Réduction en bande de base des distorsions non-linéaires dans les systèmes OFDM optiques cohérents basés sur les SOAs : analyse stochastiques et expérimentales

Jacqueline Sime

DOCTORAT BRETAGNE LOIRE / MATHSTIC

ÉCOLE NATIONALE D'INGÉNIEURS DE BREST

THÈSE DE DOCTORAT DE

L'ÉCOLE NATIONALE
D'INGÉNIEURS DE BREST

ÉCOLE DOCTORALE N° 601
Mathématiques et Sciences et Technologies de l'Information et de la Communication
Spécialité : *Télécommunications*

Par

Jacqueline Emilienne SIME

Baseband mitigation of nonlinear impairments in SOA based coherent optical OFDM systems: stochastic and experimental analyses

Thèse présentée et soutenue à Brest, le 30/11/2021
Unité de recherche : Lab-STICC, CNRS UMR 6285
Thèse N° : 11

Rapporteurs avant soutenance :

Christelle AUPETIT-BERTHELEMOT	Professeur, Université de Limoges
Yves JAOUEN	Professeur, Télécom ParisTech

Composition du Jury :

Président :	Pascal BESNARD	Professeur, ENSSAT, Université de Rennes 1
Rapporteurs :	Christelle AUPETIT-BERTHELEMOT	Professeur, Université de Limoges
	Yves JAOUEN	Professeur, Télécom ParisTech
Examinateurs :	Pascal BESNARD	Professeur, ENSSAT, Université de Rennes 1
	Laurent ROS	Professeur, INP Grenoble
Dir. de thèse :	Stéphane AZOU	Professeur, Ecole Nationale d'Ingénieurs de Brest
Co-dir. de thèse :	Noël TANGUY	Professeur, Université de Bretagne Occidentale, Brest
Co-Encadrants :	Pascal MOREL	Maître de Conférences, Ecole Nationale d'Ingénieurs de Brest
	Mihai TELESCU	Maître de Conférences, Université de Bretagne Occidentale, Brest

Acknowledgement

Throughout this thesis journey, I have received a great deal of support and assistance. Therefore, I would like to take the time to thank the people who have helped me obtain this doctorate degree.

First, I would like to thank the Ecole Nationale d'Ingénieurs de Brest (ENIB) and the Université de Bretagne Occidentale (UBO) for co-financing this thesis and the Institut Brestois du Numérique et des Mathématiques (IBNM) for financing the GEMP project. I would also like to thank my dream team: my co-directors, Prof. Stéphane AZOU and Prof. Noël TANGUY, and my co-supervisors Dr. Pascal MOREL and Dr. Mihai TELESCU. There has never been a dull moment in the past 3 years that we spent together and I could not have asked for a better experience. Thank you for your continuous help and support whether it be through our conversations about the research matter, preparing presentations, the writing of numerous papers, and general life and professional advice. Thank you for the moments we got to celebrate together and for all the laughs. I hope that you are proud of the work we have done.

I would like to thank the professors who accepted to evaluate this thesis work. Thank you to Prof. Christelle AUPETIT-BERTHELEMOT and Prof. Yves JAOUEN who have reviewed this manuscript meticulously and whose comments and discussions have allowed me to think more deeply about my research as well as think outside of the box. Thank you to Prof. Pascal BESNARD, president of the jury, and Prof. Laurent ROS who have examined this work thoroughly and allowed me to go beyond myself especially during my defense. Thank you all for a challenging yet rather enjoyable defense.

I would like to thank my parents: Prof. Jean-Marie SIME and Dr. Hortense SIME. You have always believed in me. I am able to write these acknowledgments today thanks to your advice, your encouragement, and your love. I have enjoyed connecting with you through this thesis experience that you have lived before me. I hope that one day I will be able to achieve as much as you two did with the tools you have given me.

Of course, I cannot thank my parents without thanking my siblings. Thank you to Andrée Succès, Rufin Alphonse, Paul-Dékara J.D., Marthe-Eugénie, and Claudine-Elisabeth. You have helped me keep my sanity during this thesis and seeing you all

pursue your dreams and goals has inspired me. Thank you for giving me your input on every aspect of my work and for making every day a joy to live.

Thank you to my aunt, Eugénie DJIMI who has been a part of this PhD process from the moment I started writing my letter of motivation. You have always checked on me, made sure I did my best, and inspired me through your resilience. Thank you for helping me see my objectives through.

I want to thank my grand-parents Dr. Dieudonné Dékara DJIMI, Jacqueline DJIMI, princess Emilienne MINGUE, and André NGUEHOU for watching over me, those that are here and those watching from above. Grandpa, I have appreciated your advice coming from your experience as a PhD student and an educator, and grandma I thank you for your unfailing support.

I would also like to thank both sides of my family SIME and DJIMI for being there for me, especially my uncle Alphonse DJIMI for giving me his time and advice to prepare my transition from the thesis into the professional world.

I want to thank the people who ensured this thesis ran smoothly and contributed in the scientific work that came from it: Prof. Emanuel RADOI, Dr. Joël CHARRIER, and Prof. Igor Simone STIEVANO. I also want to add those who gave me advice when doing experimental work: Dr. Thierry RAMPONE, my close friends Dr. Noor Hamdash and Dr. Dimitrios KASTRITSIS, Dr. Laura TELESCU (who helped me in general), and Prof. Ammar SHARAIHA.

I would like to thank the people who gave me guidance and helped me in my teaching journey: Céline ANSQUER who believed in my professional project to teach and gave me the tools to be a successful educator, Dr. Vincent CHOQUEUSE (who also guided me in my scientific and professional journey), Eric BOUCHARE, Dr. Kamal NASREDDINE, Dr. Abdesslam BENZINOU, and Dr. Véronique QUINTARD, and of course the students I had the chance to teach at ENIB.

During this thesis, I had not only the chance to discover Brest and Brittany but also to build meaningful friendships inside and outside the ENIB and UBO laboratories. For their support, their help during the difficult task of writing this manuscript, their scientific advice, and for all the laughs and unforgettable times we've had together, I would like to thank ShalliRosena MBWETTE, Dr. Marwa BRAIKI and Dr. Hamza BENYAHIA with little Aous, Odette Sandrine BAKAM NGUENOUHO, Tanguy LOUIS-MARIE and Marion ROULET, Dr. Olivier and Azumi AUGEREAU, Dr. Abdelouahid BEN TAMOU, Igor GLUKHOV, Alexandru FRUNZA and Abraham SOTOMAYOR FERNANDEZ who

I had the pleasure to talk to about our mutual research topics, Raed AL HASSANIEH, Olivier BALD, Maeva FRANCO, Dr. Ramez HAMIE, Dr. Yuliya DADOENKOVA, Morann MATTINA, Dr. Adan Omar ARELLANES BERNABE and Sayuri, Romildo DE SOUZA and Lis, Khawla BEN GAIED, Paulo KIOHARA, Prof. André PERENNOU, Jean-Luc (Speedy) BERTHEVAS, Dr. Jean-François FAVENNEC, Dr. Mikael GUEGAN, Dr. Florian BENTIVEGNA, Dr. Jacques CHI, Dr. Pierre-Marie MARTIN, Dr. Marc LE ROY, Dr. Mohamad YOUNES, and Lucie LE PAPE and her daughters Youna and Mannaïg who welcomed me to Brest.

I want to thank the people who made it possible for me to start this thesis by accepting to attest to my ability to do research: Prof. Farzan AMINIAN, Dr. Ronald PHLYPO, and Dr. Joshua Schwartz.

Lastly, although I tried my best to thank the people who have helped me throughout, I want to dedicate these last few lines to the people who I forgot to include here. Know that I am thankful for everything that was done for me.

Table of Contents

List of Figures

List of Tables

Acronyms

ACO-OFDM	asymmetrically clipped OFDM
ACPR	adjacent channel power ratio
ADC	analog to digital conversion
ADS	Advanced Design System
AM/AM	amplitude modulation to amplitude modulation
AM/PM	amplitude modulation to phase modulation
AMOOFDM	adaptively modulated OFDM
ANOVA	analysis of variance
AO-OFDM	all optical OFDM
ASE	amplified spontaneous emission
AWG	arbitrary waveform generator
AWGN	additive white Gaussian noise
B2B	back to back
BER	bit error rate
BPSK	binary phase shift keying
CCDF	complementary cumulative density function
CD	chromatic dispersion
CE-OFDM	constant envelope OFDM
CFO	carrier frequency offset
CLEO	*conference on lasers and electro-optics*
CMA	constant modulus algorithm
CO-OFDM	coherent optical orthogonal frequency division multiplexing
CP	cyclic prefix
CP-OFDM	cyclic prefix orthogonal frequency division multiplexing
CW	continuous wave
DAC	digital to analog conversion
DCO-OFDM	DC-biased OFDM

DMT discrete multitone
DPD digital predistortion
DPLL digital phase-locked loop
DPoD digital post-distortion
DPSK differential phase shift keying
DSP digital signal processing

EDFA Erbium doped fiber amplifier
EMP envelope memory polynomials
ENOB effective number of bits
EPON ethernet PON
ET envelope tracking
EVM error vector magnitude

FBMC filter bank multicarrier
FEC forward error correction
FFT fast Fourier transform
FIR finite impulse response
FPGA field programmable gate arrays
FTTx fiber to the x
FWM four wave mixing

GC-SOA gain clamped SOA
GEMP generalized envelope memory polynomials
GMP generalized memory polynomials
GSA global sensitivity analysis
GSOP Gram-Schmidt orthogonalization procedure
GVD group-velocity dispersion

ICCE *international conference on communications and electronics*
ICI intercarrier interference
IEEE institute of electrical and electronics engineers
iFFT inverse fast Fourier transform
IM/DD intensity-modulation/direct-detection
IPC *IEEE photonics conference*
ITU international telecommunication union

JNOG	*journée nationale d'optique guidée*
KSP	known symbol padding
Lab-STICC	Laboratoire des Sciences et Techniques de l'Information de la Communication et de la Connaissance
LAN	local area network
LARS	sparse least angle regression
LHS	latin hypercube sampling
LS	least squares
LUT	look-up table
MAN	metropolitan area network
MC	Monte Carlo
MCM	multicarrier modulation
MP	memory polynomials
MZI	Mach-Zehnder interferometer
MZM	Mach-Zehnder modulator
NF	noise figure
NMSE	normalized mean square error
NN	neural networks
OCSA	optical complex spectrum analyzer
OEO	optical-electrical-optical
OFDM	orthogonal frequency division multiplexing
OFDMA	orthogonal frequency division multiple access
OMA	optical modulation analyzer
OOK	on-off keying
OQAM	offset-QAM
OSNR	optical signal to noise ratio
PAM	pulse amplitude modulation
PAPR	peak to average power ratio
PCE	polynomial chaos expansion
PDF	probability density function
PDM	polarization division multiplexing

PMD	polarization mode dispersion
PON	passive optical network
PPM	pulse position modulation
PSD	power spectral density
PSK	phase shift keying
PTL	*photonics technology letters*
QAM	quadrature amplitude modulation
QoI	quantity of interest
RF	radio frequency
RFS	recirculating frequency shift
RSOA	reflective SOA
SCM	single carrier modulation
SEE-OFDM	spectral and energy efficient OFDM
SGM	self-gain modulation
SMF	single-mode fiber
SNR	signal to noise ratio
SOA	semiconductor optical amplifier
SPM	self-phase modulation
SPR	self-polarization rotation
SRS	stimulated Raman scattering
TDTMM	time-domain transfer matrix model
TIA	transimpedance amplifiers
TW	traveling wave
VISA	virtual instrument systems architecture
VSA	vector signal analysis
WAN	wide area network
WDM	wavelength division multiplexing
WSS	wavelength selective switches
xDSL	digital subscriber line
XGM	cross-gain modulation
XPM	cross-phase modulation

XPolM cross-polarization modulation

ZP zero padding

List of Symbols

$\boldsymbol{\beta}^{(j)}$	data aided demodulation coefficients
$\Delta\nu$	laser linewidth
γ	clipping ratio
σ	standard deviation
$\dagger$	Moore-Penrose pseudo inverse
ΔB	signal bandwidth when excluding the oversampling
ω_{LO}	laser source angular frequency
σ_{PCE}	percent uncertainty
ρ_{cp}	cyclic prefix factor
α_{dB}	fiber attenuation
ζ_{est}	estimated complex gain
ρ_{os}	oversampling factor
Δf	spacing between adjacent frequencies in multicarrier modulation
$(.)^*$	conjugate operator
A	hard-clipping clipping level
I_{pin}	generated photodiode current
N_ϵ	total number of erroneous received bits
N_{bits}	total number of bits sent
N_{sc}	number of subcarriers in an OFDM symbol
P_{ref}	reference input power into the SOA
P_s	laser source power
R_b	data rate
T_s	sample period
T_{symb}	symbol period
X	random input vector
Y	random output variable
f_k	frequency of the k^{th} subcarrier
i	imaginary number

r_k	received symbol
$s(t)$	sent time signal
$s[n]$	sent discrete signal
s_k	sent symbol
$x[n]$	predistorder input
$y[n]$	received signal for predistorter estimation

General introduction

In this day and age, data usage is rapidly extending to the masses causing an exponential increase in the global demand for internet bandwidth. In that regard, innovations in optical fiber technology have been revolutionizing world communications by exploiting the large optical bandwidth of optic fibers. The development and use of optical amplifiers is a quintessential element in responding to this need for higher bandwidth. A number of recent studies highlight that semiconductor optical amplifiers (SOAs) could become a relevant alternative to Erbium doped fiber amplifiers (EDFAs) for various application scenarios due to their attractive characteristics such as high optical bandwidth, small size, and the possibility of integration at limited cost. In particular, the practical feasibility of an SOA based access/metropolitan infrastructure has been demonstrated using advanced modulation formats such as quadrature amplitude modulation (QAM) or phase shift keying (PSK). These studies consider not only the conventional SOA only scenario, but also the cascaded configuration that allows for high capacity links in the Tbits/s range. Single carrier modulation formats are mainly considered in the literature dealing with SOA based amplification in optical networks. However, in the last decade, orthogonal frequency division multiplexing (OFDM) and its variants have been recognized as promising techniques to support high data rate in next generation optical communications, due to some important advantages such as simple compensation of channel dispersion, dynamic bandwidth allocation capability in a multi-user context, and powerful digital signal processing (DSP) based implementation. An issue arises with the non-constant envelope property OFDM signals as it makes them sensitive to nonlinear channel impairments. Due to their fast gain dynamics and intrinsic nonlinear properties, SOAs can therefore impact the system performance for such signals.

Various memory polynomial digital predistortion (DPD) schemes have recently been studied at the Laboratoire des Sciences et Techniques de l'Information de la Communication et de la Connaissance (Lab-STICC) to linearize a coherent optical orthogonal frequency division multiplexing (CO-OFDM) transmitter using an SOA as a power amplifier. Extensive simulations based on an accurate physical model of the SOA, and initial experiments have shown that it is possible to obtain a significant performance gain. How-

ever, these results were established considering a relatively simple scenario (single SOA, limited electrical bandwidth) and with a number of simplifying assumptions. Further research effort is needed for high speed and efficient SOA based CO-OFDM transmissions. This thesis work will investigate the generalized memory polynomials (GMP) predistorter and compare it to the memory polynomials (MP) and the envelope memory polynomials (EMP) which are other well-known predistorter structures. The analysis of the effectiveness of the GMP predistorter for different scenarios will be an added value. This analysis will consist of evaluating the benefits of DPD in the presence of fiber dispersion, emitter and digital to analog conversion (DAC) imperfections, for differing modulation formats, and its ability to sustain variations in the system. An experimental validation will also be the object of this thesis where the effects of laser phase noise will be studied. The conclusions drawn from this thesis work will be essential in predicting the feasibility of a transmission system using SOAs and non-constant envelope signals, the benefits to reap from that system and addressing the limitations that could come along with it.

This thesis manuscript will delve into some of the challenges encountered in SOA based CO-OFDM systems for access/metropolitan networks and will give some insight to help understand and solve these problems. The manuscript will be divided into four chapters.

Chapter 1 is an introductory chapter where some of the advances in optical communications systems and deployment are presented. The architecture of a typical optical transmission system is given from the emitter to the receiver. An emphasis is put on OFDM modulation formats and SOAs pointing out the problematic in combining them despite their undeniably significant benefits. Some techniques are laid out to respond to the issue of an OFDM/SOA transmission with DPD standing out as a viable and effective solution. The usual methods of evaluating the optical transmission system in general as well as the chosen linearization techniques are introduced. Then, the CO-OFDM system that will be studied throughout this manuscript is explained, most notably the SOA model that is used as the foundation for this thesis work.

Chapter 2 constitutes the first step into the analysis of DPD to linearize an SOA based CO-OFDM system. Chapter 2 deals with numerical simulations specifically. First, the error vector magnitude (EVM) is presented as a valid alternative to the well established bit error rate (BER) since the EVM will be one of the cornerstones of this thesis in terms of system evaluation. Second, the choice of peak to average power ratio (PAPR) to limit the high OFDM peaks is discussed. After discussing DPD implementation including the use of an algorithm to automatically find the structural parameters of a predistorter.

Then, two scenarios are explored among which one with the SOA as booster without fiber and another with the SOA in-line.

Chapter 3 builds on Chapter 2 by doing a stochastic analysis of DPD in an SOA based CO-OFDM system. A substitute to Monte Carlo (MC) simulations is proposed to perform the stochastic analysis; that alternate scheme called polynomial chaos expansion (PCE) is molded to fit a CO-OFDM system and its validity is put to the test. The sensitivity analysis is a way to study the ability of DPD to remain efficient in a real environment that is affected by component aging, manufacturing uncertainties or fluctuating environment (change in temperature, vibrations). Two scenarios are investigated: a transmission with a booster SOA without fiber and a transmission with a booster SOA with fiber added.

Chapter 4 ties in with Chapters 2 and 3 by analyzing DPD in an SOA based CO-OFDM system through an experimental setup where some of the simplifying assumptions made during numerical simulations do not necessarily translate well. First, the experimental transmission chain and testbed are described. The characteristics and limitations of specific components are then detailed. Finally, experimental results are shown for a booster SOA without fiber scenario.

At the end of this manuscript, a conclusion covering the main points and discoveries of this thesis work will be given as well as some perspectives for future research.

Chapter 1

Optical communication systems

Nowadays, optical communications are a staple in Telecommunication systems. Fiber optics went through an experimental phase in the '60s and the '70s with the first lasers seeing the light of day in the '70s. This allowed fiber optics to be put on the market in the '80s and the '90s with the knowledge that they would be the only medium that could sustain the growing requirements for higher data rates [Agr+16]. Rightfully so, contrary to its counterparts (copper/coaxial cables, wireless radio frequency (RF)), optic fibers have the ability for lower attenuation (meaning longer reach), higher data rates, and no electromagnetic interference. This makes optic fibers an ideal candidate to respond to the increasing demands of clients especially for long distance communications.

1.1 Optical communication networks

The optical communication networks can be divided into three main categories: wide area networks (WANs) or core networks, metropolitan area networks (MANs) or metropolitan networks, and local area networks (LANs) or access networks (Figure 1.1). These three groups are characterized by their distances and their cost limits with the core networks having the largest distance coverage and therefore allowing for higher costs, and the access networks having the shortest coverage thus being more restrictive in terms of costs. The scope of this thesis lies within the field of metropolitan and access networks.

Access networks are used in the context of communication inside of one building or home with distances up to 40 km [ITU18]. In France, the main network operators that provide directly to subscribers are Orange, SFR, Free, and Bouygues. To increase data rates and reduce attenuation, access networks are in the process of shifting from digital subscriber line (xDSL) to fiber to the x (FTTx) technologies. This would illustrate a relative shift from 4 Mb/s (fast ADSL) to 10 Gb/s (fiber per channel) [ITU18]. A prime way of building FTTx networks is through passive optical network (PON) systems. The standards for PONs are developed by two major groups: institute of electrical and

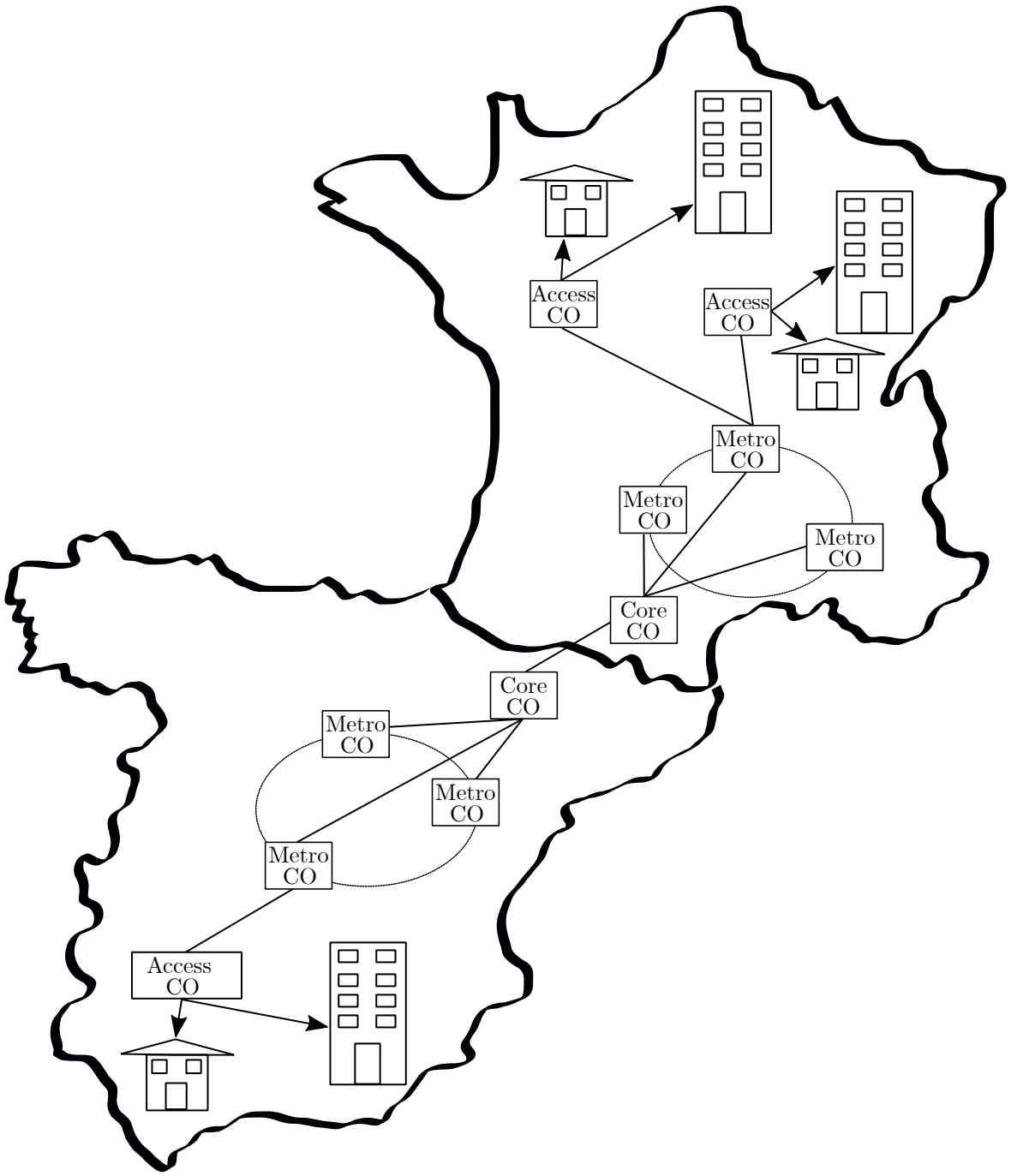


Figure 1.1 – Optical topologies. CO: Central Office

electronics engineers (IEEE) and international telecommunication union (ITU). IEEE has introduced the gigabit-capable ethernet PON (EPON), 10G-EPON, and 25/50-EPON standards. ITU has brought on the gigabit-capable GPON, XGS-PON, and NG-PON2.

