (5.11) respectively as

$$\hat{\mathcal{D}} = j \frac{\beta_2(z)}{2} \frac{\partial^2}{\partial t^2} \quad (5.13)$$

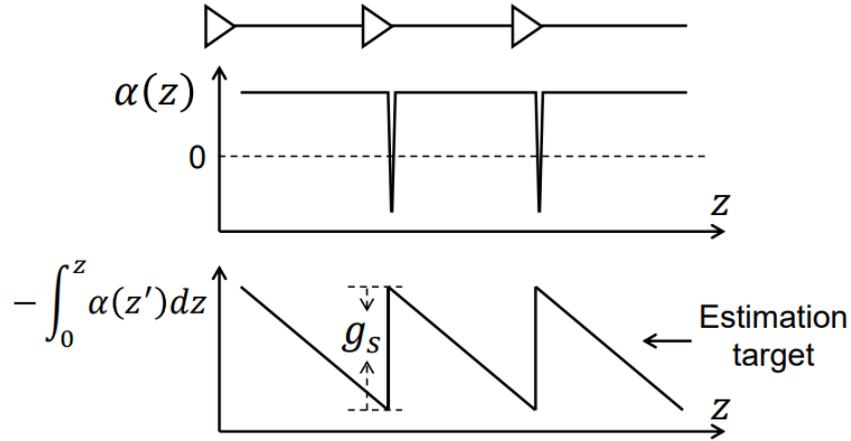
$$\hat{\mathcal{N}} = -j\gamma'(z) \|\mathbf{E}(t, z)\|^2 \quad (5.14)$$

the basic DBP step to go from position  $z_k$  to the position  $z_{k-1}$  is given by

$$\hat{\mathbf{E}}(t, z_{k-1}) = \hat{\mathbf{E}}(t, z_k) e^{\hat{\mathcal{N}}^{-1} \Delta z} e^{\hat{\mathcal{D}}^{-1} \Delta z} \quad (5.15)$$

where  $\hat{\mathcal{N}}^{-1} = j\gamma'(z)\|\mathbf{E}(t, z)\|^2$  and  $\hat{\mathcal{D}}^{-1} = -j\frac{\beta_2(z)}{2}\frac{\partial^2}{\partial t^2}$  are the inverse linear and non-linear operators. Plus,  $z_k$  represents the position on the z-axis along the fiber length, with  $k = 0, \dots, K$  and  $K$  being the total number of steps  $\Delta z = z_k - z_{k-1}$ . In particular  $z_0 = 0$  and  $z_K = L$ , where  $L$  is the total link length. DBP consists then in a cascaded structure of CD compensation and NL compensation. In the CRM method, it will be used as the model for the inverse system of the optical fiber and referred to as **channel emulator**. Besides, the target of the estimation is also defined and results to be  $\gamma'(z_k)$ , for each  $k = 0, \dots, K$ , which is related to the **longitudinal fiber loss profile**.

Note that  $\gamma'(z)$  contains the integral contribution with respect to the fiber loss  $\alpha(z)$ . Therefore, CRM does not estimate the fiber loss parameter itself, but its integral version. Figure 5.4 gives a representation of this aspect.



**Figure 5.4:** Behavior of  $\alpha(z)$  and its integral along a multi-span optical system. The spikes in the curve representing  $\alpha(z)$  are given by the amplifiers' gains  $g_s$  at the end of each span. Taken from [15].

Once defined both a model for the optical link and the target of the DLM estimation, it is necessary to formulate the estimation problem to solve.

Typically, in a plain application of DBP, the parameters  $\gamma'(z)$  and  $\beta_2(z)$  are unknown. They are set to reasonable values or according to the result of measuring procedures. In this case, they are obtained as those values which better emulate the channel so that the received signal goes back to the originally transmitted one. Hence, the optimization problem can be formulated as a classical minimization

problem in the sense of least squares. Practically speaking,

$$\begin{aligned}\hat{\gamma}'(z_k) &= \arg \min I \\ &= \arg \min \sum_{n=0}^{N_s-1} \left\| \mathbf{E}(0, nT) - \hat{\mathbf{E}}(0, nT) \right\|^2\end{aligned}\quad (5.16)$$

where  $\arg \min$  is the operator which extracts the value which minimizes its argument,  $I$  is the **cost function**,  $N_s$  is the total number of samples in the signal,  $T$  is the sampling interval,  $\mathbf{E}(0, nT)$  is the originally transmitted field at position  $z = 0$  and  $\hat{\mathbf{E}}(0, nT)$  is the output of the channel emulator. This problem can be solved by means of the SGD algorithm and considering specific expressions for  $\partial I / \partial \gamma'(z_k)$ , which implement a procedure similar to the **back-propagation** algorithm in NN's.

Note that the cost function  $I$  requires the knowledge of the transmitted signal  $\mathbf{E}(0, nT)$ . However, it can be reconstructed at the receiver side following the usual demodulation process in the DSP. This relieves of the usage of training sequences or dedicated monitoring channels.

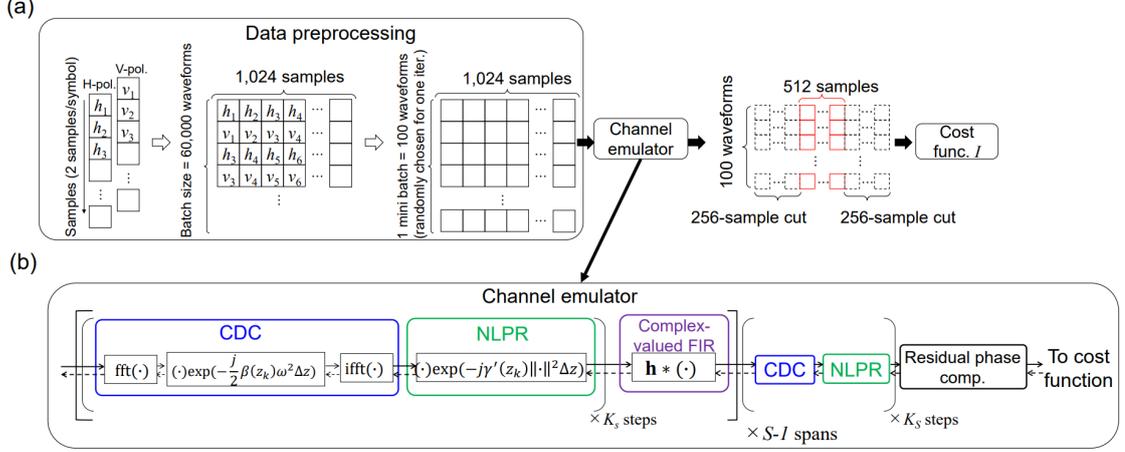
### 5.3.3 Longitudinal loss profile estimation

The DLM algorithm for longitudinal loss profile estimation is composed of four steps: **data pre-processing**, **channel emulator**, **cost function computation** and **gradient back-propagation**.

At the input of the CRM the optical signal is polarization-demultiplexed. Hence, it is composed of a vector for the x-polarization (or, H-polarization) and a vector for the y-polarization (or, V-polarization). During **data pre-processing**,  $N_{samp}$  samples are alternatively taken from each polarization with a step of 2 samples. These groups of samples are called **waveform** and get stored within a structure called **batch**. Then, a smaller structure, called **mini-batch**, of  $N_{wf}$  randomly chosen waveforms is sent as input to the channel emulator.

In the **channel emulator** the DBP is performed according to the algorithm which has already been discussed. In particular, the reconstructed channel gets divided into  $S$  spans and  $K_s$  DBP steps are performed for each span, composed of a CD compensation (CDC) block and a non-linear phase rotation (NLPR) compensation block. In the end, the residual phase term gets also compensated for and the output is sent to the cost function block. In general, due to the fact that the samples at the edges of a waveform are not properly compensated with respect to CD during DBP, 1/2 of the samples from each waveform are cut out before computing the cost function  $I$ . Half of those cut out samples are taken from the first samples and

the other half from the last samples. Figure 5.5 gives a graphical representation of this algorithm.



**Figure 5.5:** (a) DLM algorithm scheme and the involved data structures for the estimation of  $\gamma'(z_k)$ . In particular, in this example waveforms are composed of  $N_{samp} = 1024$  samples, a batch of 60000 waveforms and a mini-batch of  $N_{wf} = 100$  waveforms. (b) More detailed structure of DBP-based channel emulator. Note: the complex-valued FIR filter is not considered in this work. Taken from [15].

After computing the **cost function**  $I$ , its gradient is also computed and the learning process based on **gradient back-propagation** begins. As anticipated in Sec. 5.2.3, **error back-propagation** is an algorithm in NN's to compute the gradient of the cost function with respect to all the parameters to optimize. In particular, the term "back-propagation" refers to the fact that the gradient with respect to one parameter is obtained from the one previously computed at the successive layer. Here, the same is performed, considering the channel emulator as a NN with one neuron (CDC+NLPR, see Figure 5.3) per layer and  $S \cdot K_s$  layers.

Let us define the input waveform to each block in the channel emulator as the  $N_{samp} \times 2$  vector  $\mathbf{x} = [\mathbf{x}_H, \mathbf{x}_V]^T$  and its output as  $\mathbf{y} = [\mathbf{y}_H, \mathbf{y}_V]^T$ . The 2 in the vector dimensions represents the two polarizations. The first step is to compute the gradient of the cost function  $I$  with respect to the complex conjugated output of the last block in the channel emulator. In particular, its structure is

$$\frac{\partial I}{\partial \mathbf{y}^*} = \left[ \frac{\partial I}{\partial y_1^*}, \dots, \frac{\partial I}{\partial y_{N_{samp}}^*} \right]^T \quad (5.17)$$

and is the same for both the H- and V-polarization.

Given that complex-valued quantities are involved, this gradient, as well as all the

other ones, are computed according to Wirtinger calculus. Hence, considering the waveforms according to their real and imaginary parts, i.e.,  $\mathbf{y} = \mathbf{y}_R + j\mathbf{y}_I$ , it is possible to apply the Wirtinger linear partial differential operator

$$\frac{\partial}{\partial z^*} = \frac{1}{2} \left( \frac{\partial}{\partial z_R} + j \frac{\partial}{\partial z_I} \right) \quad (5.18)$$

where  $z = z_r + jz_I$  is a complex-valued variable. Applying this operator to the argument of Eq. (5.16) yields

$$\frac{\partial I}{\partial \mathbf{y}^*} = \mathbf{y} - \hat{\mathbf{y}} \quad (5.19)$$

where  $\hat{\mathbf{y}} = \hat{\mathbf{y}}_R + j\hat{\mathbf{y}}_I$  is the reconstructed transmitted signal. After this operation, the gradient needs to be back-propagated. For this reason, the back-propagation expressions of the channel emulator blocks, namely FFT, CDC, IFFT and NLPR, are to be defined. It is important to remember that during DBP, the input and the outputs to each block are  $\mathbf{x}$  and  $\mathbf{y}$ , respectively.

On the contrary, during back-propagation, the gradient is propagated in the opposite direction. Hence, the input to each block becomes  $\partial I / \partial \mathbf{y}^*$  and the output  $\partial I / \partial \mathbf{x}^*$ . In particular, back-propagation is indicated in Figure 5.5 by means of dashed arrows between blocks.

The first block considered is the CDC block. The operation implemented by the CDC block during DBP at position  $z_k$  is

$$\mathbf{y} = \mathcal{F}^{-1} \left\{ \mathcal{F} \{ \mathbf{x} \} e^{j2\pi^2 \beta_2(z_k) f^2 \Delta z} \right\} \quad (5.20)$$

where  $\mathcal{F}$  and  $\mathcal{F}^{-1}$  are the FFT and IFFT operators, respectively. Considering the fact that the back-propagation expression of the FFT operator is the IFFT and vice versa, the whole CDC back-propagation expression can be written as

$$\frac{\partial I}{\partial \mathbf{x}^*} = \mathcal{F}^{-1} \left\{ \mathcal{F} \left\{ \frac{\partial I}{\partial \mathbf{y}^*} \right\} e^{-j2\pi^2 \beta_2(z_k) f^2 \Delta z} \right\} \quad (5.21)$$

where the frequency domain CDC expression is computed as

$$\begin{aligned} \frac{\partial I}{\partial \mathbf{x}^*} &= \frac{\partial I}{\partial \mathbf{y}} \frac{\partial \mathbf{y}}{\partial \mathbf{x}^*} + \frac{\partial I}{\partial \mathbf{y}^*} \frac{\partial \mathbf{y}^*}{\partial \mathbf{x}^*} \\ &= 0 + \frac{\partial I}{\partial \mathbf{y}^*} e^{-j2\pi^2 \beta_2(z_k) f^2 \Delta z} \end{aligned} \quad (5.22)$$