All standards (shown in Table 1.1) strive to meet the rapidly increasing demands on bandwidth at a reasonable price.

Table 1.1 – PON standards overview [VIA21]

	Standard	Down/upstream data rates	Splitting ratio
G-PON	ITU-T G.984	2.4 / 1.2 Gb/s	Up to 1:64 (128)
XGS-PON	ITU-T G.9807.1	10 / 10 Gb/s	Up to 1:128 (256)
NG-PON2	ITU-T G.989	40 / 10 Gb/s	Up to 1:128 (256)
G-EPON	IEEE 802.3ah	1.25 / 1.25 Gb/s	Up to 1:64
10G-EPON	IEEE 802.3av	10 / 10 Gb/s	Up to 1:128
100G-EPON	IEEE 802.3ca	Up to 100 / 100 Gb/s	TBD

Metropolitan networks ensure a connection between different regions. For instance, the RENATER network links multiple public research and education entities within France. The distances for these networks can go up to several hundreds of km. The fiber technologies used in that area are moving towards 10G Ethernet or 40G Ethernet with a goal of 100 Gb/s data rates [ITU20a].

Core networks go one step further than metropolitan networks as they take care of connections between countries and trans-oceanic transmissions. Some examples are the Apollo and the ACE cable systems that connect Europe to the United States and Europe to Western Africa respectively. Both systems have a station in the French region of Brittany. The distances can reach several thousands of kilometers with data rates higher than 100 Gb/s.

1.2 Increasing demand for capacity

The previously introduced optical network architectures can differ in some respects; however, they have one main challenge to overcome: a demand for higher data rates growing exponentially (Figure 1.2). Some ways that have been found to meet this growing demand in photonics is via wavelength division multiplexing (WDM) and polarization division multiplexing (PDM). WDM takes advantage of fiber optic's large optical bandwidth by sending information onto a set of different wavelengths (Figure 1.3a). On the other hand, PDM (shown in Figure 1.3b) sends information onto one wavelength but does it in two orthogonal polarization states. Combining WDM and PDM entails that the total data rate can be increased considerably as two polarization states doubles the amount of data sent as does using two different wavelengths.

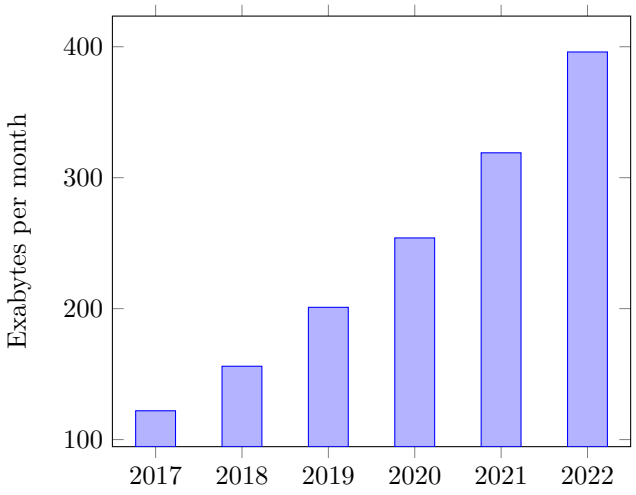


Figure 1.2 – Global traffic demand forecast [Cis20]

For the purpose of this thesis, the focus will lie on a transmission system with one polarization and one wavelength, but it is still important to be able to project oneself onto the scenario of WDM/PDM for later research.

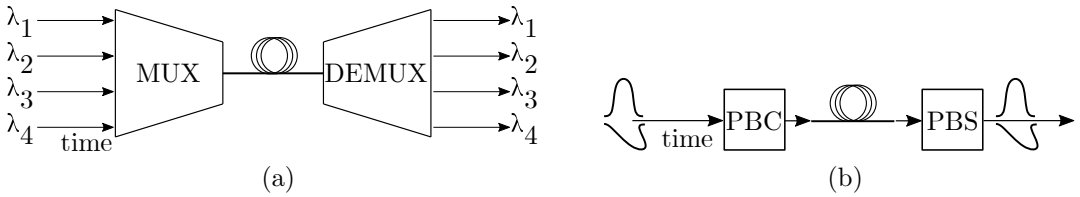


Figure 1.3 – (a) WDM for an ideal transmission. MUX: multiplexer, DEMUX: demultiplexer (b) PDM for an ideal transmission. PBC: polarization beam combining, PBS: polarization beam splitting

1.3 Architecture of optical transmission systems

The next sections will give a closer look into an optical communication architecture while expanding on ways other than WDM and PDM to increase data rate. These sections describe a conventional architecture of a transmission chain implementing the use of fiber optics as depicted in Figure 1.4.

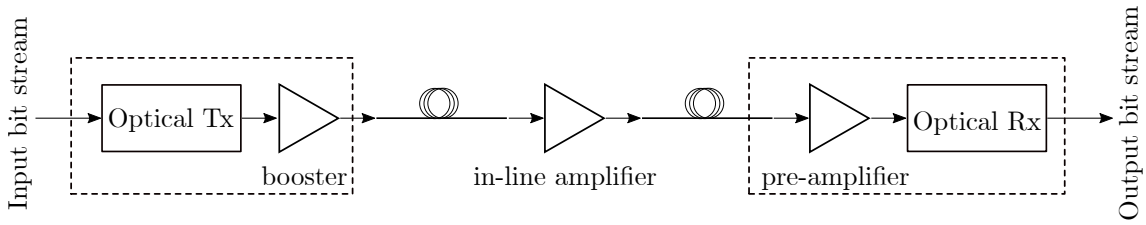


Figure 1.4 – Schematic of a general optic communication. Tx: transmitter, Rx: receiver

1.3.1 Emitter in photonics

The optical emitter (optical Tx) can be represented by Figure 1.5 with the signal undergoing digital signal processing (DSP) then digital to analog conversion (DAC) in the electrical domain to then be converted into the optical domain.

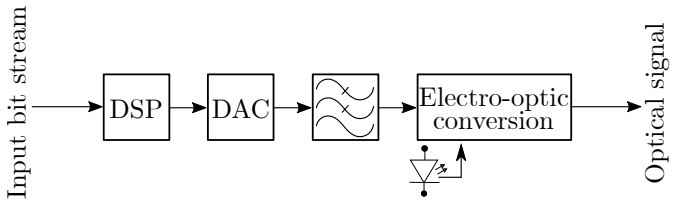


Figure 1.5 – Schematic of an optical emitter. DSP: digital signal processing, DAC: digital to analog conversion

In the digital domain, it is possible to modulate the input bit stream in two ways as depicted in Figure 1.6: intensity modulation and multidimensional modulation. The cheapest of the two is the intensity modulation where the bit stream is assigned different state levels. Some examples that are widely used in photonics are on-off keying (OOK) [Est+18] which indicates the presence and the absence of a signal and M-ary pulse amplitude modulation (PAM) [Bar+18a] which assigns a level to the signal according to its power. There has also been other forms of modulation, notably for Visible Light Communications, such as pulse position modulation (PPM) where a specific time slot within the duration of a symbol is assigned according to different levels of the signal power [Man+17b].

Intensity modulation only looks into the power of the signal, but it is also possible to exploit the phase of the signal. In photonics, M-ary phase shift keying (PSK) and M-ary quadrature amplitude modulation (QAM) are used; this is possible thanks to coherent detection. An advantage of multilevel modulation is that it increases data rate. For instance, going from a 4-QAM to a 16-QAM modulation doubles the data rate. Compared

to PSK, QAM has an extra degree of freedom in the complex domain and is more spectrally efficient although it can cause the signal to have a non-constant envelope in single carrier modulation (SCM) with a high order of modulation. Figure 1.6 illustrates that the extra degree of freedom entails less sensitivity to noise as for the same number of bits/symbol in PAM, PSK, and QAM, the QAM constellation is more spread out.

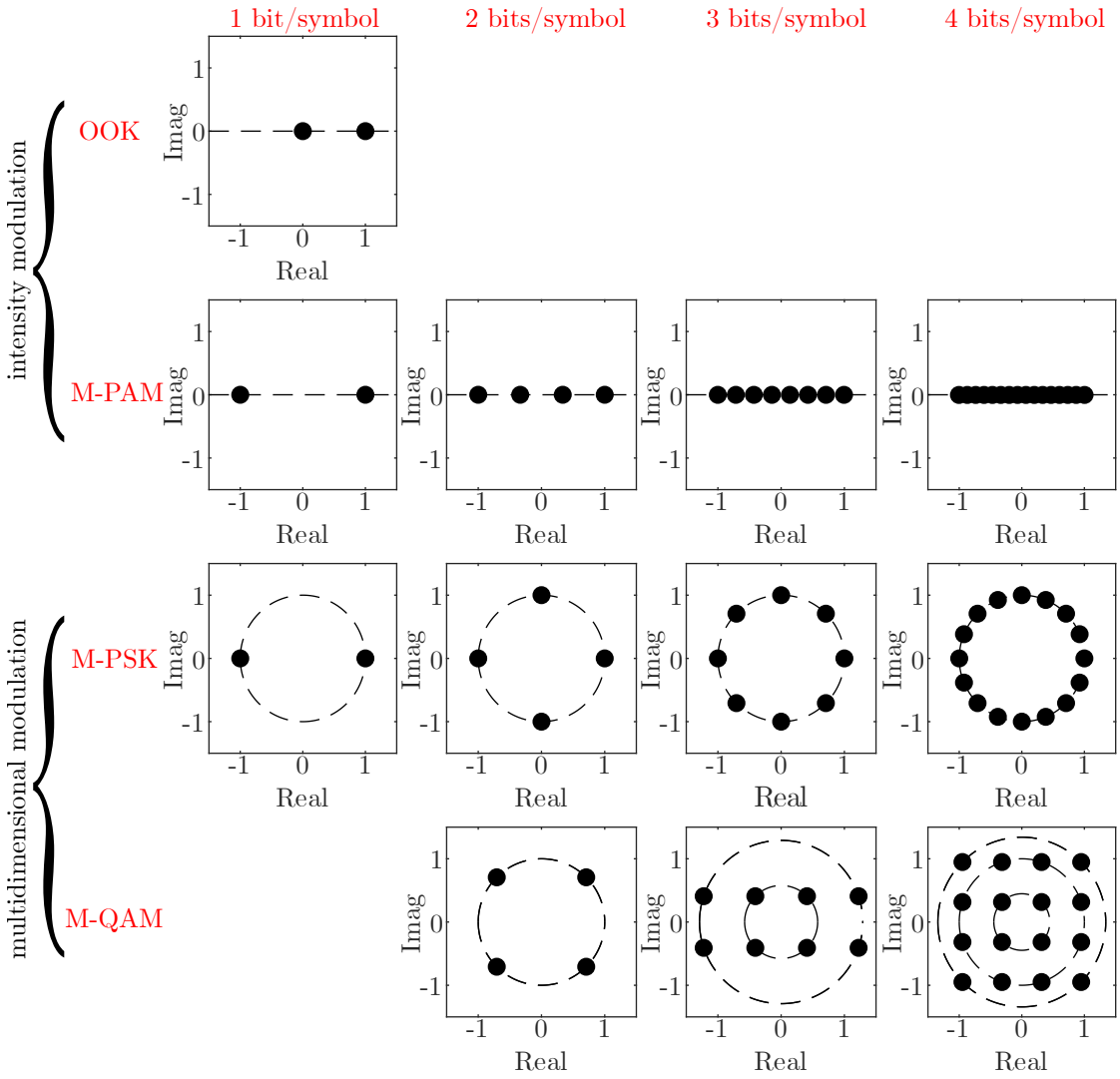


Figure 1.6 – Intensity and multidimensional modulation formats used in optical communication

The symbols s_k present in the different constellation diagrams shown in Figure 1.6 are brought into play either through SCM or multicarrier modulation (MCM). Figure 1.7 highlights the differences between the two where, for SCM, the symbols are assigned one

slot in the time domain whereas, and for MCM, the symbols are assigned one slot in the frequency domain. With the need for faster signal transmission it makes sense to look into increasing the capacity of the system and a way to achieve that goal while adding multiuser flexibility is through implementing MCM.

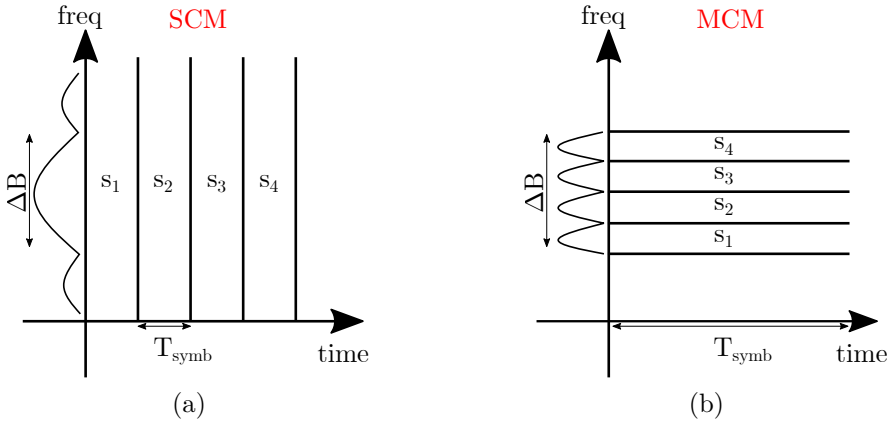


Figure 1.7 – (a) Single carrier modulation (SCM). The bandwidth here is considered to be the bandwidth after filtering. (b) Multicarrier modulation (MCM). ΔB: bandwidth, s_k: modulated symbols

In the case of MCM, the transmitted signal $s(t)$ has the following form:

$$s(t) = \sum_{k=0}^{N_{sc}-1} s_k e^{i2\pi f_k t} \qquad 0 \leq t \leq T_{symb} \tag{1.1}$$

where N_{sc} is the number of subcarriers, s_k is the modulated symbol injected onto the k^{th} subcarrier of frequency f_k, and T_{symb} is the symbol period.

Depending on the subcarrier spacing, the system could face intercarrier interference (ICI). To address that problem, a specific type of MCM called orthogonal frequency division multiplexing (OFDM) can be used where the adjacent frequencies in 1.1 have a frequency spacing $\Delta f = \frac{1}{T_{symb}}$ and $f_{k+1} - f_k = \Delta f$. This allows to greatly increase the spectral efficiency. The OFDM architecture has been used effectively in the RF field for some time and has been attracting a lot of attention in the optical field within the last decade since it introduces a method to overlap the subcarriers in such a way to increase spectral efficiency and remove ICI [Ter+02] [Arm09]. As it happens, Figure 1.8 shows that with OFDM it is possible to have twice the number of subcarriers, and thus data sent, for the same conventional MCM bandwidth where the maximum of a subcarrier spectrum

always falls on a null value for the other subcarrier spectra which avoids interference.

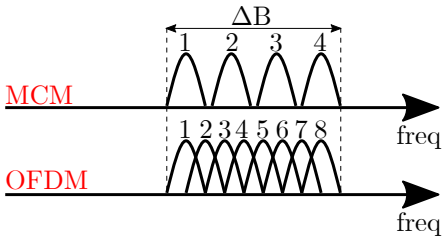


Figure 1.8 – Ordinary multicarrier vs. OFDM modulation. ΔB: bandwidth, MCM: multicarrier modulation, OFDM: orthogonal frequency division multiplexing

In its discrete form, the OFDM expression for one OFDM symbol can be written as

$$s[n] = \sqrt{N_{sc}\rho_{os}} \sum_{k=0}^{N_{sc}\rho_{os}-1} s_k e^{i2\pi \frac{k}{N_{sc}\rho_{os}} n} \tag{1.2}$$

where ρ_{os} is the oversampling factor. Equation 1.2 shows that an OFDM signal can easily be implemented via DSP through an inverse fast Fourier transform (iFFT). The iFFT is done with the frequencies centered around zero. Note that here the oversampling factor is not accounted for in the summation since the oversampling is seen as the addition of null subcarriers in the frequency domain. Also, the oversampling factor should be high enough for the sampling frequency to be well above the Shannon-Nyquist sampling theorem requirement while facilitating DSP but low enough to not require very large sampling frequencies. It was found that an oversampling factor of 4 is a good choice for OFDM [Yan+11]. The iFFT implementation reduces tremendously the number of complex multiplications required to construct a MCM signal while allowing for the use of a high number of subcarriers without needing to implement a complex architecture.

For these reasons, on top of OFDM's tolerance to transmission impairments like dispersion and capacity for allowing parallelization of DSP [Low+07], OFDM has been the object of comparative studies with SCM for photonics purposes [Vas+17] [Blu+19].

Many researchers in the photonics field are still exploring the value of OFDM [Jin+20] [El-+20]; it seems worthwhile in the context of this thesis to add to that discussion by diving into the topic of OFDM for photonics purposes.

1.3.1.1 Architecture of a cyclic prefix OFDM transmitter

In the upcoming paragraphs, we will describe the cyclic prefix orthogonal frequency division multiplexing (CP-OFDM) architecture as an example. As a few different terminologies are adopted to describe that construction, this section will serve to illustrate the chosen lexicon for OFDM based on the work of well-known researchers in the field [Kum11]. For the payload (or data) block, first a pseudo-random bit stream is modulated into symbols. An example of a constellation used to build a typical OFDM signal is shown in Figure 1.9a. There are two types of symbols in this case: the ones used for data and the ones used for pilots. The symbols used for the data can be any type of modulation. Since QAM has been shown to have better optical signal to noise ratio (OSNR) and bit error rate (BER) for OFDM than PSK and differential phase shift keying (DPSK) [Ras+18], it makes sense to focus on a QAM configuration. Here a 4-QAM modulation is shown. The symbols are coded into subcarriers as depicted in Figure 1.9b. Pilot design is shown to have an influence on system performances depending on its usage to help recover data [Zha+19] [Off+17] although in the context of this thesis work, the conventional 802.11 standard for pilot placement will be implemented [IEE17] where the pilots will be rotated 4-QAM symbols. Some subcarriers are nulled near DC and at both extremities to avoid the optical carrier and adjacent channels interference. Subsequently, some extra null subcarriers are added at both extremities in the frequency domain for the oversampling (not shown in Figure 1.9b.

Once the frequency domain construction is done, an iFFT is performed to move into the time domain as depicted in Figure 1.10a. Due to the propagation channel, a cyclic prefix is added (Figure 1.10b) to prevent what we will refer to as OFDM symbol interference in this thesis. Note that there are other types of guard intervals but cyclic prefix allows better channel estimation and has better signal to noise ratio (SNR) performance than, for instance, zero padding (ZP) and known symbol padding (KSP) [Ste+07]. The signal shown in Figure 1.10b consists of an OFDM symbol and its cyclic prefix.

Once the OFDM symbol is constructed in the time domain, a frame consisting of multiple OFDM symbols is built. Figure 1.11 details that process where the frame includes a preamble component as well as the data. The preamble can be further divided into two categories: preamble used for synchronization and preamble used for equalization. Their

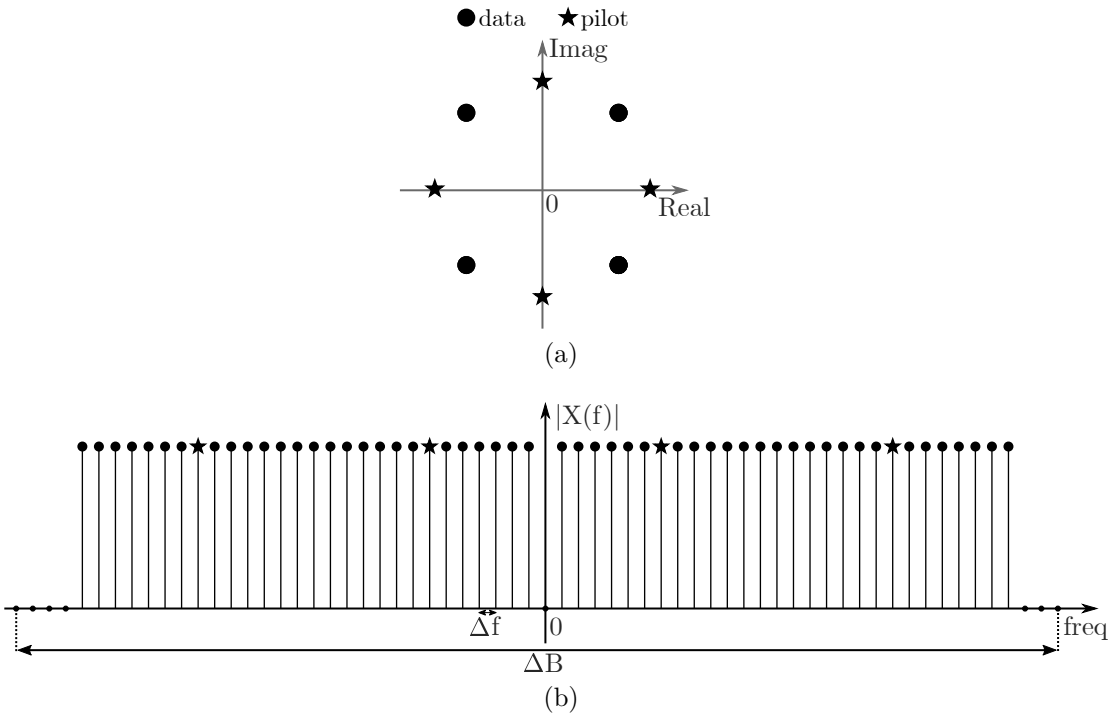


Figure 1.9 – (a) Data (4-QAM) and pilot (rotated 4-QAM) symbols for a 64 subcarriers configuration with 4 pilots per OFDM symbol (b) Subcarriers associated with the data and pilot symbols. The null subcarriers used for the oversampling are not shown. ΔB: bandwidth, Δf: frequency spacing

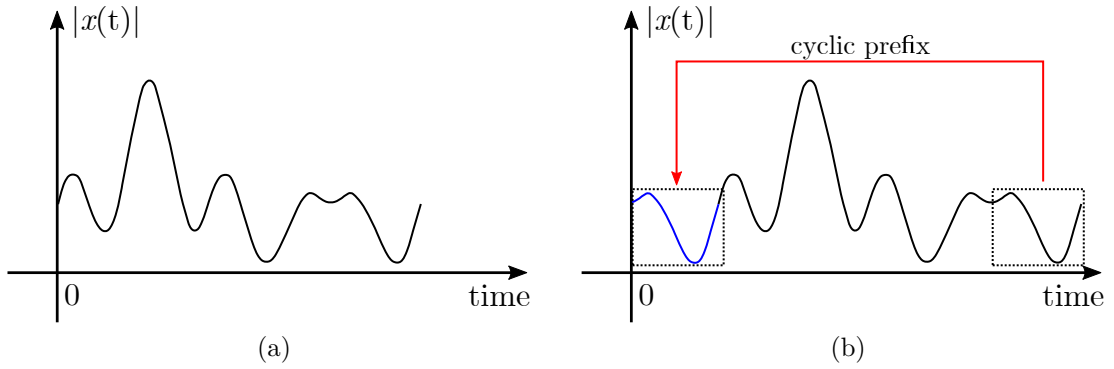


Figure 1.10 – (a) OFDM symbol (b) OFDM symbol with its cyclic prefix

specific designs will be discussed in the next paragraph. The resulting data rate is

$$R_b = \frac{N_{data}\ (1-\rho_{cp})\ \log_2(M)\,\Delta B}{N_{sc}\,N_{ofdm}} \tag{1.3}$$

where N_{data} is the total number of QAM data symbols in the OFDM frame, ρ_{cp} is the cyclic prefix factor, M is the QAM modulation order, ΔB is the signal bandwidth excluding the oversampling, N_{sc} is the number of subcarriers per OFDM symbol, and N_{ofdm} is the total number of OFDM symbols including the synchronization and equalization preambles.