Then the NLPR back-propagation expression is computed. Also in this case, the operation performed by it at position  $z_k$  is

$$\mathbf{y} = \mathbf{x} e^{j\gamma'(z_k) \|\mathbf{x}\|^2 \Delta z} \quad (5.23)$$

hence its back-propagation becomes

$$\begin{aligned}
 \frac{\partial I}{\partial \mathbf{x}^*} &= \frac{\partial I}{\partial \mathbf{y}} \frac{\partial \mathbf{y}}{\partial \mathbf{x}} + \frac{\partial I}{\partial \mathbf{y}^*} \frac{\partial \mathbf{y}^*}{\partial \mathbf{x}^*} \\
 &= \frac{\partial I}{\partial \mathbf{y}} (j\gamma'(z_k)\Delta z) \mathbf{x}^2 e^{j\gamma'(z_k)\|\mathbf{x}\|^2\Delta z} \\
 &\quad + \frac{\partial I}{\partial \mathbf{y}^*} (1 - j\gamma'(z_k)\|\mathbf{x}\|^2\Delta z) e^{-j\gamma'(z_k)\|\mathbf{x}\|^2\Delta z}
 \end{aligned} \tag{5.24}$$

where  $\partial I/\partial \mathbf{y} = (\partial I/\partial \mathbf{y}^*)^*$ ,  $\mathbf{x}^2$  denotes the pointwise square operator applied to  $\mathbf{x}$  and  $\|\mathbf{x}\|^2 = |\mathbf{x}_H|^2 + |\mathbf{x}_V|^2$ .

In order to estimate the longitudinal loss profile, also the derivative of  $I$  with respect to each  $\gamma'(z_k)$  is required. Otherwise, it is not possible to perform the SGD algorithm to update the parameters and have the system learn the optimal ones. At this purpose, it is possible to write according to Eq. (5.23)

$$\begin{aligned}
 \frac{\partial I}{\partial \gamma'(z_k)} &= \frac{\partial I}{\partial \mathbf{y}} \frac{\partial \mathbf{y}}{\partial \gamma'(z_k)} + \frac{\partial I}{\partial \mathbf{y}^*} \frac{\partial \mathbf{y}^*}{\partial \gamma'(z_k)} \\
 &= 2\Re \left\{ \frac{\partial I}{\partial \mathbf{y}^*} \frac{\partial \mathbf{y}^*}{\partial \gamma'(z_k)} \right\} \\
 &= 2\Re \left\{ \frac{\partial I}{\partial \mathbf{y}^*} \cdot \mathbf{x}^* (-j\|\mathbf{x}\|^2\Delta z) e^{-j\gamma'(z_k)\|\mathbf{x}\|^2\Delta z} \right\}
 \end{aligned} \tag{5.25}$$

where  $\cdot$  denotes the scalar product and  $\Re$  the real part operator.

At this point the algorithm is complete. At each iteration, a mini-batch is sent as input to the channel emulator, a cost function is computed, the gradient gets back-propagated and the SGD algorithm is applied to the parameters  $\gamma'(z_k)$  so as to update and optimize them. This last operation is the same as the one described in Eq. (5.7), i.e.,

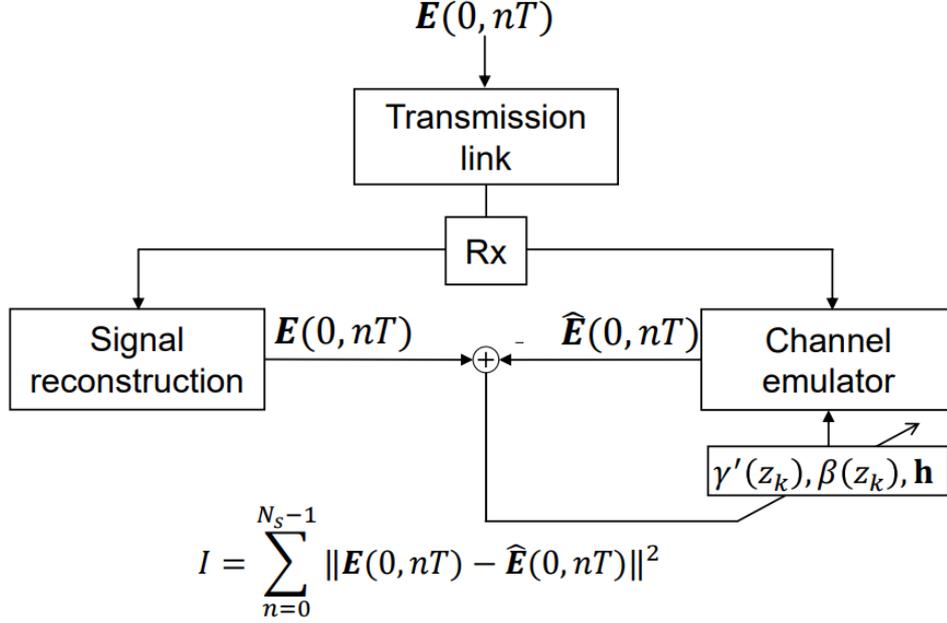
$$\gamma'(z_k) \leftarrow \gamma'(z_k) - \eta \frac{\partial I}{\partial \gamma'(z_k)} \tag{5.26}$$

where  $\eta$  is the learning rate.

Figure 5.6 shows a schematic representation of this learning process. It is important to mention the fact that  $\gamma'(z_k)$  is estimated. However, if the fiber parameter  $\gamma$  is known or measured, also the absolute longitudinal power profile can be obtained. In particular, going back to Eq. (5.11), one can write

$$P(z_k) = \frac{9}{8} \frac{\gamma'(z_k)}{\gamma} \tag{5.27}$$

If  $\gamma$  is not available, only the relative power profile is estimated, but absolute  $\gamma'(z_k)$ .



**Figure 5.6:** Scheme of the learning process in the DLM algorithm, as in [15].

### 5.3.4 Non-commutativity property

In Sec. 5.3.1 the similarity between NN's and DBP structures has been highlighted and indicated as one of the main enablers of the CRM for DLM applications. However, a second element needs to be considered. This technique leverages the **non-commutativity property** of the linear and non-linear operators. In this way, it is possible to localize system physical characteristics, such as the parameters  $\gamma'(z_k)$ , and estimate them pointwise. Going a little more in-depth into the algorithm itself, the linear and non-linear operators have been indicated as  $\hat{D}$  and  $\hat{N}$ , respectively. Therefore, when applying the DBP, the chain of compensation blocks alternatively performs the integration of the linear and non-linear part of the Manakov equation. This integration has been indicated by means of an exponential term, i.e.,  $\exp(\hat{D}(z_k)\Delta z)$  and  $\exp(\hat{N}(z_k)\Delta z)$ . It is, indeed, these operators which are non-commutative with one another. In general, a signal which has gone through a system characterized by non-commutative operators can be restored only if the inverse-order system is found. According to this view, the inverse-order system is the one which minimizes the cost function of the optimization process. Non-commutativity of operators, then, introduces in the system a concept of "ordering", which is exactly what DLM tries to achieve. This is, of course, an enabler of the CRM. On the other hand, though, it is also the main limitation. If low non-linearity is present in the system, the ordering is lost and localization is not possible anymore.

## 5.4 Simulation results

In this last section a few examples of result obtained with **MATLAB** are reported.

The system considered in this section is a single-channel system. The modulation format is PM-16QAM with a symbol rate  $R_s = 64$  GBaud. The pulses are shaped according to a SRRC filter characterized by a roll-off factor  $\rho = 0.2$  and the channel power is  $P_{ch} = 10$  dBm.

The optical link is composed of 4 spans of SMF with a span length  $L_{span} = 70$  km. The main parameters of the optical fiber are  $\alpha_{dB} = 0.2$  dB/km and  $\beta_2 = -21.285$  ps<sup>2</sup>/nm. Differently from Chapter 4 Sec. 4.3, in this scenario non-linearity is the CRM enabler, so also the parameter  $\gamma = 1.3$  1/W/m is considered.

At the end of each span is an EDFA, characterized by a noise figure  $F = 5$  dB and a gain  $G$  which perfectly compensates for the total span loss.

The signal is then resampled at 2 samples/symbol and goes into the DSP. Here, fixed CD compensation and matched filtering are performed.

Afterwards, the signal gets realigned and adaptive equalization is applied with  $N_{tap} = 26$ . In particular, the tap coefficients update strategy consists of a DA-LMS based training stage and a DD-LMS based tracking stage.

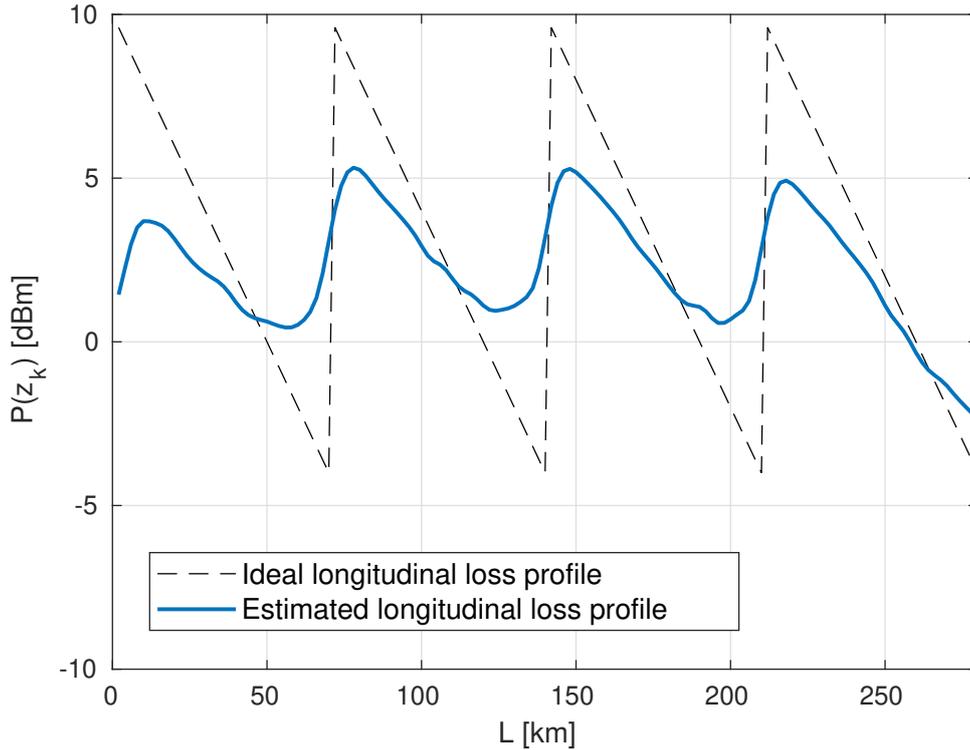
Finally, phase recovery is carried out with a BPS-ML strategy and the symbols are decoded.

Within the DSP chain, the CRM is applied right after the adaptive equalization stage. Indeed, the signal is divided into two paths. The first one is the path leading to CRM. Here, the signal goes through a CD reload stage, in which all the CD which was previously compensated for is re-inserted. Right after that, the signal goes into the data pre-processing stage, where batches and mini-batches are generated. In this simulation, the batch is composed by 60000 waveforms of  $N_{samp} = 1024$  samples, whereas the mini-batches are composed of  $N_{wf} = 100$  waveforms. Each mini-batch is finally sent as input to the channel emulator.

The second path, instead, is in charge of generating the reference (or reconstructed) transmitted signal, which will be used for the computation of the cost function during the CRM (see Eq. (5.16)) and the subsequent SGD algorithm. For simplicity, given that the originally transmitted symbol sequence is available, that has been used to reconstruct the transmitted signal. In particular, the symbol sequence has been shaped by means of a SRRC filter and then again, to perform the matched filtering. Finally, it has been re-aligned to the received signal.