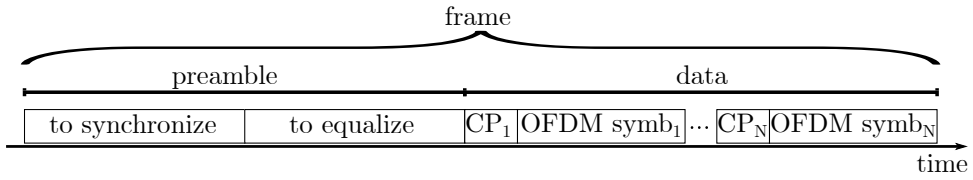


Figure 1.11 – OFDM frame consisting here of one synchronization and one equalization preamble, and N data symbols

Although CP-OFDM signals bring many advantages, they also exhibit a non-constant envelope (cf. Figure 1.10a). This results in CP-OFDM signals having intrinsically high instantaneous power characterized by a high peak to average power ratio (PAPR). PAPR is defined by

$$\mathrm{PAPR}_{dB} = 10\log_{10}\left(\frac{\max\left(|s[n]|^2\right)}{\mathrm{E}\left[|s[n]|^2\right]}\right) \tag{1.4}$$

From Equations 1.2 and 1.4, it can be inferred that as the number of subcarriers increases, so does the potential for a high PAPR; this is detailed in [Shi+09]. The maximum PAPR is encountered if all emitted symbols are identical which is unlikely since the data stream is assumed to be statistically independent and identically distributed (i.i.d); even so, it is still possible to face high PAPR. The unit of measure to determine the likelihood of having high PAPR is the complementary cumulative density function (CCDF) which, taking into consideration the Rayleigh nature of an OFDM signal's distribution, is defined by:

$$\mathrm{CCDF} = \Pr(\mathrm{PAPR} > \mathrm{PAPR}_{ref}) \tag{1.5}$$

The theoretical CCDF curve is shown in Figure 1.12 and gives an idea of the likelihood of having a particular PAPR. For example, for a N_{sc} of 128, there is a 4×10^{-1} (or 40%) probability of getting a PAPR that is greater than 8 dB for one OFDM symbol.

The CCDF curve gives statistical information only on the power maxima. The instantaneous power of the OFDM signal varies frequently; another important complementary indication to look into while studying OFDM signals will be the probability density function (PDF). Figure 1.13 represents the PDF of the amplitude of an OFDM frame with 128 subcarriers and shows that an OFDM signal is characterized by a Rayleigh distribution.

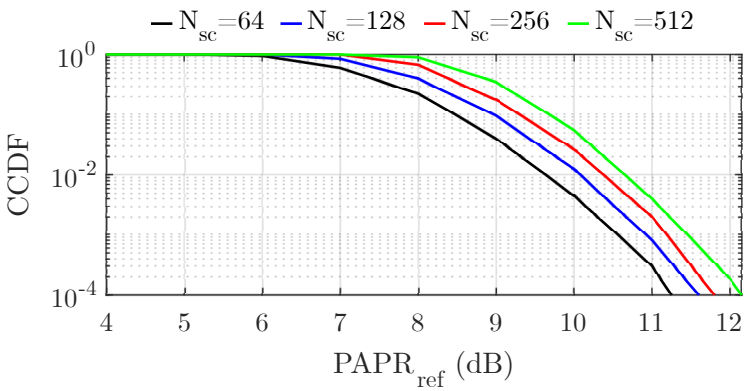


Figure 1.12 – Complementary cumulative density function of the peak to average power ratio of a 4 times oversampled OFDM signal for different numbers of subcarriers N_{sc} [Gam+16]

From Figure 1.13, it is possible to quantify the number of high power samples. For this frame, the maximum amplitude is 1.55. To put it in perspective, the theoretical maximum for $N_{data} = N_{sc} = 128$, $\rho_{os} = 4$, and $|s_k| = 1$ assuming all 4-QAM symbols are aligned would be $\max(s[n]) = \frac{N_{data}}{\sqrt{N_{sc}\rho_{os}}} = 5.65$ (with $\text{PAPR}_{max} = 10\log_{10}(N_{data}) = 21$ dB).

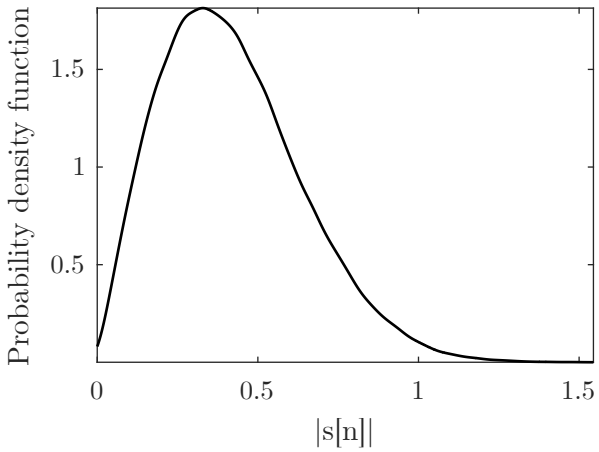


Figure 1.13 – Probability density function of the amplitude of 256 4-QAM OFDM symbols with 128 subcarriers and an oversampling factor of 4 obtained through simulation

The characteristic of high PAPR as well as the high power samples will cause some issues in the transmission as it will give rise to nonlinear effects notably while going through the DAC, the IQ modulator, and the amplifier [Ami+15]. It is possible to reduce the impact of PAPR by hard-clipping the signal. This point will be the subject of another

section. With the benefits of high spectral efficiency, a low DSP complexity and robustness to channel dispersion, OFDM places itself as a worthwhile subject of investigation for this thesis work which will aim to reduce the impact of its drawback.

1.3.1.2 Beyond conventional OFDM

As was said in the previous section, the focus of this thesis will be the conventional CP-OFDM. However, one should be aware that other OFDM formats have stemmed from the classical OFDM scheme. This thesis work should therefore be taken as a baseline that can be translated to the other configurations, some of which will be introduced in this section.

The DC-biased OFDM (DCO-OFDM) [San+15] [Per+15] and asymmetrically clipped OFDM (ACO-OFDM) [Per+15] were proposed to make RF-OFDM suitable for direct detection going from a bipolar, complex-valued signal to a unipolar, real-valued signal. The spectral and energy efficient OFDM (SEE-OFDM) [Elg+14] goes one step further by improving the spectral efficiency of that real-valued signal.

The adaptively modulated OFDM (AMOOFDM) adjusts the modulation format of the symbol assigned to a subcarrier according to the characteristics of the transmission system. For instance, a high modulation format is applied to a subcarrier experiencing a high SNR and vice versa. This scheme is shown in Figure 1.14 which illustrates one of the benefits of OFDM in improving the signal transmission capacity, network flexibility, and performance robustness [Gia+10].

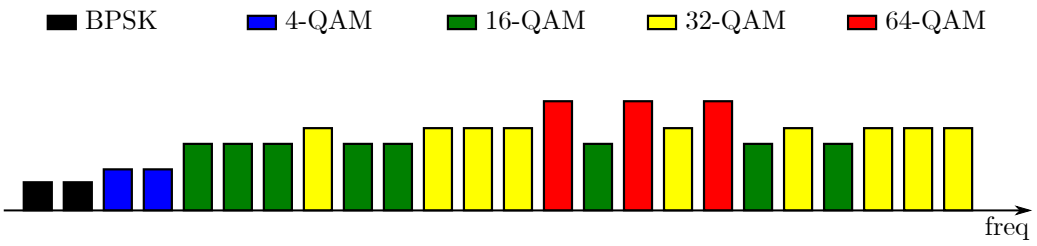


Figure 1.14 – Example of a signal modulation format distribution across the subcarriers of an adaptively modulated OFDM system

The constant envelope OFDM (CE-OFDM) which aims to solve the high PAPR of conventional OFDM by transforming the OFDM signal to one fit for efficient amplification via phase modulation [Tho+08]. This form of OFDM is currently being investigated in

photonics for direct detection [Nun+14] and coherent detection [Dia+17] showing promising results.

The all optical OFDM (AO-OFDM) which uses optical signal processing to construct the OFDM signal in an attempt to work with higher speeds while having inherently low power consumption. Different methods are implemented including optical fast Fourier transform (FFT) [Hil+10], the use of wavelength selective switches (WSS) [Sch+13a], or the use of Mach-Zehnder interferometers (MZIs) and recirculating frequency shift (RFS) [Mar+18].

The discrete multitone (DMT) approach which is a real-valued baseband version of OFDM that can be used for instance in orthogonal frequency division multiple access (OFDMA) systems to increase the tolerance to IQ imbalances and IQ time skew [Jun+16].

Offset-QAM OFDM (OFDM/OQAM) is attracting increasing attention in photonics as it has high spectral efficiency and receiver sensitivity with a higher side lobe suppression than conventional OFDM [Li+16]. The filter bank multicarrier (FBMC) is another MCM scheme that is starting to be considered for future optical networks because it is able to increase bandwidth efficiency and reduce side lobes in exchange for higher complexity in implementation [Ibr+17] [You+19].

In the case where DSP is used to construct the signal, the next step would be to convert that signal into the electrical domain.

1.3.1.3 From the electrical to the optical domain

The constructed discrete signal has to be converted into a continuous time signal via the digital to analog conversion (DAC). As far as the DAC is concerned, the resolution or the effective number of bits (ENOB) will be the parameters to look into as they will dictate the quantization intervals thus the analog levels attributed to the signal. The higher the resolution the less quantization noise will appear. At the DAC output is placed a low-pass reconstruction filter to smooth out the analog output.

Once the electrical signal has been constructed, it has to be converted into the optical domain. That step requires a laser lightwave. It can be characterized by the following equation assuming a Lorentzian spectral shape [Bar+90]

$$E(t) = \mathrm{Re}\left\{\sqrt{P_s}e^{i(\omega_{LO}t+\theta(t)+\theta_0)}\right\} = \mathrm{Re}\left\{\varepsilon(t)e^{i\omega_{LO}t}\right\} \tag{1.6}$$

where $E(t)$ is the electric field associated with the optical wave, $\varepsilon(t)$ is the complex

envelope of the signal, P_s is the source power, ω_{LO} is the laser angular frequency, θ_0 is a random initial phase in the interval $[0, 2\pi[$, and $\theta(t)$ is the laser phase noise process (in this case a Wiener process) which derivative is a zero-mean white Gaussian process with power spectral density (PSD)

$$S_{\dot{\theta}}(\omega) = 2\pi\Delta\nu \tag{1.7}$$

where $\Delta\nu$ is the laser linewidth. Since the laser lightwave is operated as a continuous wave (CW) to avoid frequency chirping caused by a change in instantaneous power in the case of a direct modulation, an external modulation needs to be performed. The IQ modulator consists of two branches that contain Mach-Zehnder modulator (MZM) and are dephased by $\phi = \frac{\pi}{2}$.

Before diving into the IQ modulator, the MZM is further explored through Figure 1.15 as this device is the basis for the optical phase and intensity modulation.

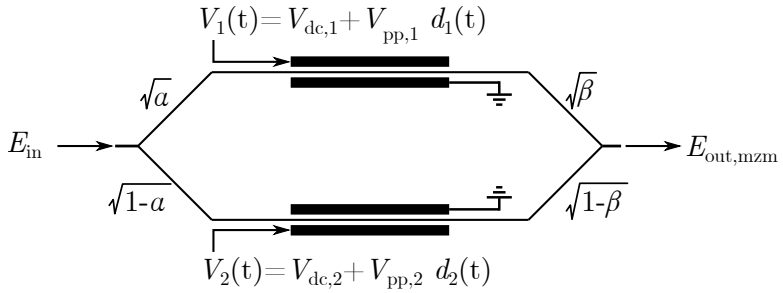


Figure 1.15 – Architecture of a Mach-Zehnder modulator. V_j is the voltage applied to the upper ($j = 1$) and lower ($j = 2$) arm of the MZM with $V_{dc,j}$ their DC components, $V_{pp,j}$ their peak-to-peak voltage and d_j their modulating signal components, E_{in} and $E_{out,mzm}$ are the electric fields associated with the optical input and output waves, α and β are the power splitting and combining ratios

That modulator is able to generate a phase shift in the light wave going through it as a result of the Pockels effect which states that if an electric field is applied to some material, its refractive index changes relative to the intensity of said electric field. That change in phase is related to the voltages driving the MZM as follows:

$$\phi_{1,2}(t) = \frac{\pi V_{1,2}(t)}{V_\pi} \tag{1.8}$$

where V_π, the half-wave voltage, is the driving voltage that provokes a phase shift of π.

This input-output relation of the MZM can be expressed by

$$\varepsilon_{out,mzm}(t) = \varepsilon_{in}(t)\left(\rho_1\sqrt{\alpha}\sqrt{\beta}e^{-i\phi_1} + \rho_2\sqrt{1-\alpha}\sqrt{1-\beta}e^{-i\phi_2}\right) \tag{1.9}$$

where $\varepsilon_{out,mzm}(t)$ is the complex envelop of the signal at the MZM output, ρ_j, α, and β determine the excess loss, and the splitting and combining ratios of the two MZM arms. Using the information collected from Equations (1.8) and Figure 1.15, and assuming that there is no imbalance in the splitting and combining ratios ($\alpha = \beta = \frac{1}{2}$), no excess loss ($\rho_1 = \rho_2 = 1$), and the MZM is operating at a chirp free biased null point ($V_{dc,1} = V_{dc,2} = V_\pi, V_{pp,1} = V_{pp,2} = V_{pp} \quad and \quad d_1(t) = -d_2(t)$), Equation 1.9 becomes

$$\begin{aligned}\varepsilon_{out,mzm}(t) &= \frac{1}{2}\varepsilon_{in}(t)\left(e^{-i\frac{\pi}{V_\pi}(V_\pi + V_{pp}d_1(t))} + e^{-i\frac{\pi}{V_\pi}(V_\pi - V_{pp}d_1(t))}\right) \\ &= -\varepsilon_{in}(t)\cos\left(\frac{\pi}{V_\pi}V_{pp}d_1(t)\right)\end{aligned} \tag{1.10}$$

where $\varepsilon_{in}(t)$ follows the Lorentzian spectral shape defined in Equation 1.6, V_{pp} is the fixed peak-to-peak voltage and $d_1(t)$ is the transmitted signal's normalized waveform ($d_1(t) \in [-\frac{1}{2}, \frac{1}{2}]$).

The previous definition of the MZM demonstrates an ability to do simple modulations like binary phase shift keying (BPSK) or PAM. In order to represent symbols from higher order modulations in the complex plane, it is necessary to return to the IQ modulator setup in Figure 1.16.

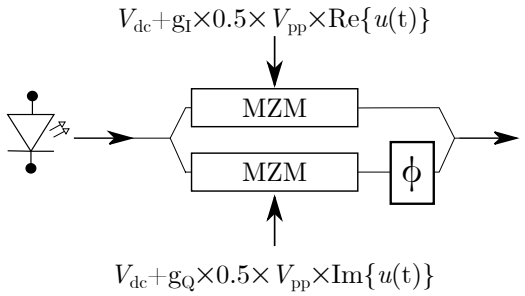


Figure 1.16 – IQ modulator based on MZMs. V_{dc} and V_{pp} are the DC component and the peak-to-peak voltage of the driving signal applied to each branch of the IQ modulator, g_I and g_Q are the gain for each branch with $g_Q = g_I = 1$, and $u(t)$ is the modulated waveform

The signal $u(t)$ is divided into its real and imaginary parts $u_I(t)$ and $u_Q(t)$ bringing

about the following output [Peu12]

$$\varepsilon_{out}(t) = -\frac{1}{2}\varepsilon_{in}(t)\left[\cos\left(\frac{\pi}{V_\pi}V_{pp}u_I(t)\right) + \cos\left(\frac{\pi}{V_\pi}V_{pp}u_Q(t)\right)e^{i\frac{\pi}{2}}\right] \tag{1.11}$$

It is important to note that variations in the instantaneous power of the signal will affect the IQ modulator and force it into its nonlinear region. Looking at the MZM's transfer function in Figure 1.17 helps explain that point. The electrical signal (in blue) goes into the MZM and is converted to the optical signal (in red). When the signal's samples have a low amplitude, the MZM is operating close to V_π, but the high peaks push the MZM into its nonlinear regime. This entails that when working with non-constant envelope signals like CP-OFDM, the transmission will experience nonlinearities coming from the IQ modulator.

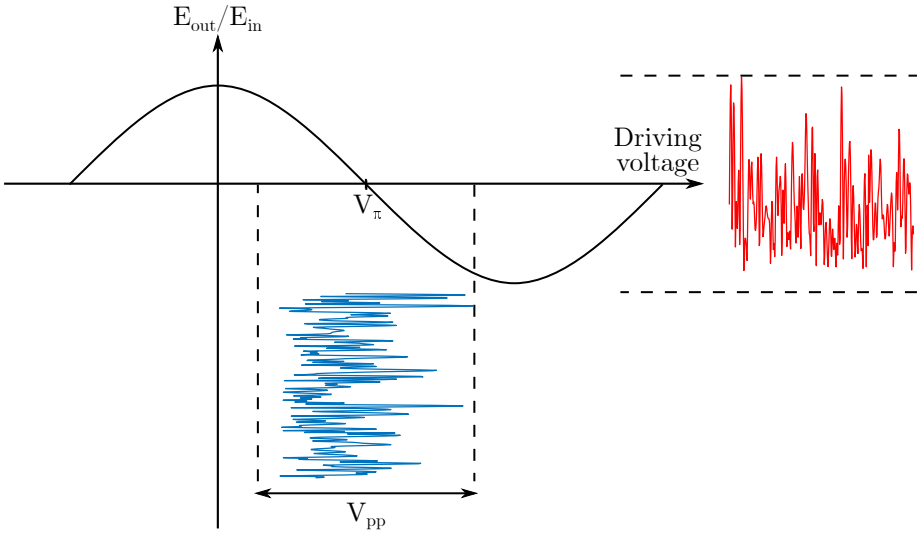


Figure 1.17 – Transfer function of the MZM biased at null point for OFDM. V_{pp} is the signal's peak to peak voltage and V_π is the half-wave voltage

Once the electrical signal has been converted into the optical domain, it is usually amplified by the booster amplifier; the amplification aspect will be described in more detail later on. After the amplification, the signal is ready to pass through the fiber link.

1.3.2 Optical channel

The fiber link is composed of spans (optic fiber + amplifier) that can add up to several hundreds of kilometers. The following section dives into the characteristics of the optic

fiber.

Light guiding in the fiber is accomplished through total internal reflection. Light is contained inside the fiber thanks to the core of the fiber having a higher refractive index than that of the cladding. There are multiple types of fibers including multimode and single-mode. The difference is found in the core sizes where the multimode fiber's core is larger allowing light to take multiple paths but causing dispersion; on the other hand, the single-mode fiber (SMF) only allows light to travel in one path with its small core meaning there will only be one mode of propagation. Some other characteristics come into play since the fiber is flawed by design and therefore will bring some disruptions in the form of:

— Attenuation which manifests into a decrease in the power of the signal caused by light absorption and scattering. The attenuation rate in dB per km, α_{dB}, depends on the wavelength with the minimum attenuation of 0.2 dB/km for silica fibers found at around 1550 nm [Agr12]

— Chromatic dispersion which results from different propagation speeds of the distinct frequency components associated to the transmitted signal. Figure 1.18 shows how this phenomenon can lead to pulse spreading which affects the quality of transmission.

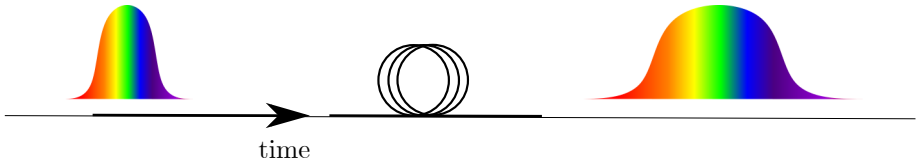


Figure 1.18 – Example of chromatic dispersion (different colors associated to distinct frequency components)

The all-pass transfer function of the fiber with chromatic dispersion is [Rob+09]

$$H_{CD}(\omega) = e^{-\frac{j\beta_2\omega^2 L}{2}} \tag{1.12}$$

where ω is the angular frequency of the signal, and L is the fiber length in km. β_2 is the group-velocity dispersion (GVD) parameter defined by:

$$\beta_2 = -\frac{D\lambda^2}{2\pi c} \tag{1.13}$$

with D the dispersion parameter of the fiber in ps/(nm-km) (with values of 15-18

ps/(nm-km) for 1550 nm), λ the wavelength of the laser source in nm, and c the speed of light in a vacuum in m/s. A more detailed study of the effect of chromatic dispersion in the context of this thesis will be shown in a later chapter.

— Polarization mode dispersion (PMD) since light will travel differently depending on its polarization which will trigger light pulse spreading not only in the case of multiple polarizations, but also in single polarization as light can rotate within the waveguide. Figure 1.19 shows an example of how PMD can contribute to pulse broadening. The PMD parameter usually denoted D_p is a measure of PMD where typical values for modern fibers are such that $D_p < 0.1$ ps/$\sqrt{\text{km}}$ [Agr12].

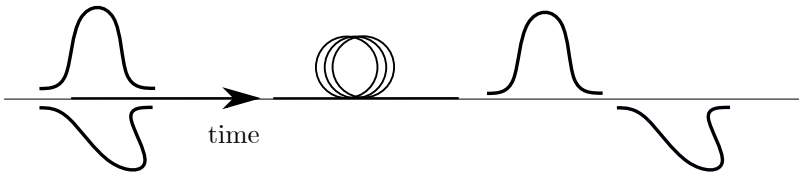


Figure 1.19 – Example of polarization mode dispersion

— Nonlinear effects that appear when the transmitted light has high power and can be divided into:
  — Kerr effects that come about because the refractive index of the material depends on the light intensity
  — Stimulated scattering due to the interaction of the signal with the material of the fiber

1.3.3 Optical amplifiers

As stated previously, the signal will experience attenuation while traveling through the fiber due to light absorption and Rayleigh scattering [Laf+07]. It means that once it reaches the receiver, the signal power will be lower than it was at the emitter. The quality of transmission depends on the receiver sensitivity, and if the signal's power at the receiver is too low, this will cause some errors when retrieving the sent signal. There are two ways to remediate this situation: using optical-electrical-optical (OEO) regenerators or using optical amplifiers. Some years ago, extensive research was made to achieve transparency (all optical communications without electrical/optical conversion) [Rou14] [Agr+16]. Nowadays, this trend prevails as optical amplifiers are being used.