As for the CRM algorithm, the DBP step  $\Delta z$  has been set to a fixed value of 2 km, which corresponds to a total of 140 DBP steps per iteration. The values for

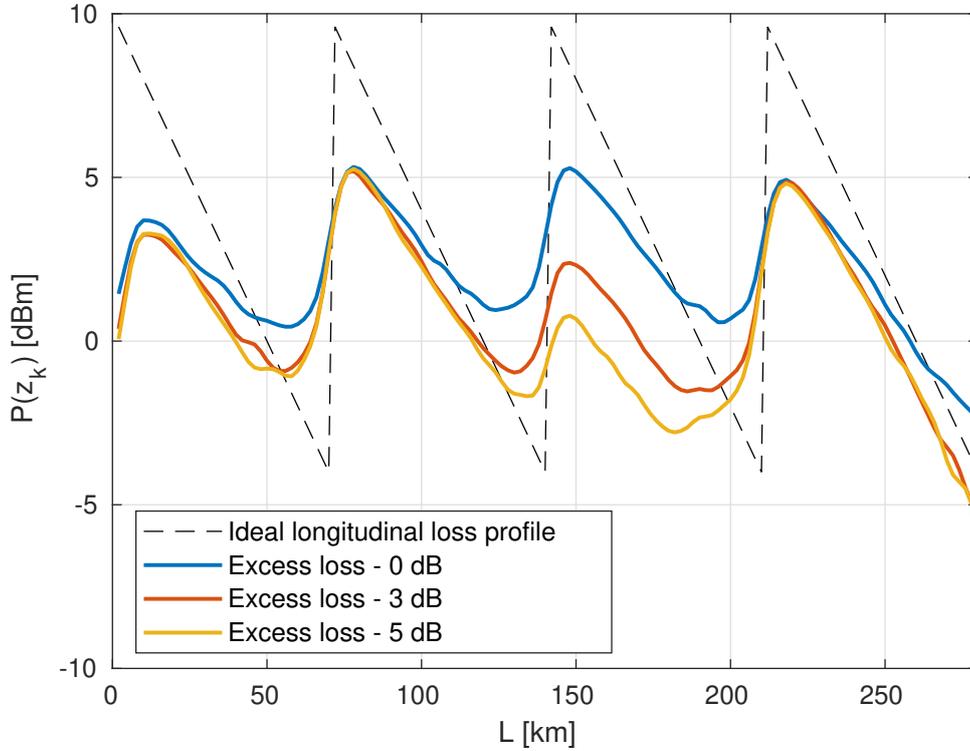
$\gamma'(z_k)$  and  $\beta_2(z_k)$ , instead, have been initialized to 0 and the nominal value of  $\beta_2$  respectively, with  $k = 1, \dots, 140$ . Figure 5.7 shows the estimated longitudinal loss profile after 600 iterations and using a learning rate  $\eta = 0.1$  for the SGD. Since the parameter  $\gamma$  is known in this scenario, then it has been possible to evaluate the absolute power loss profile by means of Eq. (5.27). In order to make comparisons, also the ideal loss profile has been reported, computed according to Eq. (5.12).



**Figure 5.7:** Estimated longitudinal loss profile (solid blue line) and the ideal longitudinal loss profile (dashed black line) after 600 iterations with step  $\Delta z = 2$  km and learning rate  $\eta = 0.1$ .

The CRM manages to detect and extract the correct behavior of the loss profile. Indeed, it is possible to distinguish between the four spans of the optical link and also the position in which EDFA amplification occurs. This is a first type of localization. Another simulation experiment consists in inserting an excess loss in some point of the optical link and check if the DLM algorithm manages to detect it and estimate its value. All the system parameters are unchanged, but an excess loss (or anomaly) is inserted at the beginning of the third span, i.e., at a distance of 140 km from the transmitter end. The loss profiles have been estimated considering

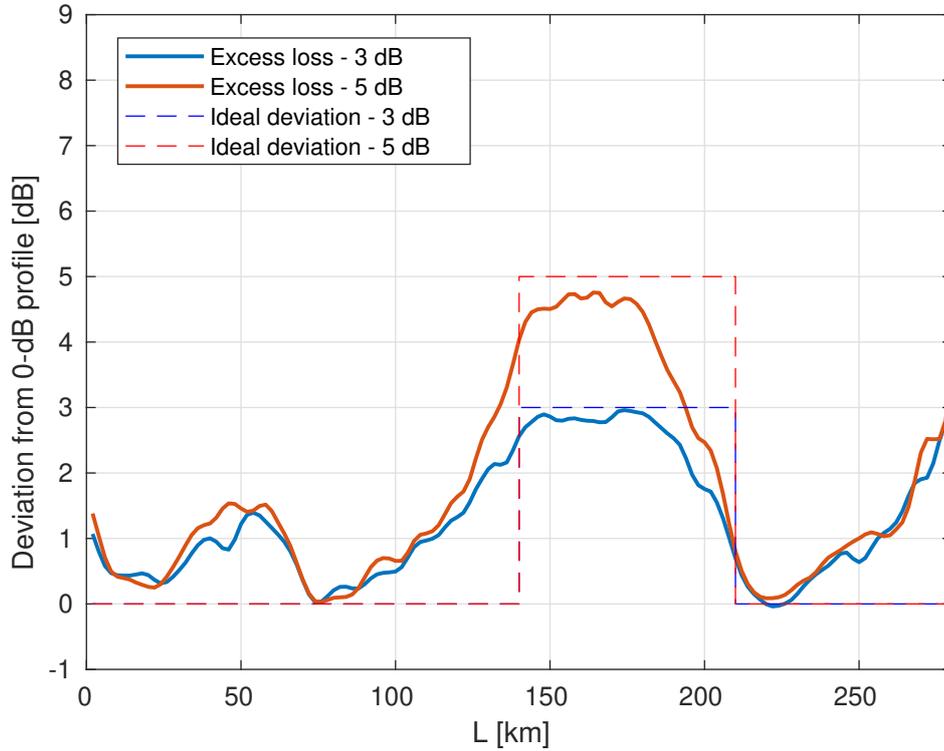
three values for the loss, namely 0 dB, 3 dB and 5 dB. The result of the estimation is reported in Figure 5.8, together with the ideal loss profile.