Assuming that on average an amplifier between fiber links has a gain of 20 dB, ac-

counting for attenuation, the signal will need to be amplified ideally about every 80 km [ITU09]. As shown in Figure 1.4, amplifiers can be placed in three different locations in a transmission system: at the transmitter where it acts as a booster, between fiber links where it is in-line, and at the receiver where it behaves like a pre-amplifier. Boosters usually have high output saturation power as the goal is to give the signal more power to have a longer reach, pre-amplifiers are focused on having a low noise factor as they are used at the end of the transmission while having a high enough gain to meet the receiver's sensitivity requirements. The in-line amplifier is a combination; it needs a moderate noise factor and gain in order to compensate attenuation properly.

Optical amplifiers can be divided into three main categories of technology: doped fiber amplifiers, nonlinear effects based amplifiers, and semiconductor amplifiers. There are multiple types of doped fiber amplifiers that operate at different bands such as Erbium and Praseodymium amplifiers with the Erbium doped fiber amplifier (EDFA) being overall the most commercially used optical amplifier with its high gain and saturation point, and low noise.

1.3.3.1 Erbium doped fiber amplifiers

An EDFA relies on population inversion to perform the amplification. Figure 1.20 illustrates the amplification process by looking at an energy level diagram of Erbium ions. A pump laser pushes the ion into an excited state. The EDFA can be pumped at 980 nm or 1480 nm. The 1480 nm pump laser puts the ion in the metastable state where the lifetime of an ion is about 10 ms. Then once a photon passes through, the ion transitions into the lower level which releases another photon in the C-band (1530-1565 nm) or the L-band (1565-1625 nm). In the case of 980 nm, the ion is put into an upper energy level with a short lifetime of 1 μs. The ion then quickly transitions to the metastable level and radiates heat without emitting any photons. Once in the metastable level, the ion goes to the lower level and produces stimulated emission. The EDFA unfortunately produces noise on top of the stimulated emission which is characterized by its noise figure (NF). The cause of the amplifier noise is amplified spontaneous emission (ASE) which appears when an ion randomly returns to the lower energy state. Since 980 nm pumping tends to yield better gain and NF, it is more often used.

Figure 1.21 portrays a schematic diagram of the EDFA. The input optical signal goes through an isolator then is coupled with a pump laser propagating in the same direction (forward pumping). The coupled signal goes through a few meters of Erbium-doped fiber

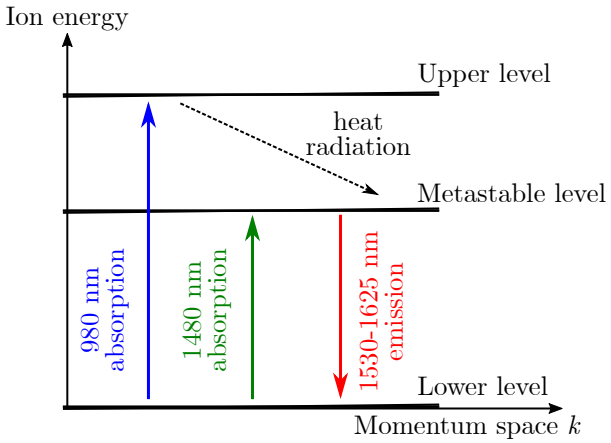


Figure 1.20 – Energy level diagram of Erbium ions

(3-15 m) to then be coupled with a pump laser propagating in the opposite direction as the signal (backward pumping). Combining forward and backward pumping serves to have a more uniform gain.

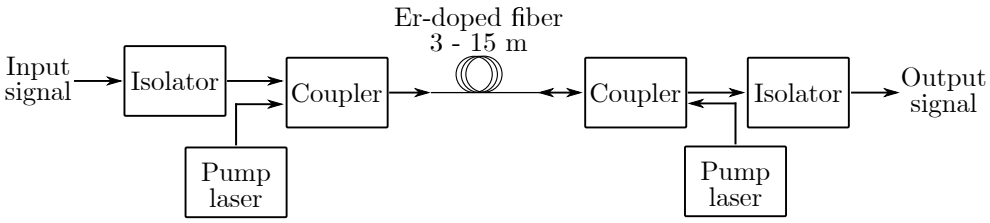


Figure 1.21 – Schematics of a bidirectionally pumped EDFA

With increasing data rates and orders of modulation, higher OSNRs are required which means the longer reach obtained from EDFAs becomes limited. This is where the need for Raman amplification can come into play.

1.3.3.2 Raman amplifiers

Nonlinear effects based amplifiers include Raman amplifiers and Brillouin amplifiers. Due to their narrow gain spectrum, Brillouin amplifiers are not considered for high data rates so far. On the other hand, Raman amplifiers are most often used notably in combination with EDFAs.

Figure 1.22 shows an example of Raman amplification. Contrary to the EDFA where a few meters of Erbium-doped fiber is pumped, the Raman amplification uses a pump laser

counter-propagating on the kilometers of transmission fiber, thus using the transmission fiber as the gain medium.

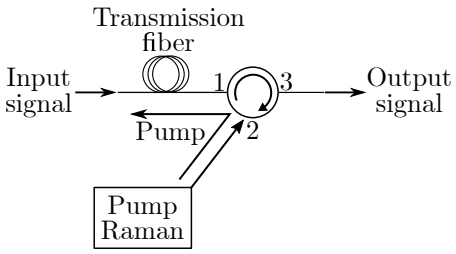


Figure 1.22 – Schematics of Raman amplification

A Raman amplifier uses a phenomenon that occurs in the fiber called stimulated Raman scattering (SRS). In the case of the Raman amplifier, Figure 1.23 shows how this phenomenon is used for amplification. A pump laser is applied at some wavelength. Some of the pump photons' energy ends up creating photons at other frequencies with lower energy. The remaining energy is absorbed by silica molecules which go into a vibrational state (optical phonon). The Raman gain spectrum depends on the pump wavelength and has a maximum gain at about 100 nm higher than the pump wavelength. Since the scattering is more likely to appear at frequencies where photons are already present, the signal to be amplified will be sent at that 100 nm mark (for a 1450 nm pump, a signal propagating at 1550 nm). Note that to achieve amplification in multiple bands, multiple pump lasers are used.

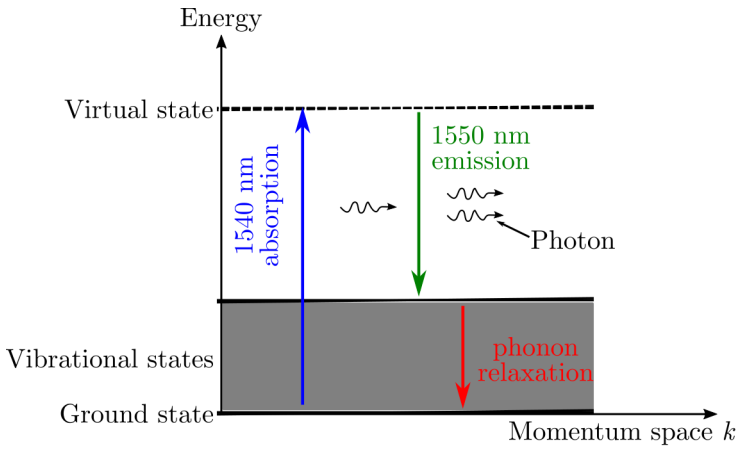


Figure 1.23 – Schematics of energy levels for Raman scattering. This diagram represents the Stokes process

With the Raman effect in silica fibers being weak, pump lasers are unfortunately required to have higher powers. The next optical amplifiers that will be presented do not require this kind of power to operate and use electrical pumping instead.

1.3.3.3 Semiconductor optical amplifiers

Semiconductor optical amplifiers (SOAs) are not as often used currently in optical communications. However, investigating the applications with the SOA is relevant in metropolitan networks as it has a higher bandwidth and a lower cost while remaining compact (they are manufactured as chips) [Sch+14] [Koe+14].

As shown in Figure 1.24, SOAs have a similar structure to that of Fabry-Perot lasers with the exception of having anti-reflection facets (with a reflectivity of less than 0.1%) making it a traveling wave (TW) device. Usually, the materials used for the SOA are InGaAs, InGaAsP, and InP as they have great quantum efficiency and thus allow for more generated photons for the amplification that will take place in the active region.

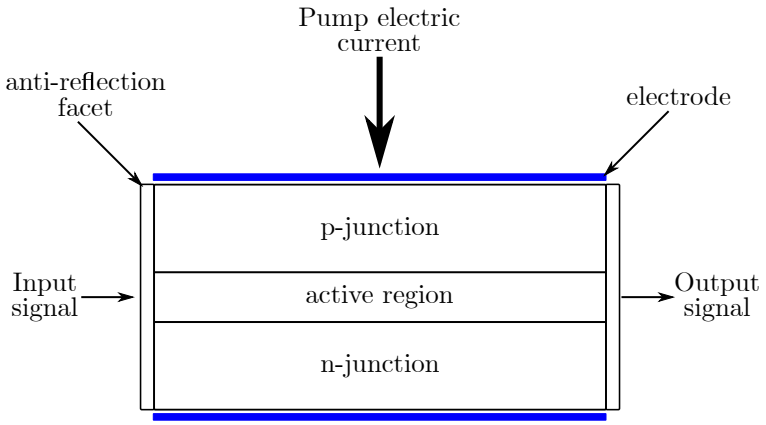


Figure 1.24 – Schematic example of an SOA

Similar to EDFAs, SOAs rely on population inversion for the amplification process. The SOA is able to amplify an incident laser when it is electrically pumped. As seen in Figure 1.25, the SOA is able to amplify the optical signal as a result of the bias current I_{bias} (the pump electric current) providing free electrons (carriers) that, when stimulated by a photon (the optical signal in this case) move toward a lower energy level and release another photon. That phenomenon is called interband interactions.

In an SOA, there will automatically be spontaneous recombination that will provoke ASE which is a source of noise that accumulates with amplifiers in-line. Compared to

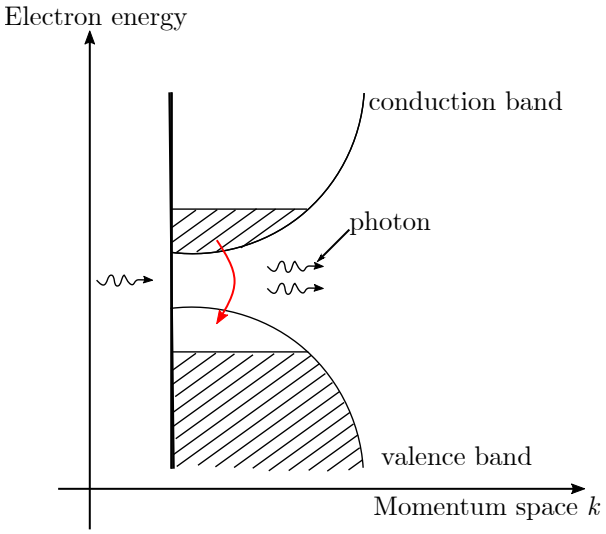


Figure 1.25 – Amplification in the SOA through stimulated emission

EDFAs, SOAs tend to have higher NF and lower gain. The gain is calculated with the following equation

$$G_{dB} = 10\log_{10}\left(\frac{P_{out}}{P_{in}}\right) \tag{1.14}$$

P_{in} and P_{out} are the input and output signals' average powers respectively.

The gain, signal output power P_{out}, and ASE depend on the bias current I_{bias} and the signal input power P_{in} as I_{bias} and P_{in} determine how many carriers (and thus gain) will be available for the amplification. Figure 1.26 shows the evolution of the SOA operating point depending on the bias current and signal input power.

The stages are the following:

— At stage A, the SOA is pumped by the bias current I_{bias} therefore there is a corresponding available gain. Since there is no input signal, there is also no output signal and all that is left is the ASE

— At stage B, an input signal is sent. A higher input power means a lower amount of available carriers leading to a saturation of the gain. The input signal is amplified which explains the increase in the output signal power. Since stimulated emission is performed, there is less instances for spontaneous emission thus reducing the ASE

— At stage C, all parameters remain unchanged

— At stage D, for the same input signal power, the bias current is raised. This increases the gain which in turn increases the output signal power. A higher bias

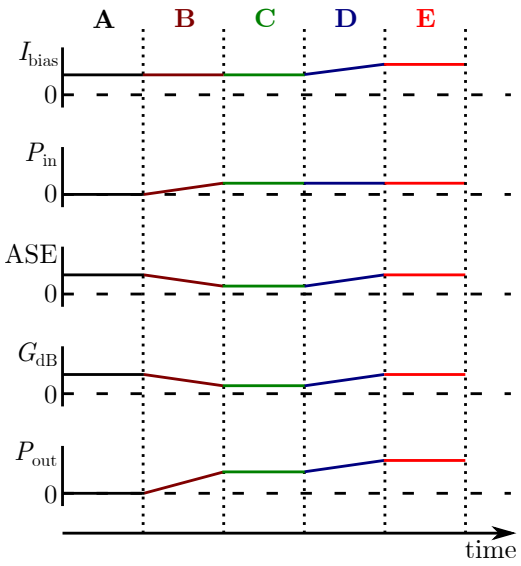


Figure 1.26 – Evolution of SOA operating point when I_{bias} and P_{in} vary. I_{bias}: bias current, P_{in}/P_{out}: signal input/output power, G_{dB}: gain, ASE: amplified spontaneous emission

current for the same input power yields more ASE

— At stage E, all parameters remain unchanged

With high input powers the SOA starts entering its nonlinear regime. There are multiple nonlinear effects that the SOA brings forth:

— Self-polarization rotation (SPR) and cross-polarization modulation (XPolM) which are a nonlinear rotation of the state of polarization of the signal on itself which depends on the signal input power and the supply current of the SOA, and a rotation brought about by another signal that is injected simultaneously respectively

— Self-gain modulation (SGM) and cross-gain modulation (XGM) which are changes in the gain of the signal due to the variation of the input power of the signal, and of the input power of a second signal respectively

— Self-phase modulation (SPM) and cross-phase modulation (XPM) which are changes in the phase of the signal due to the variation of the input power of the signal, and of the input power of a second signal respectively

— Four wave mixing (FWM) which occurs when optical signals at different wavelengths create new optical signals that can interfere with the other signals. FWM can arise from both interband and intraband effects. An example of FWM is shown for a two tone signal in Figure 1.27. The two tones having the same polar-

ization cause the amplifier gain to be modulated at the beat frequency $\Omega = f_2 - f_1$ [Con07]. This gain modulation generates new signals at frequencies f_3 and f_4 such that $f_3 = 2f_1 - f_2$ and $2f_2 - f_1$. These third order intermodulation products can fall in-band or out-band. The out-band effects can usually be filtered out, but other means will have to be used for the in-band effects. FWM tends to be more prominent in OFDM [Kha12] with its impact increasing in the case of high PAPR [Pec+09]; this will be an aspect to examine in this thesis.

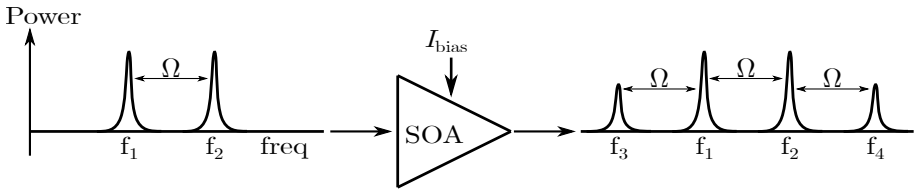


Figure 1.27 – Example of four wave mixing caused by the semiconductor optical amplifier for a two tone configuration. I_{bias} is the current supplying the SOA, $f_{1,2}$ are the frequencies of the signal, $f_{3,4}$ are the frequencies generated from FWM, and Ω is the beat frequency where $\Omega = f_2 - f_1$

All these effects are caused by the gain behavior of the SOA which can be divided into slow interband interactions (in the order of ns) and fast intraband interactions (in the order of ps). In this thesis, the focus will be on interband interactions since they will not be negligible for the purposes of this thesis. The bias current determines the gain's recovery period (interband effect). Figure 1.28 illustrates how memory effects can appear when the original gain is not yet recovered (scenario in Figure 1.28b); the gain and therefore the output will depend on previous input values.

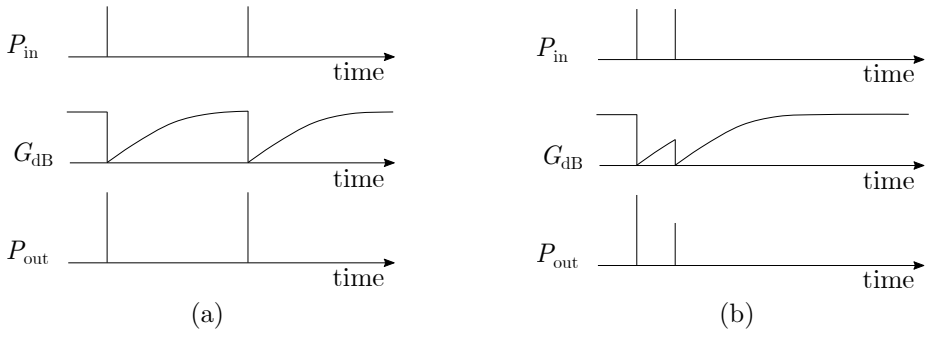


Figure 1.28 – (a) Gain for largely spaced pulses (b) Gain for tightly spaced pulses

A way to characterize both the static and dynamic behavior of the SOA is by looking at

the amplitude modulation to amplitude modulation (AM/AM) and amplitude modulation to phase modulation (AM/PM) characteristics of the SOA shown in Figure 1.29. Figure 1.29a represents the amplitude of the output signal vs. the amplitude of the input signal while Figure 1.29b displays the phase shift between the input and output signal vs. the amplitude of the input signal. In both cases, the red line portrays the behavior of a linear amplifier. The static effects are apparent as the curve does not match that of a linear amplifier. The shaded region shows the dynamic (memory) effects.

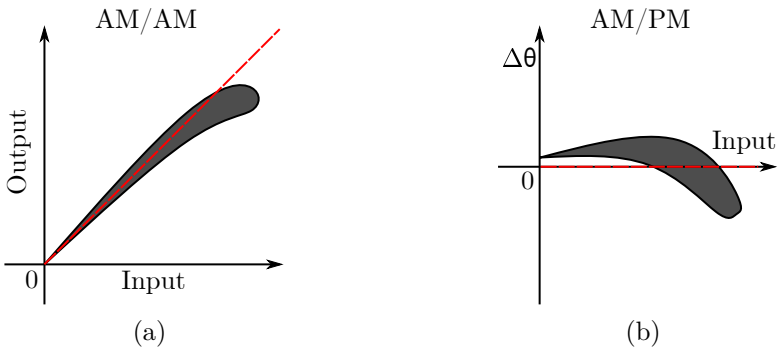


Figure 1.29 – (a) AM/AM characteristic of an SOA (b) AM/PM characteristic of an SOA. The red line is the curve for a linear amplifier. The shaded region shows the dynamic behavior

Now that the main optical amplifiers have been described, the next chapter opens the discussion regarding what types of optical amplifiers would be more suited for access/metropolitan networks.

1.3.3.4 Optical amplifiers comparison for access/metropolitan networks

With limited space, one of the challenges of optical communications is the size of components. Figure 1.30 shows the relative dimensions of a commercial module EDFA, a commercial module Raman amplifier, and a commercial SOA with their packaging. Commercial module amplifiers are designed to be more compact than rack-mounted and bench-top amplifiers. It can be seen that the SOA is much more compact. It is possible to reduce the EDFA coil diameter with a slight trade-off in efficiency, but the extent of this reduction is limited by the fiber size [Lia+06]. The SOA on the other hand has a chip length in the range of hundreds of microns [Sai+11] with the added value that it can easily be integrated with other optical components [Bud+15].

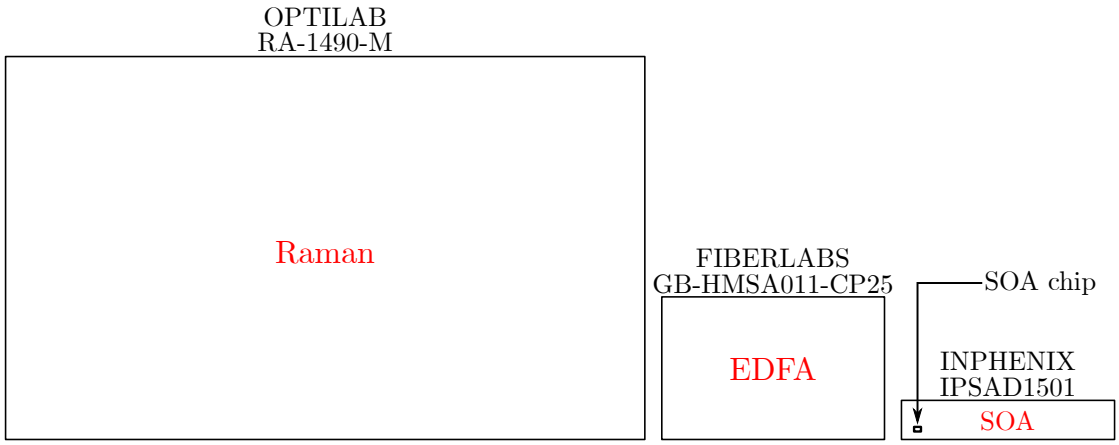


Figure 1.30 – Relative size comparison of commercial module amplifiers with their packaging. EDFA: Erbium doped fiber amplifier, SOA: semiconductor optical amplifier

In addition to component size, other parameters should be taken into consideration. Table 1.2 shows an overview of some optical amplifiers. The EDFA is beneficial with regards to its high gain and low noise figure. However, it can be fairly costly, bulky, and is not able to operate in the O-band (1260-1360 nm). This could be a disadvantage because although EDFA can be deployed for WDM-PONs which use the C-band [Liu+16], a lot of transmissions for short distances use the O-band where chromatic dispersion is minimal.

Table 1.2 – Comparison of optical amplifiers [Hor+20]

Parameter	EDFA	RAMAN	SOA
Gain [dB]	> 40	> 30	> 30
Center wavelength [nm]	1530-1625	1280-1650	1280-1650
Bandwidth (3 dB) [nm]	30-60	up to 100	60
Max. saturation [dBm]	30	0.75× pump power	18
Polarization sensitivity	No	No	Yes/No
Noise figure [dB]	> 3.5	5	8
Pump power	25 dBm	> 30 dBm	< 400 mA
Excited state duration [s]	1×10^{-1}	1×10^{-14}	2×10^{-9}
Size	Rack-mounted	Bulk module	Compact
Switchable	No	No	Yes
Cost factor	Medium	High	Low

Raman amplifiers have the advantage of having a larger optical bandwidth and can have high saturation power. However, they require a substantial amount of pump power

and have a relatively high cost.

On the other hand, despite their lower saturation power and high noise figure, SOAs are low cost. They are also compact, can be used for optical switches/routers, and have a large optical bandwidth (e.g. 60 nm corresponding to approximately 7.59 THz in electrical bandwidth). Depending on the manufacturing SOAs can be polarization sensitive, but polarization independent SOAs are also commercially available.

For these reasons, SOAs are potential prospects for access networks as the possibility for overall reach increases and the requirements for low cost, simple physical layers are of utmost importance. Many researchers are currently exploring the implementation of SOAs as optical amplifiers for access/metropolitan networks [Tal+19] [Zha+20] [Bar+18b]. As was discussed in Section 1.3.3.3, SOAs do not have good linearity characteristics; improving the linearity of SOAs will be a key factor in this thesis work.

1.3.4 Receiver in photonics

A general representation of an optical receiver is shown in Figure 1.31. An optical

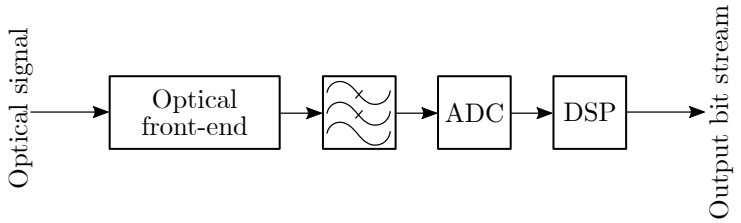


Figure 1.31 – Schematic of an optical receiver. DSP: digital signal processing, ADC: analog to digital converter

receiver can make use of either direct, heterodyne, or coherent detection. The considered modulation techniques at the emitter and detection technique at the receiver need to match, i.e. direct detection requires OOK or PAM, coherent detection would expect PSK or QAM [Kah06]. The cheapest and easiest to implement is the direct detection. Nevertheless, coherent detection has better performances in terms of adaptability (providing amplitude and phase profile to the DSP) and robustness to transmission impairments [Kah06]. The receiver's conventional architecture for a coherent detection is depicted in Figure 1.32. Note that the DSP implemented depends on the system operator.

The optical signal at the receiver goes through a polarization beam splitter which separates the horizontal and vertical polarization modes. Then a reference laser and

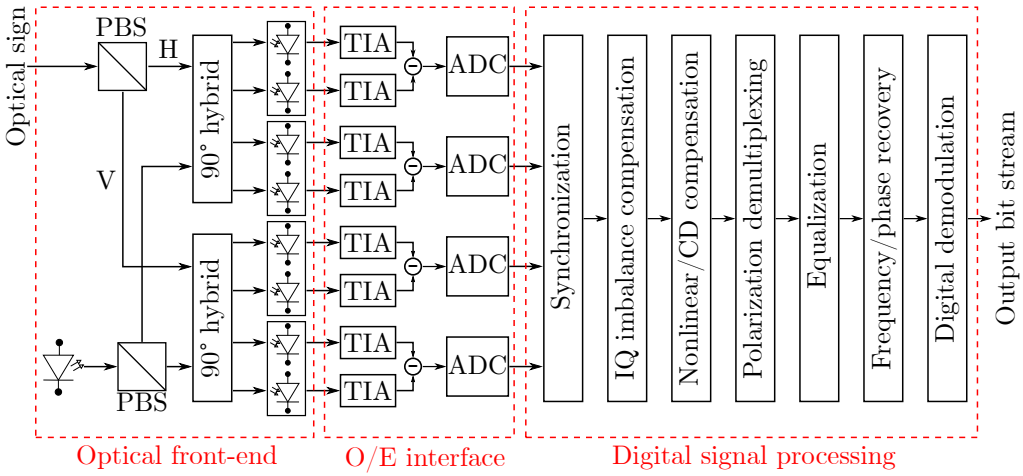


Figure 1.32 – Outline of an optical digital coherent receiver. PBS: polarization beam splitter, H: horizontal polarization, V: vertical polarization, TIA: transimpedance amplifiers, ADC: analog to digital conversion, CD: chromatic dispersion, O/E: optical/electrical

photodetectors are used to convert the received optical signal to a current. For the photodetectors, photodiodes can be used to convert the optical signal into an electrical current. The usual photodiodes are p-n , p-i-n, and avalanche photodiodes with p-i-n photodiodes having the highest level of stability. The conversion is performed following the below relation for p-i-n photodiodes

$$I_{pin} = \eta \frac{\lambda q}{hc} P_{in} = R P_{in} \tag{1.15}$$

with I_{pin} being the photodiode current generated from the incident light power P_{in}, η being the quantum efficiency, q the electron charge, $\frac{hc}{\lambda}$ the energy of the photon, and R the photodiode sensitivity.

Then, the transimpedance amplifiers (TIA) transform that current into a voltage that is then changed into the digital domain by the analog to digital conversion (ADC). Similar to the DAC, the ADC will also have a specific resolution or ENOB that will affect the quantization levels attributed to the signal. The ADC is also preceded by an anti-alias low-pass filter. The DSP then performs:

— Synchronization because the signal at the receiver might have experienced some delays. There are several ways to align the signal properly in terms of sample; a simple one is to perform cross-correlation [Bor+14]. In case of fractional delays, interpolation is usually required. For an OFDM signal, the synchronization

preamble is used then removed at this stage

— IQ imbalance compensation because the received I and Q signals will most likely not be orthogonal due to imperfections in the emitter, notably the IQ modulator not operating at the quadrature point, and the receiver. A common algorithm to compensate these imperfections is through orthonormalization [Fat+08]

— Nonlinear/chromatic dispersion compensation because the signal can be affected by chromatic dispersion in the fiber and nonlinearities in the channel. Chromatic dispersion mitigation can be done either in the time domain or the frequency domain. The advantage of using an OFDM based transmission is that it has high tolerance to chromatic dispersion [Shi+06]. Nonlinearities can not only come from the amplifier, but also from the fiber in high power transmissions. This thesis will focus on the nonlinearities caused by the amplifier. The ways to compensate nonlinearities will be discussed in Section 1.4

— Polarization demultiplexing to compensate residual chromatic dispersion impairments as well as polarization mode dispersion (PMD). To do so, the constant modulus algorithm (CMA) can be used [Kik08]

— After polarization demultiplexing, equalization can be used to counter static linear channel impairments as well as polarization mixing. In an OFDM setup, the equalization preamble is used then removed at this point

— Carrier frequency offset (CFO) and residual phase recovery since there will be some differences between the lasers at the transmitter and receiver. Two widely used methods to address these issues are the digital phase-locked loop (DPLL) to acquire the frequency offset [Zib+08] and *Viterbi* for the phase offset correction [Vit83]. In OFDM, the pilots are used for the phase correction

— Demodulation/detection of data symbols: hard decision uses minimum Euclidean distance

This section described the coherent optical receiver in a fairly general manner. A more detailed description of OFDM demodulation will be given in later sections.

1.3.5 Criteria for system evaluation

To evaluate the usefulness of the proposed schemes, different criteria will be put to use:

— The error vector magnitude (EVM) shown in Figure 1.33 gives the error between

the sent symbol and the received symbol. Its equation is

$$\text{EVM} = \sqrt{\frac{\sum_{k=1}^{N_{sc}} |s_k - r_k|^2}{\sum_{k=1}^{N} |s_k|^2}} \tag{1.16}$$

Figure 1.33 displays the error between a sent and received QAM symbol

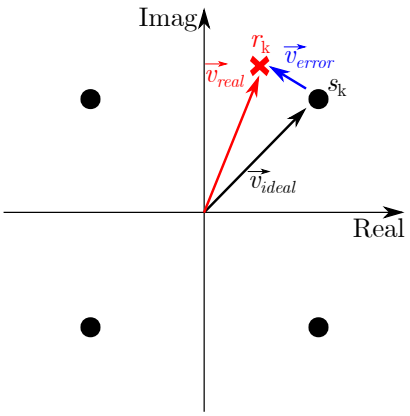


Figure 1.33 – Example of a 4-QAM constellation with error. s_k and r_k are the sent and received symbols

— The constellation diagram which gives a visual representation of the distance of the received symbols compared to the sent symbols. Figures 1.34a and 1.34b show constellation diagrams for 4-QAM and 16-QAM. The distortion in the constellation can give further information on the transmission. For example, in these figures the shape is due to noise and nonlinearities. A phase spread in the constellation could have indicated the presence of laser phase noise

— The BER which looks into the error between the sent and the received data stream, i.e.

$$\text{BER} = \frac{N_\epsilon}{N_{bits}} \tag{1.17}$$

where N_ϵ indicates the total number of erroneous received bits and N_{bits} is the total number of bits sent. Contrary to the EVM, the BER requires a substantial amount of data to be accurate. This is an issue for both numerical and experimental data which makes it a difficult metric to operate with. This thesis work will therefore verify the relationship between the BER and the EVM to justify prioritizing the EVM when presenting results

— The normalized mean square error (NMSE) gives the error between the sent and

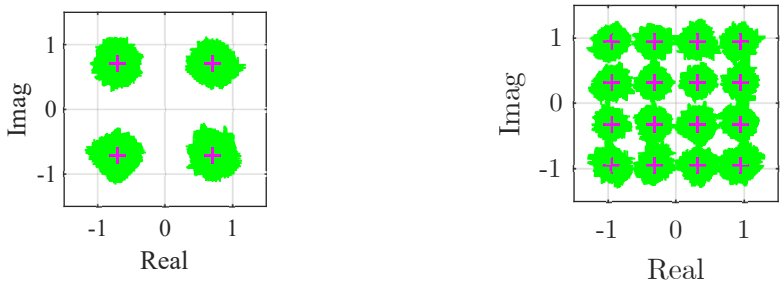


Figure 1.34 – Example of a constellation diagram for (a) a 4-QAM configuration (b) and a 16-QAM configuration

received time signals. NMSE is defined by

$$\mathrm{NMSE}_{dB} = 10\log_{10}\left(\frac{\sum_{n=1}^{N}|y[n]-x[n]|^2}{\sum_{n=1}^{N}|x[n]|^2}\right) \tag{1.18}$$

This criterion will be used to evaluate the validity of the linearization structures that will be employed throughout this thesis (more details will be given in the next sections). It will be useful when laying the groundwork for a complete optical transmission. However, once the fiber's chromatic dispersion will be taken into consideration, it will be more appropriate to focus on the other criteria.

— The power spectral density (PSD) of the OFDM signal at the emitter (Figure 1.35) and at the receiver which will show the in-band and out-band (for multi-user purposes) effects of digital predistortion (DPD) using a rectangular window of length T_{symb} (symbol period) with no overlap [Wel67].
Our interest will be focused on the changes observed in-band, but we will also take a look at the effects seen out-band when anticipating multi-user applications

— The OSNR which measures, as indicated by its name, the ratio between the signal and noise in an optical channel defined by the following:

$$\mathrm{OSNR}_{dB} = 10\log_{10}\frac{P_{out}}{P_{ASE}} \tag{1.19}$$

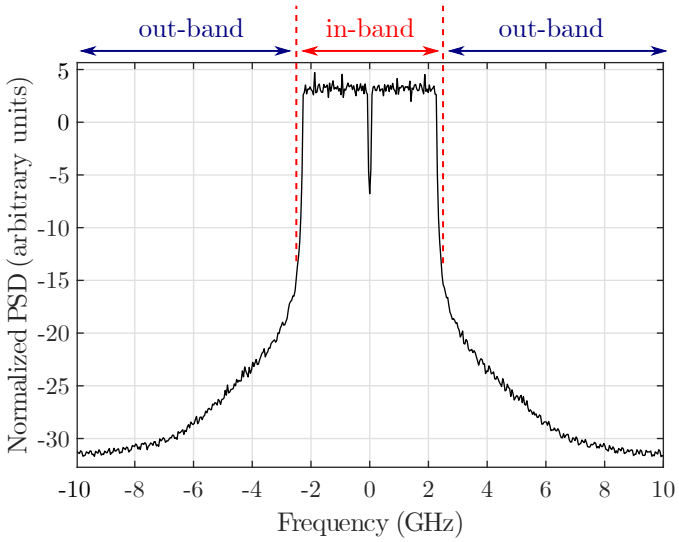


Figure 1.35 – Example of the power spectral density of an OFDM signal at the emitter with a 5 GHz bandwidth

1.4 Linearization schemes

The previous sections described the transmission system. In Section 1.3.3.3, the undesirable effects of the SOA were shown. It would be tempting to operate the amplifier in its linear region to avoid amplifier distortions; however, it would also entail that the SNR at the receiver will not be as high leading to stricter requirements in terms of receiver sensitivity and shorter reach. A means of mitigating the nonlinearities of the SOA therefore needs to be found.

There has been a plethora of linearization techniques in different domains and at different stages of the transmission (Figure 1.36).

The next paragraphs are dedicated to comparing these schemes. One important characteristic to look into will be the capacity of the linearization scheme to eliminate intermodal distortions. Figure 1.37 depicts a simple illustration of the signal spectrum after amplification. The output signal consists of the desired signal which contains all linear combinations of the signal. The intermodal distortions designate the nonlinear parts of the amplification. In Figure 1.37, the 2^{nd} (IMD2), 3^{rd} (IMD3), 4^{th} (IMD4), 5^{th} (IMD5), 6^{th} (IMD6), and 7^{th} (IMD7) order distortions are displayed. The goal here would be to remove these intermodal distortions and only keep the desired signal. This is done by generating IMDs that are equal in amplitude and 180° out of phase from the ones produced by the amplifier. Note that the general consensus is that even order IMDs can be

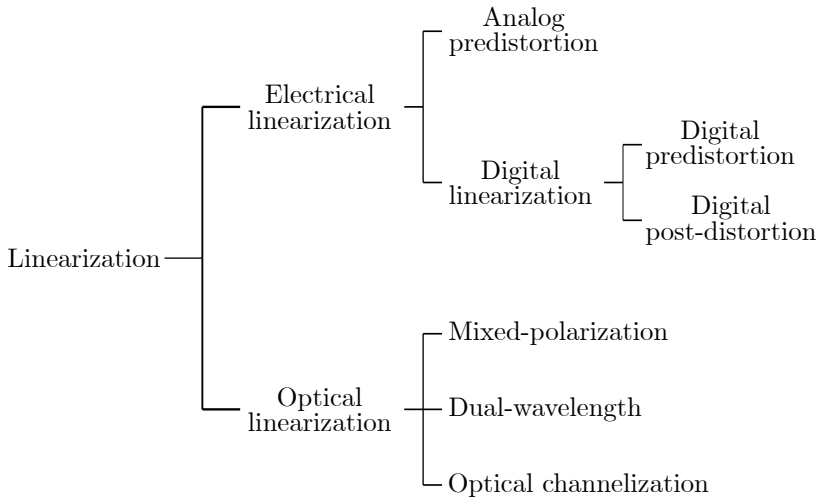


Figure 1.36 – Different linearization techniques [Zha+14]

naturally filtered out and thus only the odd order IMDs should be addressed.

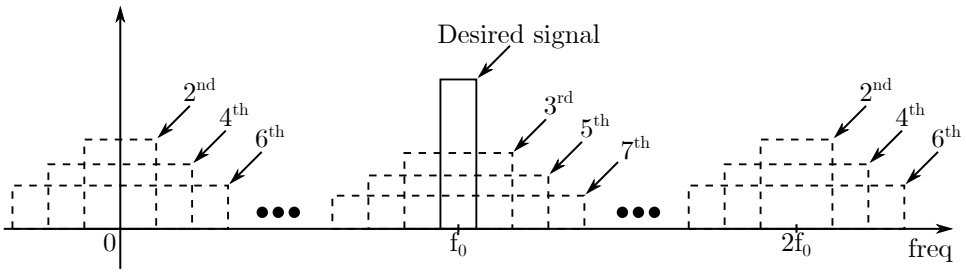


Figure 1.37 – Illustration of intermodulation modes showing 2^{nd} to 7^{th} order distortions

Linearizing the amplification process is not an easy fit. Section 1.3.3.3 clearly outlines the challenges that arise when an SOA is put in its nonlinear region including static and dynamic nonlinearities. Many factors change the behavior of the SOA thus will affect the effectiveness of the linearization method. The phase-amplitude coupling, SGM, SPM, XGM, XPM, and FWM effects will have to be considered. In the case of the CP-OFDM that will be studied, the FWM effects will be more prominent and will be affected by the frequency spacing and power per subcarrier of the CP-OFDM signal. Note that we will see in later chapters that many effects in the SOA are assymetrical frequency-wise which will add to the difficulty. The ASE will also be a factor that will influence the linearization as well as the small signal frequency response, the number of subcarriers for an CP-OFDM signal, the OSNR, and the data rate. A signal with a large bandwidth sending high data rates will bring forth memory effects. The change in the operating point of the SOA

can bring about memory effects as well and will be dependent on the bias current, the operating wavelength, and the input power of the signal. The non-constant envelope will be another influence on the matter. The next paragraphs will go over different methods used to mitigate these effects.

Before moving forward, one should note that the linearization schemes that will be described in this section do not directly alter the amplifier's function. However some methods that linearize the amplifier by changing the operating point exist. An example is envelope tracking (ET) which, as its name suggests, tracks the envelope of the signal and uses that information to change the bias current of an SOA, thus improving power efficiency and linearity [Ort+20]. These kinds of methods are quite dependent on time alignment between the input RF signal and envelope signal, and require an envelope tracker with a high bandwidth. The focus of the rest of this section will be on linearization methods that do not affect the amplification function.

Optical linearization can be done through:

— Mixed-polarization [Che+13] where a front polarizer, rear polarizer and modulator are used such that adjusting incident polarization angles suppresses nonlinearities
— Dual wavelength linearization [Zhu+14b] which exerts two wavelengths onto an external modulator such that nonlinear components generated at two wavelengths cancel out
— Optical channelization [Xie+12] where the transmitted optical signal is pre-distorted in such a way that the intermodulation components get canceled
— Hybrid polarization which combines mixed-polarization with dual wavelength linearization by connecting two or more modulators in cascade or parallel [Zhu+09]

A more detailed overview of optical linearization techniques can be found in [Sin+19]. Optical linearization has the advantage of having high bandwidth, but it has high complexity and has lower effectiveness in terms of linearization compared to digital linearization. It also has a higher cost than electrical analog predistortion.

Electrical analog predistortion circuits can be implemented to linearize a system [Zhu+13] [Zhu+14a] [Ros+17]. Circuits can be cascaded to get rid of intermodulation distortions (for example IMD3 and IMD5). Although these circuits are usually low in complexity, they introduce power losses in the system and are not as efficient and flexible as digital pre/post distortion.

Digital pre/post-distortion which can be implemented through DSP yields the best results in terms of linearization and has good efficiency [Zha+14] [Sin+21]. They can also

easily be adapted to changing modulation formats and parameters of the system contrary to electrical analog predistortion; a set of digital predistorters can be kept in memory and selected accordingly. This thesis work will look into their use while keeping an eye on low complexity solutions. Digital predistortion will be favored in this thesis work as digital post-distortion (DPoD) tends to require higher complexity since they have to expand their structure according to received clipped/saturated amplitude peaks on top of doing corrections on a noisy signal [Car+15]. In general, it is preferable to focus on optical transmitter compensation in order to launch a more desirable signal into the optic fiber. This is a concept that is applied by other researchers [Ami+15] [Ber+15]; DPD will thus be the focus of this thesis.

1.4.1 Digital predistortion techniques

The concept of predistortion is depicted in Figure 1.38. It can be first considered an inverse behavioral modeling problem (most cases being black-box models here).

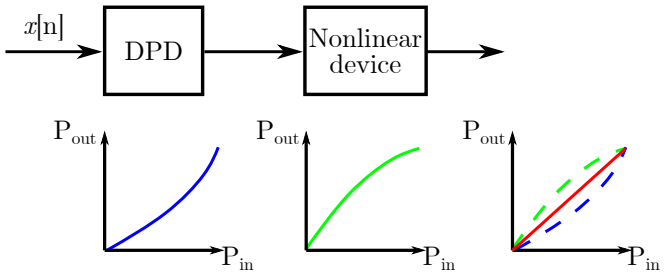


Figure 1.38 – Basic concept of predistortion. DPD: digital predistortion

Predistortion relies on a relatively basic principle. One may see the channel as a nonlinear function altering the modulated signal in an undesired fashion. It appears quite natural to try to implement the inverse nonlinear function and add it into the system so that the effects of the two blocks simply cancel each other out. While the concept comes across as being simple, many factors come into play and putting it into practice can prove quite complex. To be effective, DPD will have to properly identify the SOA nonlinear behavior described in Section 1.3.3.3. It means one would need to consider static and dynamic effects, the variations of SOA characteristics due to a change in its operating point or in the signal's input power. Additional difficulties can also come from the signal bandwidth with respect to the frequency response of FWM and from the added ASE.

In terms of identification two grand strategies are possible: direct learning and indirect learning [Gha+15]. Direct learning implies building a mathematical model for the channel and analytically computing its inverse function. The approach is generally cumbersome and is not relevant to the scope of this thesis. Indirect learning (shown in Figure 1.39) bypasses the analytical inversion requirement by essentially recasting the problem as classical optimization. One starts with a well-chosen nonlinear parametric structure and subsequently computes its coefficients by minimizing an error criterion. Obviously the choice of the predistorter structure is crucial and non-trivial. However, a lot of relevant knowledge is already available in fields such as control theory or behavioral modeling where nonlinear system identification is a well explored topic.

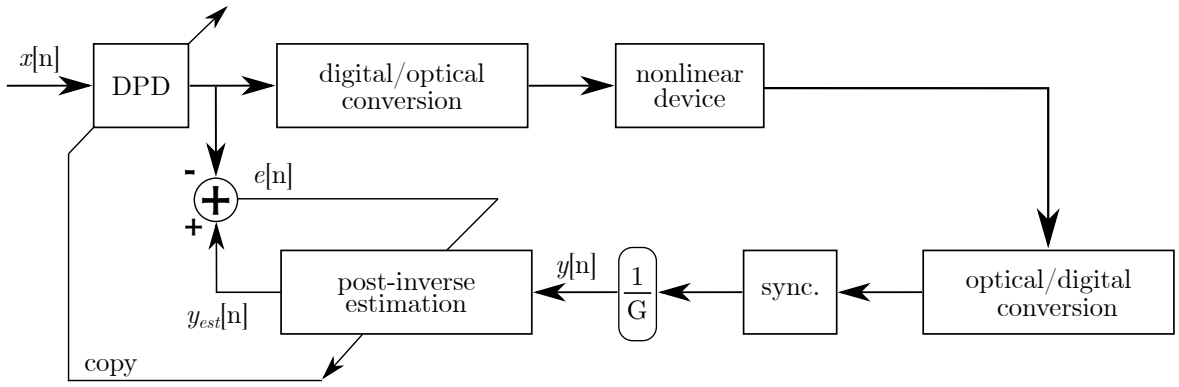


Figure 1.39 – Indirect learning approach. DPD: digital predistortion, sync.: synchronization

Some algorithms that are used as predistorters where an output sample only depends on the current input sample are the following

— Look-up table (LUT)

$$y_{LUT}[n] = H(|x[n]|)x[n] \tag{1.20}$$

where $x[n]$ and $y_{LUT}[n]$ are the input and output of the predistorter and H is the applied gain associated with the sample $x[n]$. Usually this structure is used without taking into account the statistical profile of the input signal. Research shows that constructing a LUT based on the PDF of the signal improves the benefits of this type of structures [You+20]

— Polynomial

$$y_{poly}[n] = x[n] \sum_{k=0}^{K-1} a_k |x[n]|^k \tag{1.21}$$

where $x[n]$ and $y_{poly}[n]$ are the input and output of the predistorter, and K is the order of the polynomial function. These types of predistorters are memoryless functions.