**Figure 5.8:** Estimated longitudinal loss profiles when an excess loss of 0 dB (solid blue line), 3 dB (solid orange line) and 5 dB (solid yellow line) is inserted at the beginning of the third span (at a distance of 140 km from the transmitter end). Also the ideal loss profile (dashed black line) is reported.

Though the estimated profiles are not always perfectly overlapping, it is possible to coarsely identify the point in which the excess loss has been inserted. By looking at the third span, the peak of the loss profile lowers by approximately 3 dB and 5 dB, which are exactly the value of the inserted loss for the two cases.

A possible method to estimate the position and the value of the excess loss is that of computing the deviation of the 3-dB and 5-dB loss profiles from the 0-dB one. Figure 5.9 shows exactly this result, along with their ideal value. In particular, the inflection points of the curves represent the position of the excess loss. Despite some fluctuations, the anomaly position can be roughly estimated around 140 km. Overall, the curves are in good agreement with the ideal ones, represented by the dashed lines.



**Figure 5.9:** Deviation of 3-dB loss profile (blue line) and 5-dB loss profile (orange line) from the 0-dB one. Ideal deviation for 3 dB case (dashed blue line) and 5 dB case (dashed red line) are also reported.

## 5.5 Conclusions

These examples conclude the section about simulation results on CRM. In its most basic form, the CRM has proved to be a valid technique to move from a cumulative to a distributed OPM approach for optical fiber communication systems. Indeed, it manages to extract some useful longitudinal quantities, such as the loss profiles and the localized excess loss, solely relying on digital techniques implemented in the DSP chain and making use of the received signal only. Of course, this represents just a starting point for these DLM applications, which have a lot of potential to still develop, as well as implementation issues to solve. All of these aspects are discussed in Chapter 6.

# Chapter 6

## Future works

In this chapter, we discuss on the main issues which came up during CRM implementations, together with some possible improvements suggested by recent scientific research works. Then, more examples of results with CRM are shown, taken from the main reference for this method, i.e., [15]. Finally, an overview of an alternative DLM method is presented, namely the correlation method (CM), as described in [17].

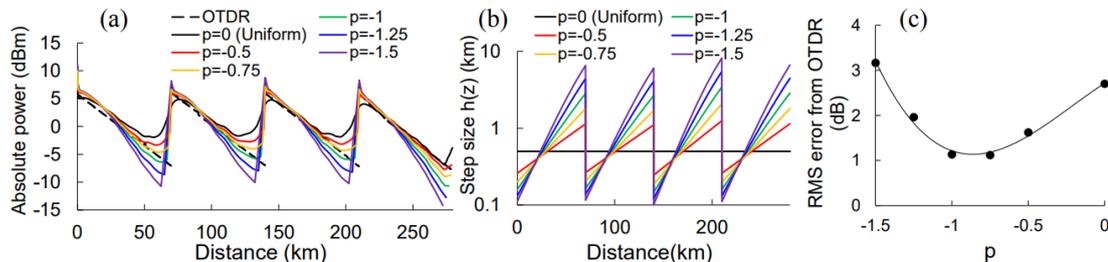
### 6.1 Issues of CRM and possible improvements

The first issue emerged during simulations of CRM is related to longer optical communication links. In Sec. 5.4 an optical link with  $4 \times 70$ -km spans has been considered and used to present the main results. However, moving to a longer system, e.g.,  $8 \times 70$ -km spans, the CRM still manages to detect the correct behavior of the loss profile, but tends to diverge more easily. Of course, this represents a major impairment for long-haul optical systems, since they are in general made of a large number of spans and reach transmission distances of up to thousands of kilometers. What is more, the estimated profiles suffer from dead-zones at the beginning and at the end of profiles themselves. This aspect can be noticed by looking at Figure 5.9, where a higher deviation from the 0-dB profile occurs at both edges of the estimated profile, though no particular loss is inserted.

As for convergence issues, an improved version of SGD is suggested in [15], namely the adaptive moment (ADAM) algorithm. It also exploits the first and second order moment of the system parameters gradients, allowing for a faster and more steady convergence of the learning process. Hence, this is one aspect to look further into.

A possible solution to the dead-zone issue, instead, is proposed in [18], from the

same authors. The idea is similar and coherent with what has been discussed for the optimization of DBP steps in Sec. 3.4.1. It consists in adjusting the DBP steps performed in the channel emulator proportionally to the power of the signal during propagation. Figure 6.1 shows an example of results taken from this article.

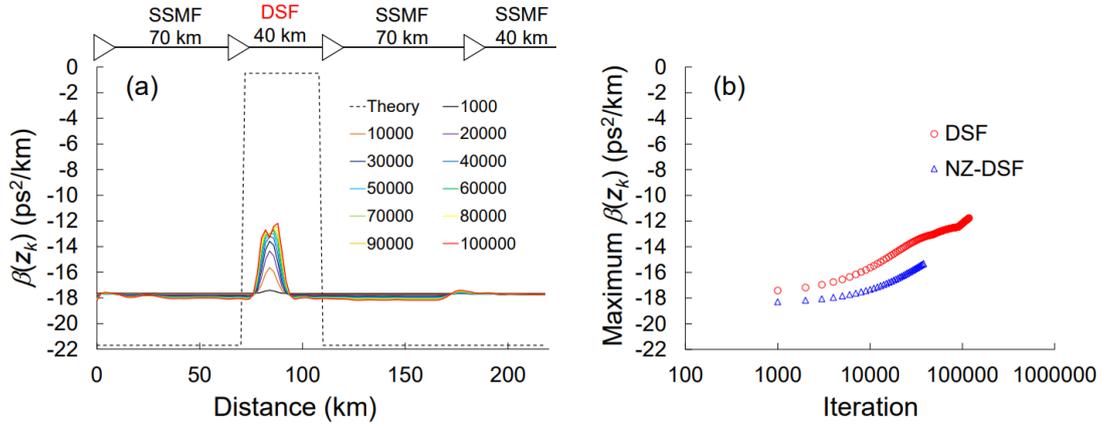


**Figure 6.1:** (a) Estimated profile over a  $4 \times 70$ -km optical link for different step sizes, according to the rule  $h(z) \propto P(z)^p$ . (b) Step size for each case. (c) Estimation error for each  $p$ . Taken from [18].