An issue that could be encountered by choosing to use the aforementioned predistorters is that it does not take into consideration the memory effects of the SOA. It seems more suitable to use algorithms that take account of this behavior. Some examples of dynamic nonlinear predistorters are neural networks (NN) and the Volterra series. The baseband equivalent of the Volterra series is denoted [Ben+79] [Sch+13b]

$$y_{volt}\left[n\right] = h_0 + \sum_{k \in \boldsymbol{I}_K} H_k\left[x\right]\left[n\right] \tag{1.22}$$

where $\boldsymbol{I}_K = \{1\,, 3\,, 5\,, \ldots, K\}$ is a set containing the orders of nonlinearity with K the maximum nonlinearity order. Note that only the odd terms are taken into consideration. The k-th order Volterra operator is

$$\begin{aligned} H_k\left[x\right]\left[n\right] = \sum_{m_1=0}^{M_k-1} \cdots \sum_{m_Q=m_{Q-1}}^{M_k-1} \sum_{m_{Q+1}=0}^{M_k-1} \cdots \sum_{m_k=m_{k-1}}^{M_k-1} \\ h_k\left(m_1\,, \ldots, m_k\right) \prod_{q=1}^{Q} x\left[n - m_q\right] \prod_{r=Q+1}^{k} x^*\left[n - m_r\right] \end{aligned} \tag{1.23}$$

where $Q = \frac{k+1}{2}$, $(.)^*$ is the conjugate operator, h_k is the Volterra kernel of order k and M_k is the corresponding memory depth.

The Volterra series are a generalized function that encompasses all aspects of a nonlinear system with memory effects which makes it too complex in the context of high data rate communications. One simplification of that model is the memory polynomials (MP) where only the diagonal terms of the Volterra series are deemed useful

$$y_{MP}[n] = \sum_{k=0}^{K-1} \sum_{l=0}^{L-1} a_{kl} x[n-l]|x[n-l]|^k \tag{1.24}$$

where K and M are respectively the model's nonlinearity order and memory depth.

Another algorithm called the envelope memory polynomials (EMP) that looks solely at the input's envelope will be of interest. Its input-output relation is

$$y_{EMP}\left[n\right] = x[n]\left(c_0 + \sum_{k=1}^{K-1} \sum_{l=0}^{L-1} a_{kl}|x[n-l]|^k\right) \tag{1.25}$$

Generalized memory polynomials (GMP) are a recent and elegant solution that includes MP and EMP while providing an extra degree of freedom. It should be noted that the generalization implied here is by no means a return to the Volterra series with its kernels that increase in dimension as the order of nonlinearity increases. This structure rather aims to take into account the signal's envelope, a lagging envelope and a leading envelope [Mor+06]. Its mathematical equation is

$$\begin{aligned} y_{GMP}[n] = &\sum_{k=0}^{K_a-1}\sum_{l=0}^{L_a-1} a_{kl}x[n-l]|x[n-l]|^k \\ &+\sum_{k=1}^{K_b}\sum_{l=0}^{L_b-1}\sum_{m=1}^{M_b} b_{klm}x[n-l]|x[n-l-m]|^k \\ &+\sum_{k=1}^{K_c}\sum_{l=0}^{L_c-1}\sum_{m=1}^{M_c} c_{klm}x[n-l]|x[n-l+m]|^k \end{aligned} \tag{1.26}$$

where K_j are nonlinearity orders, L_j and M_j are the memory depth, and a_{kl}, b_{klm} and c_{klm} are the coefficients to be estimated. The MP structure is obtained by solely taking the double sum term and the EMP structure is obtained by setting $L_a = 1$, $K_b = K_a - 1$, $L_b = 1$, $L_c = 1$, $K_c = 0$, $M_c = 0$. The evaluation of the relevance of this predistorter in the context of photonics will be one of the contributions of this thesis work.

Other predistorter structures are available in the literature such as Wiener, Hammerstein, and other combinations of predistortion blocks (in parallel or cascaded). A synthesis of predistorter structures is presented in [Gha+09].

All these polynomial based predistorters display linear equations with respect to their coefficients. These relations can be rewritten in vector form as

$$\boldsymbol{y} = \boldsymbol{\Phi a} \tag{1.27}$$

$\boldsymbol{a}$ is a column vector containing the model coefficients and $\boldsymbol{\Phi}$ is a $N \times N_{coeff}$ matrix constructed from the input signal according to the model's basis function. N is the length of the signal and N_{coeff} is the number of coefficients of the predistorter model.

The coefficient values can be estimated via the least squares (LS) algorithm and calculated using the Moore-Penrose pseudo inverse (denoted †) of $\boldsymbol{\Phi}$ as

$$\boldsymbol{a} = \boldsymbol{\Phi}^{\dagger}\boldsymbol{y} \tag{1.28}$$

Several pre-compensation schemes have already been explored in [Dio+17]. We will

venture into comparing some of the predistorters tested during previous works with the GMP in the photonics context. For practical reasons, we will not be implementing an adaptive algorithm that would allow to cope with slow variations of the SOA behavior (notably temperature variation and component aging). Indeed, with the need for a coherent receiver to estimate the predistorters (Figure 1.39) a closed loop implementation cannot be easily considered. A discussion of DPD identification algorithms is made in [Gil+20]. We will also evaluate the ability of DPD to remain effective in the face of variability in operating parameters through a sensitivity analysis in a later chapter.

1.5 Thesis objectives

The previous sections went over the challenges and advancements of optical transmission systems within the last years. In the context of access/metropolitan networks more specifically, the increasing demand for capacity, flexibility and longer reach while keeping simple, low cost physical layers presents its share of difficulties.

A typical optical transmission architecture is presented while keeping in mind these challenges. At the emitter, several modulation formats are described with an emphasis on OFDM. Indeed, OFDM is a format that is widely used in RF and has garnered a lot of attention from the photonics community over the years for its ability to be constructed with low complexity and allowing parallelization of DSP, its capacity for the dynamic allocation of frequency bands, its potential for high spectral efficiency, and its resilience towards channel dispersion. CP-OFDM is chosen as the modulation format to study during this thesis; however, other modulation formats stemming from this conventional OFDM can be substituted in later investigations.

It is shown that the fiber causes attenuation in the channel; the CP-OFDM signal will therefore have to be amplified to meet receiver sensitivity requirements and avoid worsening the quality of transmission. It makes sense to study the inclusion of SOAs as amplifiers in an optical communication system as there are substantial advantages exploring this cheap, integrable alternative in the case of metro/access networks. Furthermore, this component goes along the aim for higher data rates as it does provide a large optical bandwidth.

On the other hand, a coherent optical orthogonal frequency division multiplexing (CO-OFDM) configuration will allow us to take advantage of the added-value of OFDM modulation while putting ourselves in a "worst case" scenario in terms of nonlinearity for

the SOA. For operators, the findings and proposed solutions in this thesis can largely be applied to a majority of multicarrier modulations as well as high order single carrier modulations affected by non-constant envelopes.

As can be gathered from the previous statements, this thesis deals with the linearization of SOAs in CO-OFDM systems for metro/access networks. Some preliminary work has already been done in the Laboratoire des Sciences et Techniques de l'Information de la Communication et de la Connaissance (Lab-STICC) to show the benefits of predistortion. The added aim for this thesis is to solidify the robustness of that setup by simulating a more realistic CO-OFDM transmission, exploring new DPD structures, examining techniques to find the structural parameters of these predistorters, and performing both stochastic and experimental analyses. In addition to pre-compensation, it will be pertinent to examine post-compensation. Note that there will also be a compromise to be made between keeping linearity and needing a higher OSNR.

The performance metrics presented in Section 1.3.5 will serve to properly evaluate the benefits of predistortion. The DPD parameter complexity will also be looked into in a general sense since a more specific view on complexity dealing with field programmable gate arrays (FPGA) implementation needs careful computation beyond the scope of this thesis. However, since optical communication deals with high data rate, this criteria will have substantial importance in the evaluation of DPD performance.

1.5.1 Semiconductor optical amplifier model

To evaluate DPD in SOA based CO-OFDM systems, a co-simulation is performed between MATLAB and Keysight Advanced Design System (ADS). ADS houses the SOA model while the rest of the CO-OFDM transmission is coded on MATLAB. This means that the obtained numerical results of this thesis are contingent on the validity of the ADS SOA model. This section will serve to explain the working hypotheses that were used to construct that model and the comparison that were made with a commercial SOA.

The ADS SOA model is a baseband generalized time-domain transfer matrix model (TDTMM) whose implementation is done under an equivalent circuit. That model is designed to handle high bit rate modulations and relies on bi-directional propagation, an ASE propagation equation, and a carrier density evolution rate equation.

The SOA's active region is divided into several sections along the propagating optical signal axis. The carrier density is considered to be constant within one section, and the signal and ASE are assumed to travel along each section without being reflected. Note

that the uniformity of the carrier density in one section leads to the number of sections being dependent on pulse duration (a short pulse requires more sections). The shortest pulse that can be treated by this model is 1 ps since subpicosecond dynamics are not considered. In each section, the carrier density, optical signal field, and ASE intensity is determined.

The baseband optical signal field propagation equations are the following [Mor+08]

$$F_k^+(t) = \left[e^{-i\kappa^{(0)}_{N_k,\omega_j} L_s}\right]\left[e^{-i\kappa^{(1)}_{N_k,\omega_j} L_s \frac{d}{dt}}\right]\left[e^{-i\kappa^{(2)}_{N_k,\omega_j} \frac{L_s}{2}\frac{d^2}{dt^2}}\right] F_{k-1}^+(t) \tag{1.29}$$

$$F_{k-1}^-(t) = \left[e^{-i\kappa^{(0)}_{N_k,\omega_j} L_s}\right]\left[e^{-i\kappa^{(1)}_{N_k,\omega_j} L_s \frac{d}{dt}}\right]\left[e^{-i\kappa^{(2)}_{N_k,\omega_j} \frac{L_s}{2}\frac{d^2}{dt^2}}\right] F_k^-(t) \tag{1.30}$$

where $F_k^{\pm}$ is the envelope of the optical signal field co-propagating and counter-propagating along the active region in the k^{th} section, L_s is the length of one section, and $\kappa^{(m)}_{N_k,\omega_j} = \frac{d^m \kappa(N_k,\omega)}{d\omega^m}\Big|_{\omega_j}$ designates the m order derivative of the wavevector $\kappa(N_k,\omega)$ describing the action of the medium on the phase and amplitude of the electromagnetic waves. Its expressions can be decomposed into its Taylor expansion given in [Mor+08].

The wavevector depends on the optical signal pulsation ω_j and the carrier density N_k in the k^{th} section. The carrier density evolution is given by

$$\frac{dN_k}{dt} = \frac{I_{el,k}}{q_e L_s w d} - R_{sig,k}(N_k,\omega_j) - R_{ase,k}(N_k,\omega_j) - R_{nst,k}(N_k) \tag{1.31}$$

where $I_{el,k}$ is the electrical current flowing through the k^{th} section, $L_s wd$ the section's volume in the active region, q_e the elementary electrical charge, $R_{sig,k}(N_k,\omega_j)$ is the recombination rate due to the amplification of the signals, $R_{ase,k}(N_k,\omega_j)$ the recombination rate due to the amplified spontaneous emission, and $R_{nst,k}(N_k)$ the nonstimulated recombination rate. In this model, the stimulated recombination rates are considered to be proportional to light intensity.

The SOA has a fast gain dynamic due to interband and intraband effects. These effects are modeled by looking at the global compression term ϵ_j. The compressed material gain is defined by

$$g_{m,comp}(N,\lambda) = \frac{g_m(N,\lambda)}{1 + \sum_{j=1}^{M_{sig}} \epsilon_j S_j} \tag{1.32}$$

with M_{sig} being the number of signals propagating inside section k, $g_m(N,\lambda)$ being the material gain and S_j the photon density associated with each signal carried by wavelength

λ_j. The compression due to the ASE is neglected in this model since it is considered that the ASE does not have a high enough power spectral density to induce it in SOAs.

The ASE in this SOA model is considered as a saturating intensity thus the spontaneous emission is deemed constant in the time domain. The propagation equations for the ASE intensity are as follows

$$I^{+}_{ase,k} = e^{g_n(N_k,\lambda_{ase})L_s} I^{+}_{ase,k-1} + I_{sp}\left(N_k, \lambda_{ase}\right) \tag{1.33}$$

$$I^{-}_{ase,k-1} = e^{g_n(N_k,\lambda_{ase})L_s} I^{-}_{ase,k} + I_{sp}\left(N_k, \lambda_{ase}\right) \tag{1.34}$$

with $g_n\left(N_k, \lambda_{ase}\right)$ the net gain (difference between the material gain and the waveguide internal losses), and $I_{sp}\left(N_k, \lambda_{ase}\right)$ represents the generated spontaneous emission in one section.

The ADS SOA model has been successfully validated using numerical vs. experimental data obtained in the SOA's static and dynamic regimes [Mor+08], and by experimentally validating the model not only in the component level but also the system level of a CO-OFDM transmission [Kha+12]. It has also been experimentally validated for a reflective SOA in an intensity modulated OFDM configuration [Ham+15].

The ADS SOA model was validated for different types of SOAs [Mor06]. The one that will be used throughout this thesis work is based on a commercial bulk 750 μm long SOA with a 19 dB gain at 1510 nm with a bias current of 200 mA.

1.5.2 Implemented CO-OFDM setup

The CO-OFDM setup implemented in this thesis is based on a MATLAB/ADS co-simulation with the SOA being modeled on ADS and the rest of the communication chain being implemented on MATLAB.

Figure 1.40 depicts the entire implemented CO-OFDM transmission whose implementation and usage has evolved throughout this thesis. A bit stream is modulated into symbols that are then used to construct the payload (associated with the sent symbols s_k) and the preamble in parallel. The payload designates the data being carried and the preamble serves to aid in synchronizing and equalizing. Note that although the payload and preamble are built in parallel for simplicity, in practice it would be done progressively in real-time.

Once the DPD has been chosen and identified (the identification of the predistorter

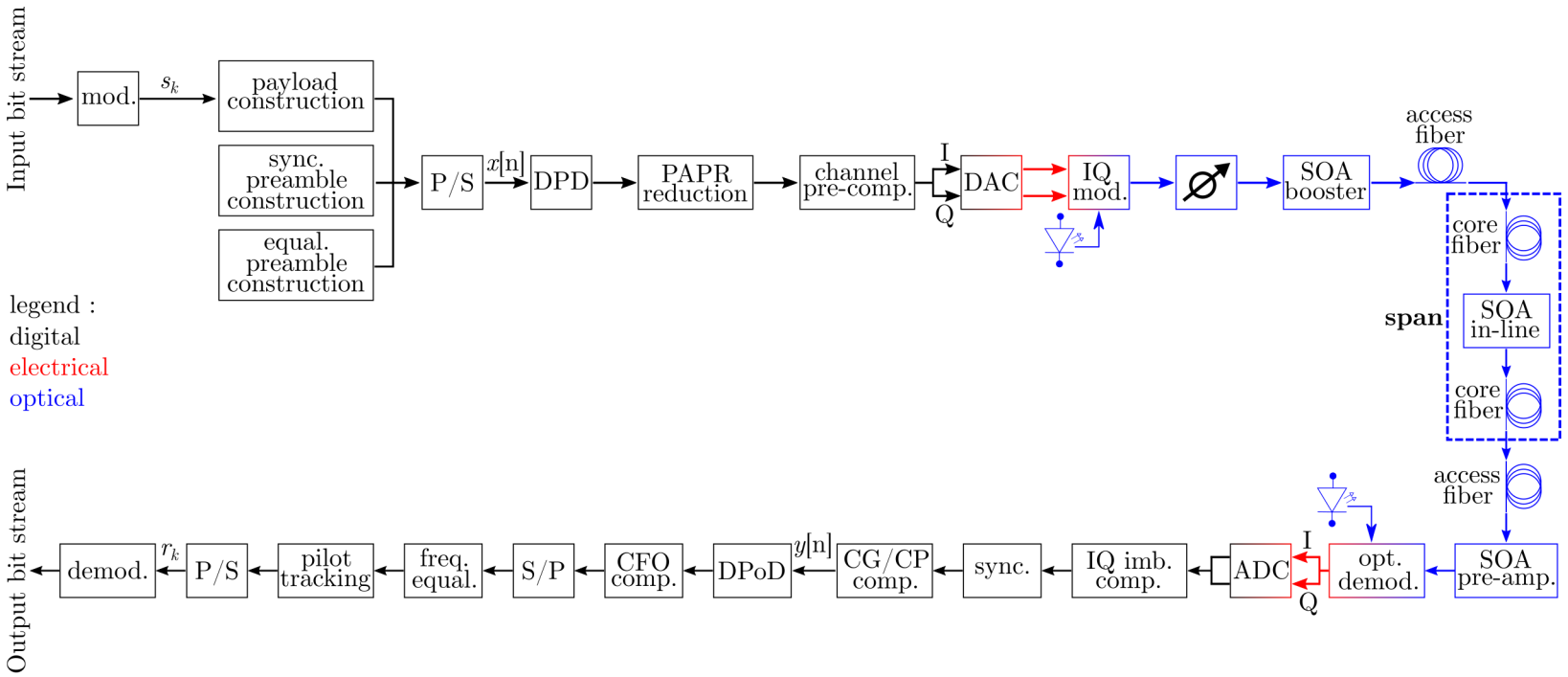


Figure 1.40 – Complete implemented CO-OFDM transmission. Mod: modulation, equal.: equalization, sync.: synchronization, P/S: parallel to serial, DPD: digital predistortion, PAPR: peak to average power ratio, comp.: compensation, DAC: digital to analog conversion, opt.: optical, att.: attenuation, SOA: semiconductor optical amplifier, demod.: demodulation, ADC: analog to digital conversion, imb.: imbalance, CG/CP: common gain/common phase, DPoD: digital post-distortion, CFO: carrier frequency offset, S/P: serial to parallel, freq.: frequency. s_k and r_k are the sent and received QAM symbols, and $x[n]$ and $y[n]$ are the time signals used to estimate the predistorter

will be explained later), a validation phase can be performed. During the validation phase of DPD, the time signal $x[n]$ is predistorted before going through PAPR reduction. As the focus of this thesis is on the DPD aspect, hard-clipping was chosen as a simple PAPR reduction technique. However, other more elaborate schemes have been shown to work well with DPD [Bej+15].

The predistorter is usually identified in an optical back to back (B2B) configuration meaning without the presence of the optical channel. The predistorter only compensates emitter imperfections; a complementary block can be used to pre-compensate channel imperfections such as chromatic dispersion. This will be discussed in more detail in later sections.

Then the signal undergoes DAC before being converted into the optical domain using the IQ modulator described in Section 1.3.1. The optical signal goes into a booster SOA whose input power is controlled by an optical attenuator. The characterization and modeling of the simulated INPHENIX-IPSAD1501 SOA has been discussed in Section 1.5.1. This same model will be used throughout this thesis work. As stated in earlier sections, DPD's purpose in this thesis is to mitigate the emitter impairments, most notably the SOA's nonlinearities.

The amplified signal passes through one span consisting of several kilometers of fiber and an in-line amplifier to counter fiber attenuation. At the receiver, the signal is pre-amplified to meet the receiver sensitivity threshold when it gets converted back to the electrical domain, and subsequently, the digital domain.

Several DSP algorithms are brought into effect at the receiver. First an IQ imbalance compensation is applied followed by timing synchronization, and common phase and gain compensation. The common phase and gain are corrected using a complex gain that we estimate as such

$$\zeta_{est} = \mathbf{x}_{eq}^{\dagger}\mathbf{y}_{eq} \tag{1.35}$$

where $\mathbf{x}_{eq}$ is the sent time equalization preamble and $\mathbf{y}_{eq}$ is the received time equalization preamble. During the identification phase of DPD, the signal denoted $y[n]$ can be extracted to identify the predistorter which can then be applied at the emitter during the validation phase. That process has been described in section 1.4 where the predistorter's coefficients are estimated through least squares using the signal at the emitter $x[n]$ and the signal at the receiver $y[n]$. Another linearization scheme that will be explored in the thesis is DPoD which intends to mitigate nonlinearities at the receiver instead of the emitter. Its attributes with respect to DPD will be looked into further in section 2.5.

Following DPoD is the carrier frequency offset (CFO) compensation. Then comes the frequency equalization, and the phase and amplitude tracking. The recovered symbols r_k which will be used for EVM calculations are ultimately demodulated. The resulting bit sequence is compared with the input bit sequence to compute the BER.

So far, the setup that was described is the one used to obtain numerical results. Getting the experimental data is quite simple in our case. The same implemented MATLAB setup is used for the emitter and receiver DSP. The digital signal is sent to the experimental setup, goes through the experimental channel, then is sent back to MATLAB for the digital processing. Chapter 4 will dive deeper into how this is achieved.

Figure 1.40 represents the integral CO-OFDM transmission whose implementation and usage has evolved throughout this thesis. When presenting numerical and experimental results, the running setup will be shown where the deactivated blocks will be crossed out, the ideal blocks will be grayed out, and any differing algorithms will be pointed out.

Chapter 2

Digital predistortion for SOA based CO-OFDM systems: simulation results

This chapter goes over the implementation of a coherent optical orthogonal frequency division multiplexing (CO-OFDM) transmission in the context of this thesis and serves to present the obtained simulation results. First, a link between error vector magnitude (EVM) and bit error rate (BER) will be established in the context of this thesis. Second, the choice of clipping will be presented. Third, the implementation of digital predistortion (DPD) will be described with an emphasis on finding a way to automatically obtain the predistorters' structural parameters. Finally, the results for a booster semiconductor optical amplifier (SOA) then an in-line SOA will shown, putting particular interest in the generalized memory polynomials (GMP) predistorter. Figure 2.1 displays the adjusted setup of the CO-OFDM transmission introduced in Section 1.5.2; this setup will be used for the simulation results of Chapter 2.

For the emitter digital signal processing (DSP), everything is performed apart from the chromatic dispersion compensation. After the orthogonal frequency division multiplexing (OFDM) modulation (described in Section 1.3.1.1), the DPD (based on if it is the identification or validation phase), and the peak to average power ratio (PAPR) reduction (hard-clipping) ensues the digital to analog conversion (DAC). For now the DAC and analog to digital conversion (ADC) are considered ideal meaning the quantization error will be defined only by the computer in use. For the electro-optic and opto-electronic conversions, the lasers are also considered ideal, thus there will be no laser phase noise. The scenario here is optical back to back (B2B); after the booster SOA, the signal goes directly into the optical front-end which is considered ideal as well. The photodiode sensitivity and efficiency will therefore not be a factor to take into account. The receiver DSP includes the synchronization which is a simple cross-correlation between the received

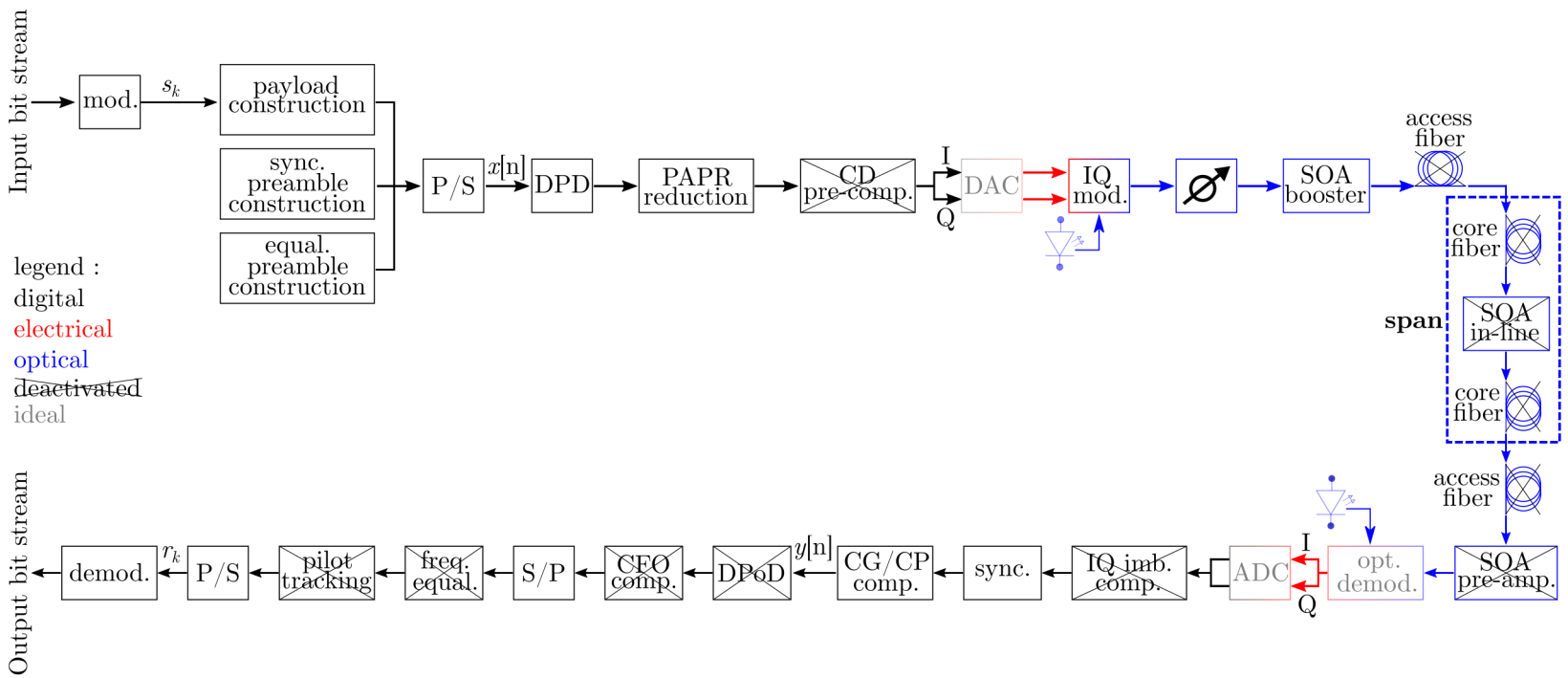


Figure 2.1 – Booster SOA implemented CO-OFDM transmission. Mod: modulation, equal.: equalization, sync.: synchronization, P/S: parallel to serial, DPD: digital predistortion, PAPR: peak to average power ratio, CD: chromatic dispersion, comp.: compensation, DAC: digital to analog conversion, opt.: optical, att.: attenuation, SOA: semiconductor optical amplifier, demod.: demodulation, ADC: analog to digital conversion, imb.: imbalance, CG/CP: common gain/common phase, DPoD: digital post-distortion, CFO: carrier frequency offset, S/P: serial to parallel, freq.: frequency. s_k and r_k are the sent and received QAM symbols, and $x[n]$ and $y[n]$ are the time signals used to estimate the predistorter

signal and the OFDM synchronization preamble, the common gain/phase compensation (Equation 1.35), and the quadrature amplitude modulation (QAM) symbols demodulation (using hard decision). The demodulated symbols can then be used to obtain the BER.

2.1 Linking BER and EVM

The BER is usually looked into to assess the quality of transmission in photonics, but its computation especially in the case of high quality transmission can be demanding. The EVM has been established as a faster alternative which can be related to the BER in the context of a system limited by optical additive white Gaussian noise (AWGN) [Sch+11]. This thesis work intends to directly use the EVM metric throughout in the context of a system limited by optical nonlinearities instead. It therefore seems wise to inquire about the EVM's relationship with BER still holding even when straying away from an optical AWGN channel. Several simulations are run using the optical B2B scenario described earlier. The input power into the SOA is changed from -17 dBm to -10 dBm to obtain different values of EVM/BER. Figure 2.2 attempts to compare the 4-QAM approximated BER vs. EVM of [Sch+11] to the BER and EVM obtained by simulating the nonlinear channel used in this thesis.

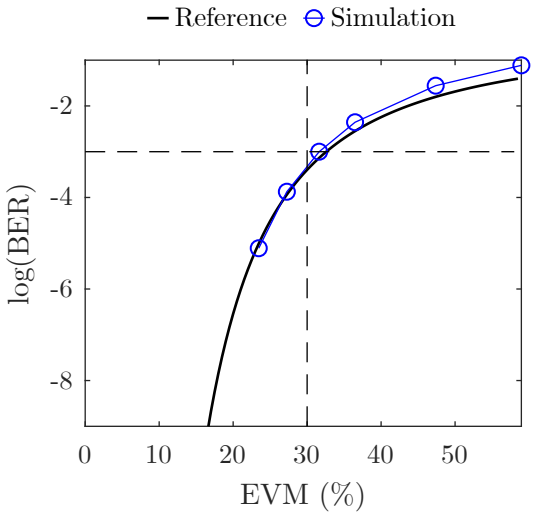


Figure 2.2 – $\log_{10}$(BER) vs. EVM. The blue curve is created by calculating the error between 1,221,120 bits and the dashed lines outline the 30% EVM limit corresponding to a BER of 10^{-3}

First, one can see that the EVM values shown in Figure 2.2 are rather high compared to what is customary for radio frequency (RF). With higher data rates and bandwidth, and a cleaner transmission medium, optical communications in general see higher EVM values than in RF [Idl+17]. As the EVM decreases, the simulated curve converges into the reference curve. This makes sense since when the EVM is lower, the SOA is not as nonlinear. In that case, the scenario does starts to resemble the system affected by optical AWGN. On the same figure, it is shown that although, the values in EVM tend to be slightly higher in our case, using the value of 30% as the limit for a BER of 10^{-3} is more than reasonable. Note that the chosen BER value of 10^{-3} is a pessimistic outlook as future transmissions expect a BER of 10^{-2} before forward error correction (FEC) [ITU20b]. With these results, it is believed that the EVM will be a meaningful metric to use in place of the BER.

2.2 Choosing the clipping level for PAPR reduction

Before going into the simulation results that will be presented in the next section, it is important to first clarify the PAPR reduction block. As the focus of this thesis is not on PAPR reduction, the simplest technique is executed: hard-clipping. However, a plethora of PAPR techniques are available in the literature [Rah+13]. Hard-clipping is defined as

$$y_c[n] = \begin{cases} Ae^{i\angle x[n]}, & \text{if } |x[n]| > A \\ x[n], & \text{otherwise} \end{cases} \tag{2.1}$$

where $\angle$ symbolizes the phase, and A, the clipping level, is related to the clipping ratio γ (given in dB) like so

$$A = 10^{\frac{\gamma}{20}}\sqrt{P_{avg}} \tag{2.2}$$

where P_{avg} is the average power of the signal $x[n]$. P_{avg} can be estimated using the OFDM spectral configuration such that $P_{avg} \approx \frac{N_{sc}-N_{null}}{N_{sc}\rho_{os}}$ where N_{sc} is the number of subcarriers per OFDM symbol, N_{null} represents the null subcarriers, and ρ_{os} is the oversampling factor. Following the 802.11 standard for the configuration of an OFDM frame, Equation 2.2 suggests that as the number of subcarriers increases, so does the clipping level.

As depicted in Figure 2.3, hard-clipping introduces undesirable effects as it creates more and more spectral regrowth as the clipping ratio decreases. It also leads to undesirable effects in-band as it does increase EVM in an ideal transmission. However, when looking at a transmission with the SOA, it can also lead to favorable conditions.

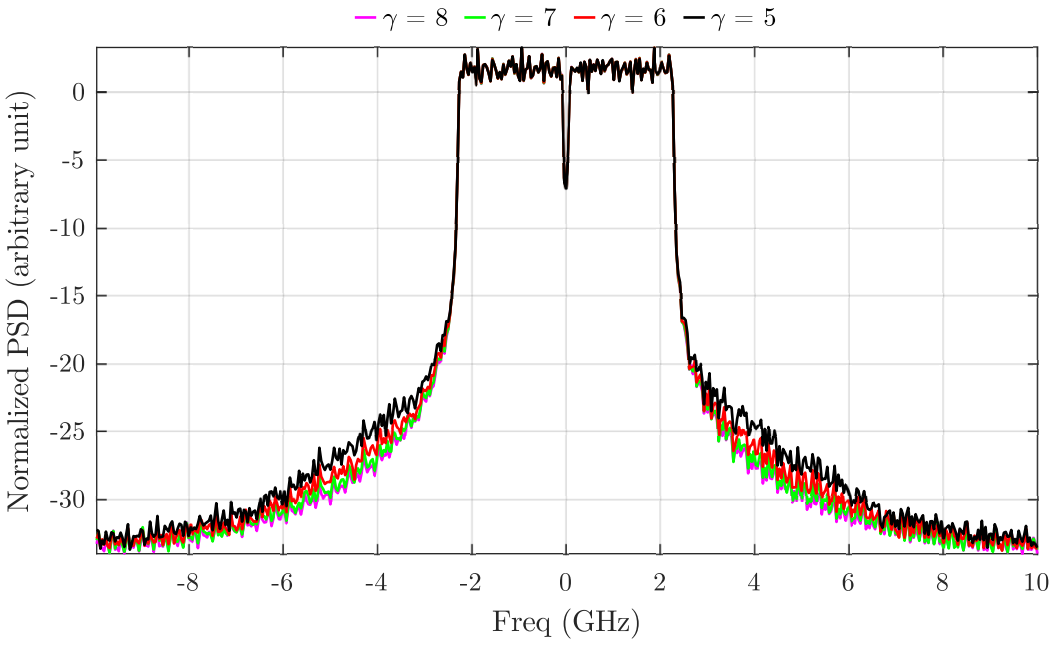


Figure 2.3 – Normalized power spectral density (PSD) of a sent 4-QAM OFDM signal for different clipping ratios γ. The signal has 128 subcarriers, a bandwidth of 5 GHz, and an oversampling factor of 4

If we look at Figure 2.4a and Figure 2.4b representing the mean EVM and the standard deviation σ of that EVM from 10 different simulation runs at different levels of clipping, it can be seen that when the SOA is operating in its nonlinear region, a lower clipping ratio leads to a lower EVM and lower standard deviation of that EVM. This observation does translate to BER as well since it is seen in [Azo+14] that clipping can produce lower BER.

Clipping the signal is also effective for the DAC. High amplitude samples degrade quantization noise as they induce larger DAC quantization levels. In [Ami+15], it is shown that at lower DAC resolutions, the benefits of clipping become more apparent.

After having observed the advantages and disadvantages of clipping, it seems clear that a trade-off needs to be evaluated. In our case, we chose to pick a clipping ratio that is high enough that we can work with high powers but also at least leaves the synchronization preamble unaltered. Figure 2.5a shows 2000 samples of an OFDM signal at the emitter and Figure 2.5b shows the probability density function (PDF) of an OFDM signal computed from 256 OFDM symbols with the parameters as in Figure 2.5a. Different clipping ratios are shown.

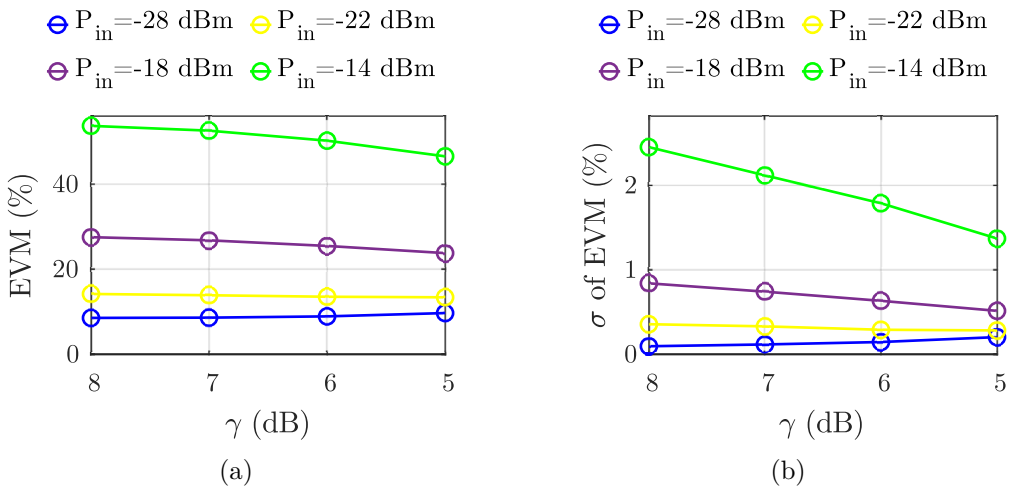
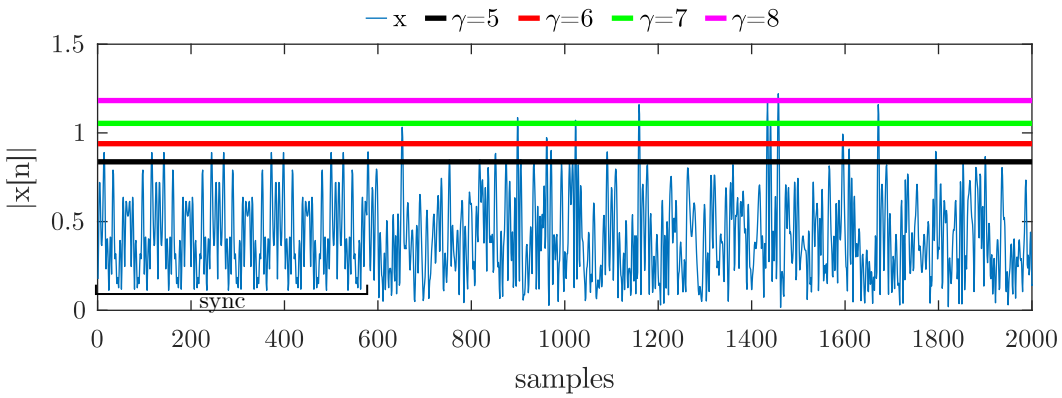


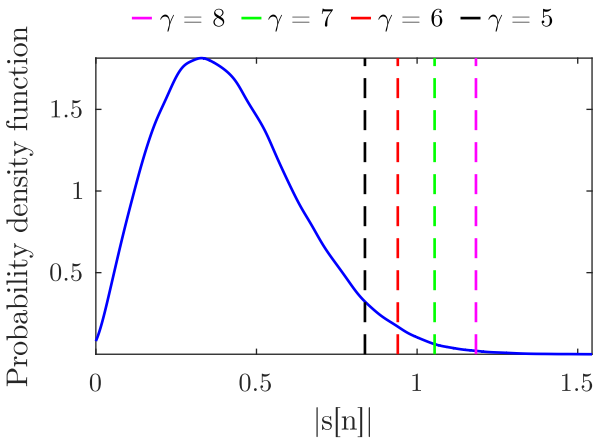
Figure 2.4 – (a) Average EVM and (b) standard deviation σ of the EVM taken from 10 different simulation runs vs. clipping ratio γ at different input powers into the SOA

Considering Figure 2.5a, a clipping ratio of 6 dB seems to be a good trade-off since it does not clip the synchronization preamble (marked "sync" in Figure 2.5a). According to Figure 2.5b, a clipping ratio of 6 dB corresponds to clipping roughly less than 5.45% of the signal. Note that this means the Rayleigh distribution of OFDM is disturbed since there will be a peak at the clipping level and no samples above that level.

PAPR reduction schemes can help limit but are not effective in mitigating the non-linearities from an SOA and Mach-Zehnder modulator (MZM) [Ami+15]. DPD therefore needs to be used to take care of these nonlinear effects, especially since we want to work at high powers.



(a)



(b)

Figure 2.5 – (a) $|x[n]|$ (2000 samples) and (b) PDF of a sent 4-QAM OFDM signal and different clipping ratios γ. The signal has 128 subcarriers, a bandwidth of 5 GHz, and an oversampling factor of 4

2.3 Choices of DPD implementation

In Section 1.4.1 the concept of DPD was described. Among the widely known predistorter structures that were presented, three will be of particular interest to us: the memory polynomials (MP), the envelope memory polynomials (EMP), and the GMP. This section will go over how DPD is implemented using these 3 structures.

2.3.1 Finding the predistorters' structural parameters

DPD is done through two phases: the identification/learning phase and the validation phase. The learning phase where the predistorter's coefficients are estimated need to be

done before the predistorter can be applied during the validation phase.

The learning phase is performed using the indirect learning approach in an optical B2B setup (Figure 1.39) where the DPD block at the emitter is bypassed since the predistorter has not been identified yet. At the receiver, after synchronization and common gain/phase compensation, the coefficients can be estimated through least squares using the received signal and the sent signal (1.28). The number of coefficients and thus the complexity of the predistorter is determined by the nonlinearity order and the memory depth. Using the MP structure in Equation 1.24 as an example, the number of coefficients is given by $M \cdot K$.

Optical communication systems work with high bandwidth and high data rates. In these conditions, the requirements for low computational costs, low latency need to be rather strict especially since the complexity of DPD usually increases with high bandwidth. It is therefore imperative to limit the potential parametric complexity of predistorters which in our case means we will limit the nonlinearity order and memory depth. It is important to note that although complexity is equated to parametric complexity in this thesis for simplicity, the complexity of a real-time implemented predistorter depends on other factors like the operations needed per sample [Teh+10].

In previous works at the Laboratoire des Sciences et Techniques de l'Information de la Communication et de la Connaissance (Lab-STICC), an optimal predistorter was chosen by doing a sweep of the nonlinearity order and memory depth and finding the set of parameters that yielded the lowest EVM values over a range of different input powers into the SOA [You19]. To explain further with the MP as an example, if the maximum nonlinearity order is limited to $K_{max} = 8$ and the maximum memory depth to $M_{max} = 8$, finding the proper predistorter would require testing 64 combinations. In total, at least 128 simulation runs would be needed: 1 run to identify the predistorter and 1 run to obtain the EVM value when applying the predistorter. Since the MP and EMP only have two parameters to set, it does not seem completely impractical to do such comprehensive analysis for these two predistorter structures.

When dealing with the GMP the size of the problem becomes daunting; venturing into a grid search with $8^8 = 16\,777\,216$ combinations (when limiting all 8 parameters to a maximum value of 8) should preferably be ruled out for any practical purpose. Indeed, performing these sets of simulations would take roughly 173 hours in our case.

One simple solution to this problem is given by the so-called Hill-Climbing algorithm which presents an automated means of selecting the structural parameters [Wan+16]. The

general stages of the Hill-Climbing algorithm used to choose the eight GMP structural parameters K_a, L_a, K_b, L_b, M_b, K_c, L_c, and M_c are described in Figure 2.6. The GMP structure shown in Equation 1.26 is reminded below

$$\begin{aligned} y_{GMP}[n] = &\sum_{k=0}^{K_a-1}\sum_{l=0}^{L_a-1} a_{kl}x[n-l]|x[n-l]|^k \\ &+\sum_{k=1}^{K_b}\sum_{l=0}^{L_b-1}\sum_{m=1}^{M_b} b_{klm}x[n-l]|x[n-l-m]|^k \\ &+\sum_{k=1}^{K_c}\sum_{l=0}^{L_c-1}\sum_{m=1}^{M_c} c_{klm}x[n-l]|x[n-l+m]|^k \end{aligned}$$

First, the 8-tuple is initialized as a linear function with no memory. A linear function with no memory means that: $L_a = 1$, $K_a = 1$, $L_b = 1$, $K_b = 0$, $M_b = 1$, $L_c = 1$, $K_c = 0$, and $M_c = 1$. Then the neighborhood of that current point is searched. The neighborhood is defined as the $(-1, 0, +1)$ range of each parameter and consists of a maximum of $3^8 = 6\,561$ points. The predistorter coefficients are estimated for each point belonging to the neighborhood as well as the current point. The cost function to minimize is defined as

$$J = \lambda \,\mathrm{NMSE} + N_{coeff} \tag{2.3}$$

where λ serves to put more or less weight on the normalized mean square error (NMSE) with respect to N_{coeff}, the number of coefficients. The idea behind this criterion J is to find a trade-off between the predistorter's modeling accuracy and its parametric complexity. Keep in mind that J is a unitless criterion which is minimized according to a fixed λ value. If a point belonging to the neighborhood has the smallest J, then that point now becomes the new current point and the algorithm goes back to exploring the neighborhood. If on the contrary, the current point has the minimum J, then the algorithm stops and the structural parameters associated to that point are picked. In general, the number of iterations is limited to a certain value (10 in our case), but the solution is usually obtained before the algorithm reaches that limit. As stated previously, a limit is also put on the maximum value allowed for a parameter.

As it might be difficult to visualize what this looks like, an illustration of that algorithm for a 2-dimensional predistorter structure is shown in Figure 2.7. Starting from the leftmost side, the predistorter is initialized as a linear function with no memory. The current point is considered the 1$^{\text{st}}$ current point. The point with the lowest J at that

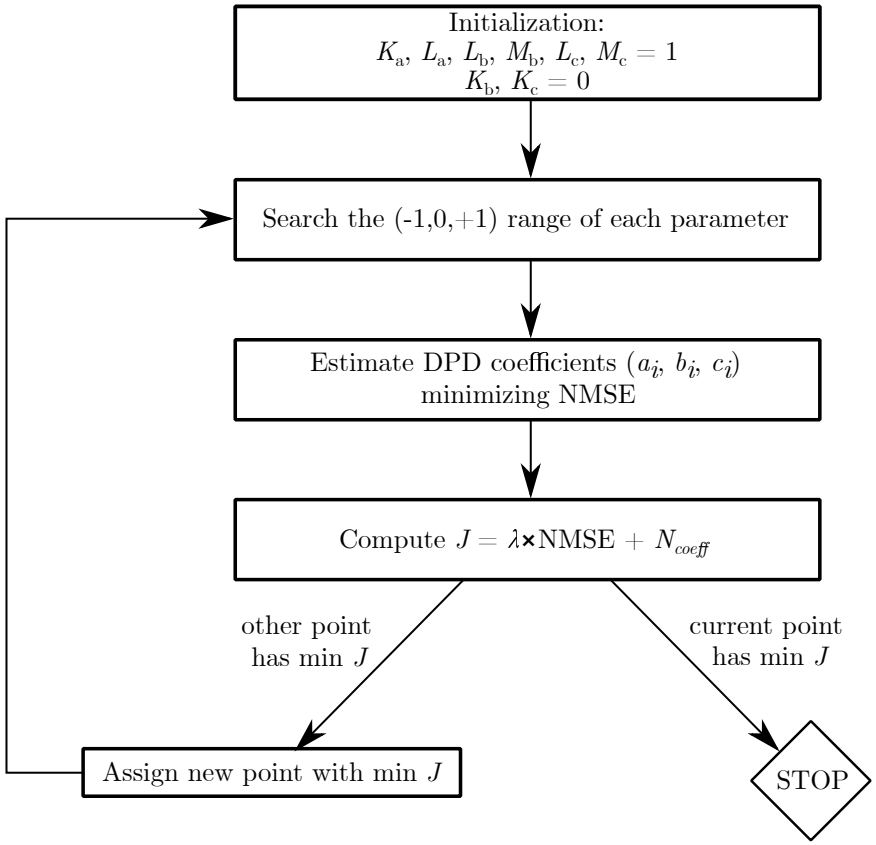


Figure 2.6 – Hill-Climbing algorithm

stage is the point where the memory depth and the nonlinearity order are both of value 1 which becomes the new current point (in Figure 2.7 it is shown as the 2$^{\text{nd}}$ current point). Here the minimum J is found once again in the neighborhood when the memory depth and nonlinearity order are both of value 2. In the last stage, the current point has the smallest J, thus a memory depth of 2 and nonlinearity order of 2 are chosen as the structural parameters.