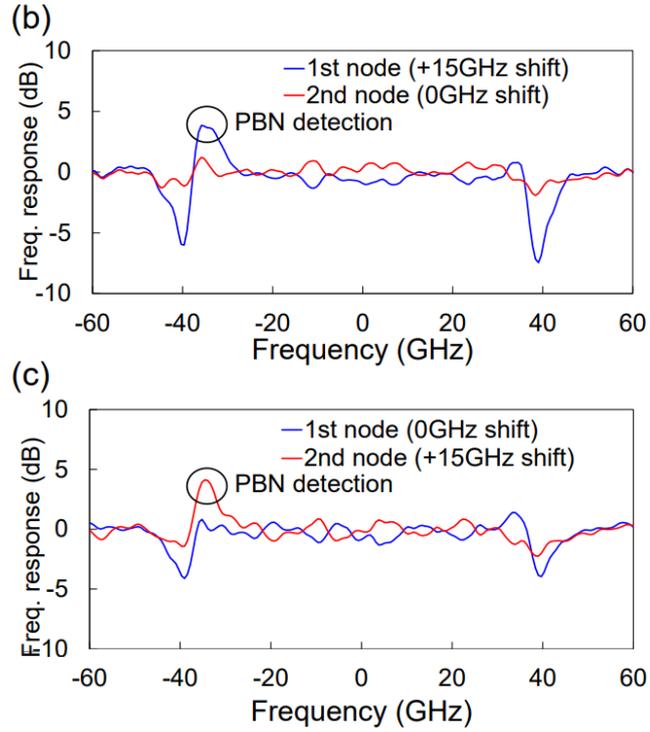
Finally, the application of the CRM method on experimental data is still to be performed.

## 6.2 CD maps and multiple filters' responses extraction with CRM

In [15] also a method to extract the CD map and multiple filters' responses over the communication link is presented. The procedure is the same as the one described in Sec. 5.3.3. For the CD map, it is necessary to add the  $\beta_2(z_k)$  parameters in the optimization process and update them computing their back-propagation expression. The filters' transfer functions, instead, are computed by modeling the filters at the end of each span as complex-valued FIR filters, and optimizing their tap coefficients during the learning process. Despite the fact that the CD map extraction method has proved to be very slow, it can still be used to detect if a span is made of a different fiber type. On the other hand, filters' transfer function extraction is more interesting. Indeed, non-linearity in the fiber allows to separate the contributions of each individual filter and check if some detuning is present. In this way, it is possible to act on just that particular filter. Figure 6.2 and Figure 6.3 show some results.



**Figure 6.2:** (a) Estimated  $\beta_2(z_k)$  profile for a different number of SGD iterations when a DSF span is inserted between SSMF spans. (b) Maximum value reached by  $\beta(z_k)$  versus number of iterations. Taken from [15]



**Figure 6.3:** Estimated FIR filters responses in a  $3 \times 50$ -km system with an optical bandpass filter at each node when a detuning of +15 GHz is inserted (b) in the first filter and (c) in the second filter. Taken from [15]

### 6.3 An alternative approach: the correlation method

Finally, an alternative approach is described in [17]. This DLM algorithm belongs to the correlation method (CM) approach and is referred to as "in-situ power profile estimator" or "in-situ PPE". The scheme of the algorithm is reported in Figure 6.4.

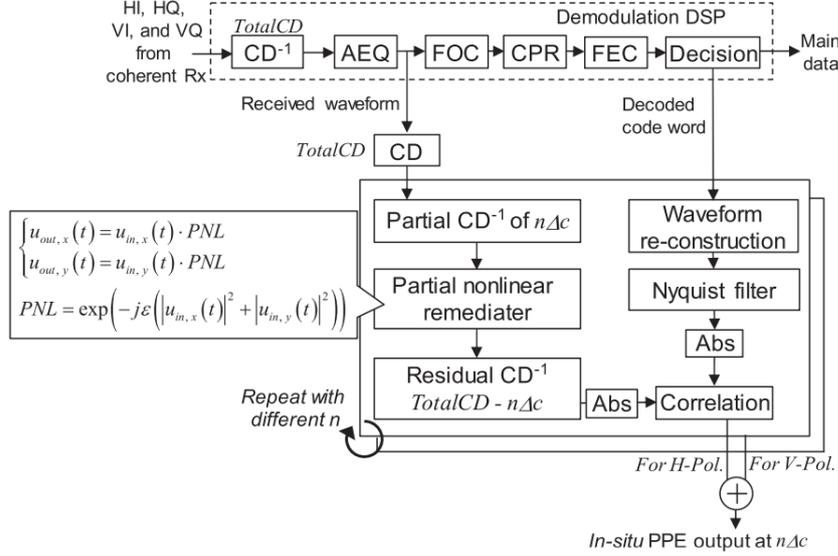


Figure 6.4: In-situ PPE algorithm scheme, as in [17].

This type of approach to extract the longitudinal fiber loss profile is not very different from the CRM method, which is the subject of this work. Indeed, at each step the algorithm performs a partial DBP and then evaluates the power by means of a correlation operation. In any case, it would be extremely interesting to try to implement also this DLM approach and eventually compare the performance of the two.

### 6.4 Conclusions

To sum up, OPM is a field full of potential. This work has presented an advance in terms of what can be estimated, moving from a cumulative to a distributed OPM and has proved to be successful in its very basic implementation and application. For this reason, it would be of great interest to keep developing and improving these techniques, in order to contribute to the progress of optical fiber communication systems and telecommunications more in general.

# Bibliography

- [1] J. M and M.Y. Jamro. *Optical Fiber Communications: Principles and Practice*. Prentice Hall Internacional series in optoelectronics. Financial Times/Prentice Hall, 2009. ISBN: 9780130326812. URL: <https://books.google.it/books?id=7HtqsEOMkXUC> (cit. on pp. 3–8, 11, 12, 14–16, 19, 22, 23, 33).
- [2] Y. Sun, J.L. Zyskind, and A.K. Srivastava. «Average inversion level, modeling, and physics of erbium-doped fiber amplifiers». In: *IEEE Journal of Selected Topics in Quantum Electronics* 3.4 (1997), pp. 991–1007. DOI: 10.1109/2944.649527 (cit. on p. 13).
- [3] Tingye Li. «Advances in Optical Fiber Communications: An Historical Perspective». In: *IEEE Journal on Selected Areas in Communications* 1.3 (1983), pp. 356–372. DOI: 10.1109/JSAC.1983.1145944 (cit. on p. 16).
- [4] Peter J. Winzer, David T. Neilson, and Andrew R. Chraplyvy. «Fiber-optic transmission and networking: the previous 20 and the next 20 years [Invited]». In: *Opt. Express* 26.18 (Sept. 2018), pp. 24190–24239. DOI: 10.1364/OE.26.024190. URL: <http://opg.optica.org/oe/abstract.cfm?URI=oe-26-18-24190> (cit. on pp. 17, 26).
- [5] R. Clark Jones. «A New Calculus for the Treatment of Optical SystemsI. Description and Discussion of the Calculus». In: *J. Opt. Soc. Am.* 31.7 (July 1941), pp. 488–493. DOI: 10.1364/JOSA.31.000488. URL: <http://opg.optica.org/abstract.cfm?URI=josa-31-7-488> (cit. on p. 25).
- [6] R. S. Kerdock and D. H. Wolaver. «Atlanta fiber system experiment: results of the atlanta experiment». In: *The Bell System Technical Journal* 57.6 (1978), pp. 1857–1879. DOI: 10.1002/j.1538-7305.1978.tb02131.x (cit. on p. 26).
- [7] Kazuro Kikuchi. «Fundamentals of Coherent Optical Fiber Communications». In: *Journal of Lightwave Technology* 34.1 (2016), pp. 157–179. DOI: 10.1109/JLT.2015.2463719 (cit. on pp. 29, 32, 34–36, 39, 41).

- 
- [8] Seb J. Savory. «Digital filters for coherent optical receivers». In: *Opt. Express* 16.2 (Jan. 2008), pp. 804–817. DOI: 10.1364/OE.16.000804. URL: <http://opg.optica.org/oe/abstract.cfm?URI=oe-16-2-804> (cit. on pp. 29, 37, 39, 40).
- [9] Alex Alvarado, Tobias Fehenberger, Bin Chen, and Frans M. J. Willems. «Achievable Information Rates for Fiber Optics: Applications and Computations». In: *Journal of Lightwave Technology* 36.2 (2018), pp. 424–439. DOI: 10.1109/JLT.2017.2786351 (cit. on p. 38).
- [10] Ezra Ip and Joseph M. Kahn. «Compensation of Dispersion and Nonlinear Impairments Using Digital Backpropagation». In: *Journal of Lightwave Technology* 26.20 (2008), pp. 3416–3425. DOI: 10.1109/JLT.2008.927791 (cit. on pp. 42, 44).
- [11] Junwen Zhang, Xinying Li, and Ze Dong. «Digital Nonlinear Compensation Based on the Modified Logarithmic Step Size». In: *Journal of Lightwave Technology* 31.22 (2013), pp. 3546–3555. DOI: 10.1109/JLT.2013.2285648 (cit. on p. 44).
- [12] Dario Pileri, Antonino Nespola, Fabrizio Forghieri, and Gabriella Bosco. «Non-Linear Phase Noise Mitigation Over Systems Using Constellation Shaping». In: *Journal of Lightwave Technology* 37.14 (2019), pp. 3475–3482. DOI: 10.1109/JLT.2019.2917308 (cit. on pp. 47, 48).
- [13] Fabian. N. Hauske, Maxim Kuschnerov, Bernhard Spinnler, and Berthold Lankl. «Optical Performance Monitoring in Digital Coherent Receivers». In: *Journal of Lightwave Technology* 27.16 (2009), pp. 3623–3631. DOI: 10.1109/JLT.2009.2024960 (cit. on p. 52).
- [14] G. Bosco, R. Cigliutti, E. Torrenco, A. Carena, V. Curri, P. Poggiolini, and F. Forghieri. «Joint DGD, PDL and chromatic dispersion estimation in ultra-long-haul WDM transmission experiments with coherent receivers». In: *36th European Conference and Exhibition on Optical Communication*. 2010, pp. 1–3. DOI: 10.1109/ECOC.2010.5621379 (cit. on p. 52).
- [15] Takeo Sasai, Masanori Nakamura, Etsushi Yamazaki, Shuto Yamamoto, Hideki Nishizawa, and Yoshiaki Kisaka. «Digital Longitudinal Monitoring of Optical Fiber Communication Link». In: *Journal of Lightwave Technology* 40.8 (2022), pp. 2390–2408. DOI: 10.1109/JLT.2021.3139167 (cit. on pp. 71, 78, 80, 83, 88–90).
- [16] Takeo Sasai, Masanori Nakamura, Etsushi Yamazaki, Shuto Yamamoto, Hideki Nishizawa, and Yoshiaki Kisaka. *Physics-oriented learning of nonlinear Schrödinger equation: optical fiber loss and dispersion profile identification*. 2021. DOI: 10.48550/ARXIV.2104.05890. URL: <https://arxiv.org/abs/2104.05890> (cit. on pp. 71, 76).

- [17] Takahito Tanimura, Setsuo Yoshida, Kazuyuki Tajima, Shoichiro Oda, and Takeshi Hoshida. «Fiber-Longitudinal Anomaly Position Identification Over Multi-Span Transmission Link Out of Receiver-end Signals». In: *Journal of Lightwave Technology* 38.9 (2020), pp. 2726–2733. DOI: 10.1109/JLT.2020.2984270 (cit. on pp. 88, 91).
- [18] Takeo Sasai, Masanori Nakamura, Etsushi Yamazaki, and Yoshiaki Kisaka. «Precise Longitudinal Power Monitoring over 2,080 km Enabled by Step Size Selection of Split Step Fourier Method». In: *2022 Optical Fiber Communications Conference and Exhibition (OFC)*. 2022, pp. 1–3 (cit. on pp. 88, 89).