Though using the Hill-Climbing algorithm reduces the time spent finding the structural parameters, it would prove useful to test out how optimal it is compared to a full grid search. Figure 2.8 displays the criterion J computed for all explored points with different imposed coefficient limits at $P_{ref} = -12$ dBm into the SOA for $\lambda = 26$ and $\lambda = 4$. For example, for a maximum number of coefficients limited to 60, 3 187 551 points were explored and therefore 3 187 551 J values are depicted at the number of coefficients $N_{coeff} = 60$. Note that this number is different from the 16 777 216 combinations that were mentioned earlier because a limit is put on the maximum number of coefficients a

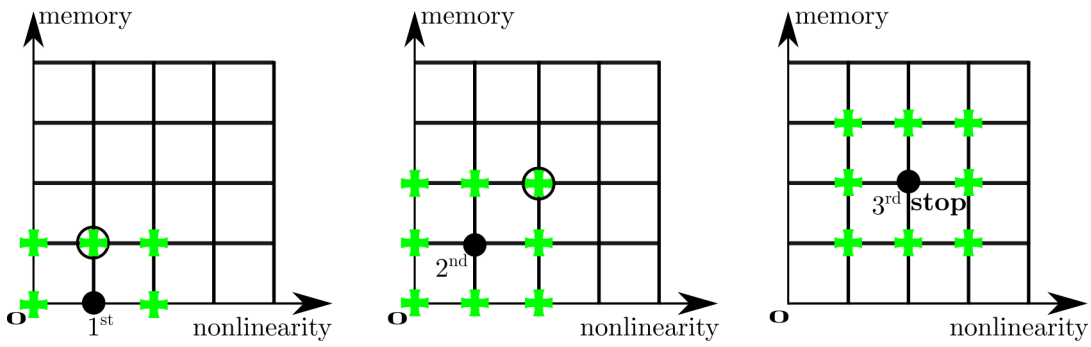


Figure 2.7 – Hill-Climbing algorithm for a 2-dimensional predistorter structure. The black dot is the current point, the green crosses are the neighbors, and the black circle is the best point

predistorter is allowed to have. Figure 2.8 shows that the lowest J is found when the number of coefficients is at 33 and the structural parameters are the following for $\lambda = 26$: $L_a = 3$, $K_a = 3$, $L_b = 1$, $K_b = 3$, $M_b = 8$, $L_c = 1$, $K_c = 0$, and $M_c = 1$. For the case with $\lambda = 4$, the cost function is minimized when the number of coefficients is at 16 and the structural parameters are the following: $L_a = 2$, $K_a = 2$, $L_b = 1$, $K_b = 2$, $M_b = 6$, $L_c = 1$, $K_c = 0$, and $M_c = 1$. The parameters all had a maximum value limited to 8 in these cases.

Using the same limit value of 8 on the parameters, but instead going through the Hill-Climbing algorithm yields the exact same predistorter which means that the Hill-Climbing is able to find the best J successfully without needing to do a full grid search. It is nice to point out the fact that compared to a grid search with a limit of 60 coefficients for instance, the Hill-Climbing algorithm would significantly reduce computation times. The full grid search requires about 33 hours (computing 3 187 551 predistorters) while the Hill-Climbing only requires 4 minutes (computing 6 366 predistorters for $\lambda = 4$) or 8 minutes (computing 12 323 predistorters for $\lambda = 26$) to find the structural parameters.

So far, the Hill-Climbing heuristic has been compared with a full grid search for the identification phase of DPD. The next step is to compare Hill-Climbing and a full grid search at the validation phase. As was stated earlier, the full grid search is done by computing multiple predistorters for different maximum number of coefficients allowed to a predistorter; this yields Figure 2.8. For the validation phase, the best predistorter (lowest J) is selected for each number of coefficients, i.e. 60 predistorters are picked. Then, these 60 predistorters are applied at the emitter and a value for the EVM is obtained. The predistorters obtained from Hill-Climbing are applied at the emitter as

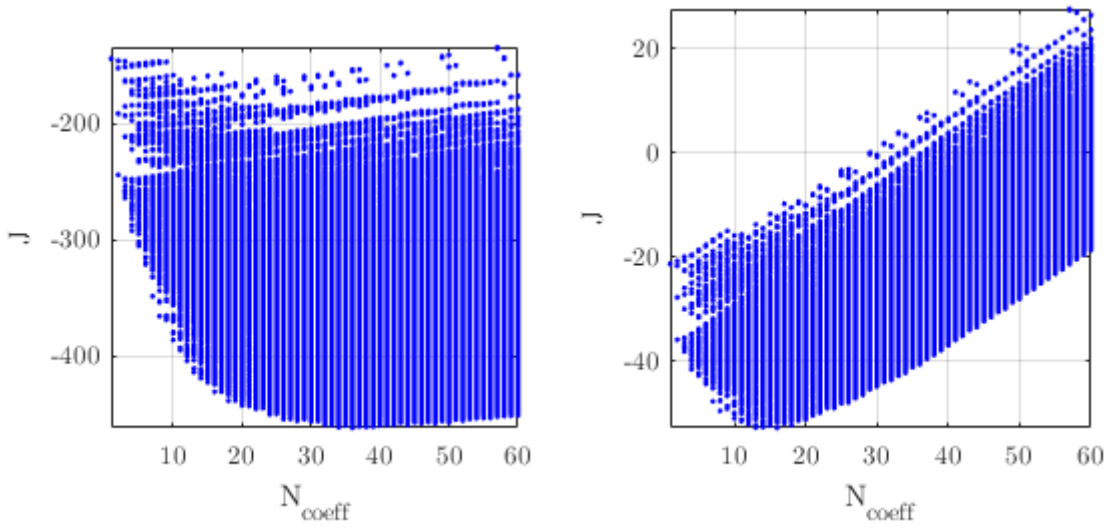


Figure 2.8 – (a) Full grid search of GMP structural parameters at $P_{ref} = -12$ dBm with (a) $\lambda = 26$ and (b) $\lambda = 4$. J is the criterion to minimize and N_{coeff} the number of coefficients

well. Figure 2.9 shows the EVM obtained during the validation phase of DPD using the selected predistorters from the full grid search at different input powers into the SOA. The black markers indicate the values obtained from Hill-Climbing for $\lambda = 4$ and $\lambda = 26$.

Since the predistorter is identified at the receiver then applied at the transmitter in a nonlinear system and the estimation is done on a signal with limited duration, it makes sense to observe some differences between the error during the learning phase and that of the validation phase. In Figure 2.9, it can be seen that for $P_{ref} = -14$ dBm and $P_{ref} = -12$ dBm, an increasing number of coefficients allows to improve predistortion until a certain point of convergence. Beyond that point, an increase in the number of coefficients does not lead to better results. For $P_{ref} = -14$ dBm and $P_{ref} = -12$ dBm, $\lambda = 26$ gives a more complex predistorter for a similar EVM value. The $P_{ref} = -10$ dBm case shows that with a higher input power of the SOA, the minimum EVM is reached with rather few coefficients. In fact, increasing the number of coefficients seem to be more detrimental to the transmission. The NMSE at the identification phase which becomes lower as the number of coefficients increases does not necessarily give the best EVM during the validation phase. Overall, it seems preferable to pick a λ small enough as to limit the

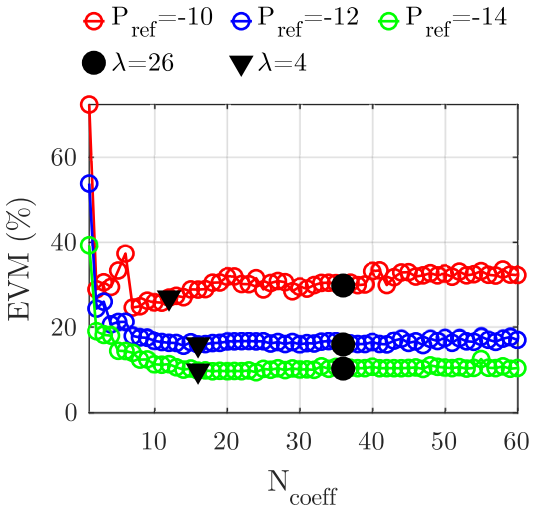


Figure 2.9 – GMP at the validation phase for different P_{ref}. Best EVM is the minimum EVM for each P_{ref}, and $\lambda = 4$ and $\lambda = 26$ are the EVM obtained using predistorters found with Hill-Climbing

number of coefficients of the predistorter, especially at high input powers of the SOA. This proves that the choice of λ is not trivial. Research has been done to automatically find a desirable value for the weight coefficient λ [Wan+18] though during this thesis work, the value for λ is simply picked in such a way to limit a predistorter's number of coefficients.

So far, the predistorter coefficients have been estimated using least squares in the time domain. Using a quadratic error based on the EVM to calculate the predistorter coefficients could be of value. Indeed, it is a way to work with smaller matrices since the predistorter learning phase would require the received data symbols r_k instead of the received samples $y[n]$. Additionally, the SOA is part of a much bigger system with the conversion from the digital to the optical domain meaning we could be faced with delays. If the synchronization is not done properly, the impact on the NMSE can be significant whereas with the EVM a predistorter could be able to adjust more easily to a rotation. An analysis is done using the EVM as a base to estimate the predistorter coefficients as well as to find the structural parameters. The formula for the new cost function becomes

$$J_{EVM} = \lambda \, \mathrm{EVM} + N_{coeff} \tag{2.4}$$

Drawing a comparison between the Hill-Climbing using the NMSE and the Hill-Climbing using the EVM as the error criterion leads us to generate Figure 2.10 where the EVM at

the receiver is looked at for different input powers into the SOA. Even though it makes sense to use the EVM instead to estimate the coefficients and as the error criterion, it seems like it does not produce better performances. The distortions experienced by the signal are in the time domain therefore an identification process done in the time domain seems logical.

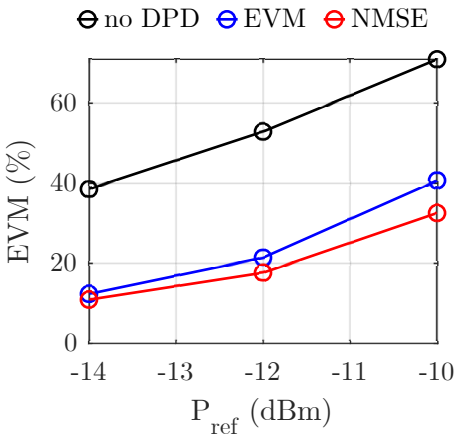


Figure 2.10 – Hill-Climbing using different error criteria

Since using the NMSE, gives slightly better performances and it does not require converting the time signal into data symbols to obtain the EVM, it seems more appropriate to choose it as the error criterion for Hill-Climbing.

All in all, the Hill-Climbing heuristic is proven to be an appropriate algorithm to rapidly find the structural parameters of predistorters. There has also been research on evaluating its robustness and advantages by using other search criteria, controlling the search path and neighborhood definition of the algorithm [Wan+18]. Being able to use the error criterion J allows us to keep an eye on parametric complexity. Another way to limit the parametric complexity of preidstorters is to only consider odd order of nonlinearities. This point will be discussed in the next section.

2.3.2 Even orders in digital predistortion

In the literature, the polynomial structures can be simplified by only considering odd order nonlinearities [Mor+06] [Jia16]. The assumption that only odd orders have influence is compatible with the fact that analytically, all even order frequency mixing falls in the out-band region (Figure 1.37) meaning they can simply be filtered out at the receiver

while for odd orders some can be filtered out but the rest falls in the in-band region requiring DPD [Yan+19] [You19]. In previous works performed by the Lab-STICC, we chose to follow the consensus of only considering odd nonlinearity orders to lower the complexity of predistorters.

However, other studies show that even orders can help increase the accuracy of predistortion [Din+04] [Che+16]. We therefore decided to test out the inference that even orders can be ignored for DPD by running some simulations. Figure 2.11 displays the EVM obtained at different input powers into the SOA for a GMP found with Hill-Climbing when taking into account the even orders. Compared to the odd orders only, it is seen that even orders do add value to the predistorter notably at high input powers. For the case where $P_{ref} = -10$ dBm, with 8 added coefficients to the predistorter, an improvement of 8.3% in EVM is found. Seeing the added value of even orders through simulation, we decided to include them in all predistorters.

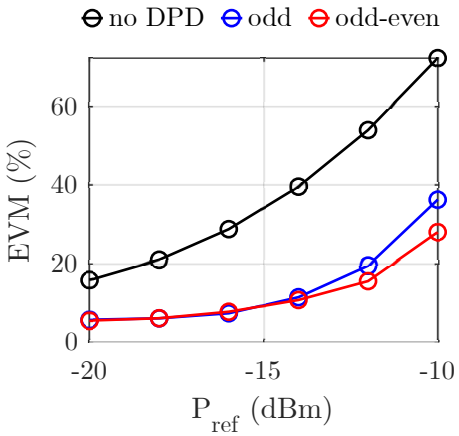


Figure 2.11 – Influence of even orders on the GMP. Black: no DPD, blue: GMP with only odd orders, red: GMP with odd-even orders

2.4 Digital predistortion for a booster SOA

The results shown in this section are based on the setup portrayed in Figure 2.1. These results serve to evaluate the performance of the GMP predistorter and compare it to the MP and EMP for photonics purposes. Their input/output relationships are given in Equations 1.24, 1.25, and 1.26.

2.4.1 Numerical results for a booster SOA

In this comparative analysis, the purpose of the MP, EMP, and GMP is to mitigate the effects of, most importantly, the SOA, but also the MZM which has a static nonlinear characteristic (Equation 1.11). The parameters shown in Table 2.1 were used to perform the comparative analysis.

Table 2.1 – Parameters used for DPD evaluation (booster SOA)

Parameter	Symbol	Value	Units
Modulation order	M	4	-
Number of subcarriers	N_{sc}	128	-
Bandwidth	ΔB	5	GHz
Cyclic prefix factor	ρ_{cp}	$\frac{1}{8}$	-
Total number of QAM data symbols sent	N_{data}	3180	-
Oversampling factor	ρ_{os}	4	-
Clipping ratio	γ	6	dB
MZM peak-to-peak voltage	V_{pp}	$1.25 \times V_{\pi}$	volts
Laser wavelength	λ_{LO}	1540	nm
SOA bias current	I_{bias}	150	mA

The objective of the first sets of results is to evaluate the efficiency of the three predistorters. First, for each input powers into the SOA and each structure (MP, EMP, GMP) a predistorter is identified using Hill-Climbing to find the structural parameters. The same frame is used to perform the learning phase. Then ensues the validation phase where the predistorters are applied at the emitter on several other OFDM frames. The same frames are used for all cases (no DPD, MP, EMP, GMP). The EVM is computed for each considered input power into the SOA and the results are plotted in Figure 2.12a. As seen in Section 2.1, for 4-QAM OFDM in optical fiber access networks, EVM values in the 30% range are considered acceptable [Sch+11]. Network reach is thus directly related to the highest input power that still guarantees the 30% EVM boundary. For each scenario, 10 simulations are carried out featuring different OFDM sequences consisting of 1 synchronization preamble, 1 equalization preamble, and 30 OFDM data symbols corresponding to 3 180 4-QAM data symbols or 6 360 data bits.

The different plots in Figure 2.12a show average EVM (dark thin lines) and standard deviation (shading). The predistorter coefficients are recomputed for each input power level. One notices that while MP does provide some improvement, EMP and GMP predistorters perform much better with GMP leading particularly at lower power levels. One

can assume that the GMP converging to an EMP at high powers shows a higher influence of the signal's envelope in causing nonlinearities with the SOA. The analysis is repeated using NMSE (Figure 2.12b); the ranking of the various predistorters is confirmed. It is interesting to observe that the standard deviation is reduced when DPD is included in the system (note thinner shading for blue, red and green plots). This entails that the system is more robust with respect to the characteristics of the input signal.

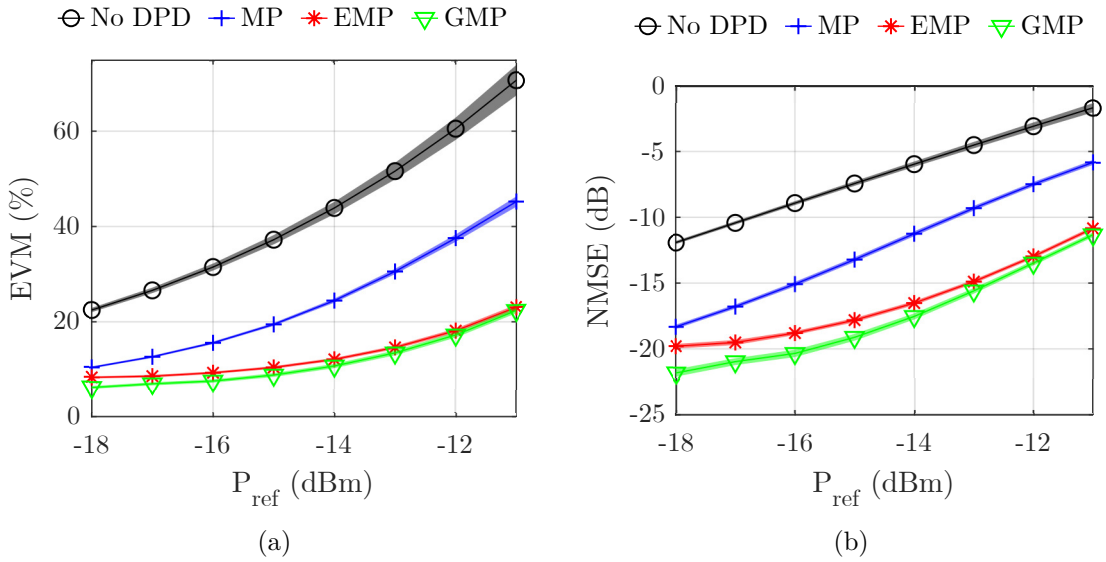


Figure 2.12 – (a) EVM and (b) NMSE for different predistorters at different P_{ref}

Figure 2.13 shows the 4-QAM constellation on the receiver side (in an optical back-to-back scenario) for an input power into the SOA $P_{ref} = -14$ dBm. Qualitatively the ranking of the three DPD solutions in terms of performance is confirmed. Visually the effect is even more apparent than in the previous figures.

Figure 2.14 shows OFDM spectra for the same four scenarios at the same reference power of -14 dBm. Note that frequency values are relative to the 1540 nm optical carrier. Also, since the plot is meant for comparison of SOA input and output, the magnitude scale is normalized hence the values are not physically representative. Observation of the in-band behavior shows that the DPD acts as a pre-equalizer (the tilt visible on the purple waveform is compensated). Equally interesting is the out-band behavior. The input of the SOA (black curve) shows some spectral regrowth due to clipping as well as the IQ modulator nonlinearity. One may also see that adjacent channel power ratio (ACPR) is visibly reduced once DPD is applied. In this respect GMP performs best (particularly

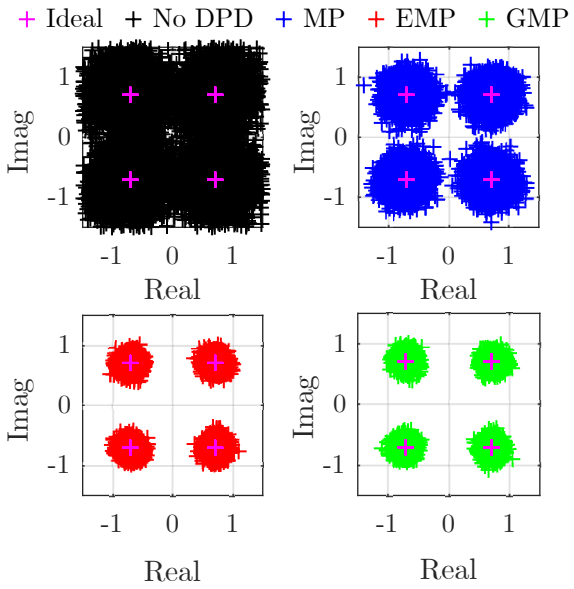


Figure 2.13 – Constellation diagram for different predistorters at $P_{ref} = -14$ dBm

visible on the left-hand side).

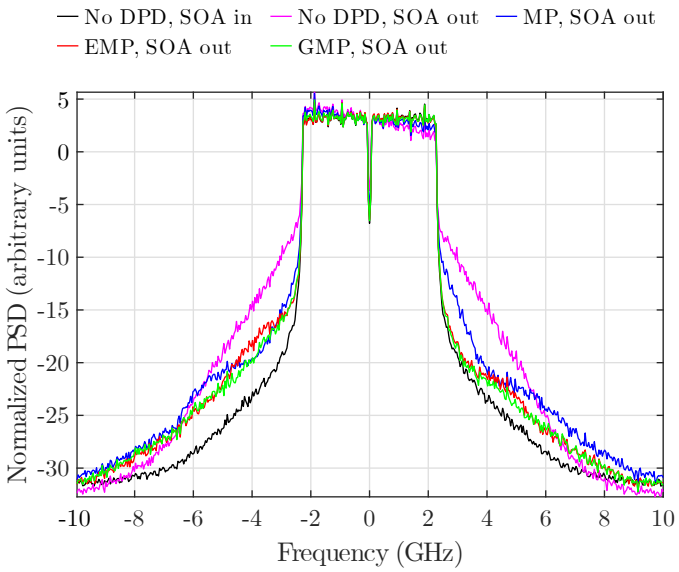


Figure 2.14 – PSD for different predistorters at $P_{ref} = -14$ dBm

The hypothesis when doing the efficiency analysis of the predistorters is that the physical layer is capable of constantly and accurately monitoring the input power into the SOA and is thus able to choose the appropriate predistorter depending on the input

power. In the case where the physical layer does not allow for such complexity, it seems important to evaluate the robustness of DPD when the input power into the SOA strays from the identification power. This analysis is carried out and illustrated in Figure 2.15 where this time the parameters of the three predistorters are computed only once for a reference power of −14 dBm. They are then used unaltered for system validation at various other input power levels in what is essentially a robustness analysis. EMP and GMP clearly outperform MP, with GMP now exhibiting a more robust behavior at higher power levels. This might mean that the phase of the input signal becomes more influential when the input power is higher than the reference power. The number of coefficients of the three different predistorters at $P_{ref} = -14$ dBm are the following: MP 6, EMP 15, and GMP 16. The GMP thus adds little parametric complexity for better performances. Also note that the gain in performance from DPD decreases rapidly when the input power is lower where we see that at $P_{in} = -18$ dBm, the no DPD case gives the same EVM values as the EMP and GMP cases. At this point, the input power into the SOA is quite far from P_{ref} and the SOA is much more linear at $P_{in} = -18$ dBm than at $P_{ref} = -14$ dBm. A wise decision would then be to either change the predistorter's coefficients or turn the predistorter off.

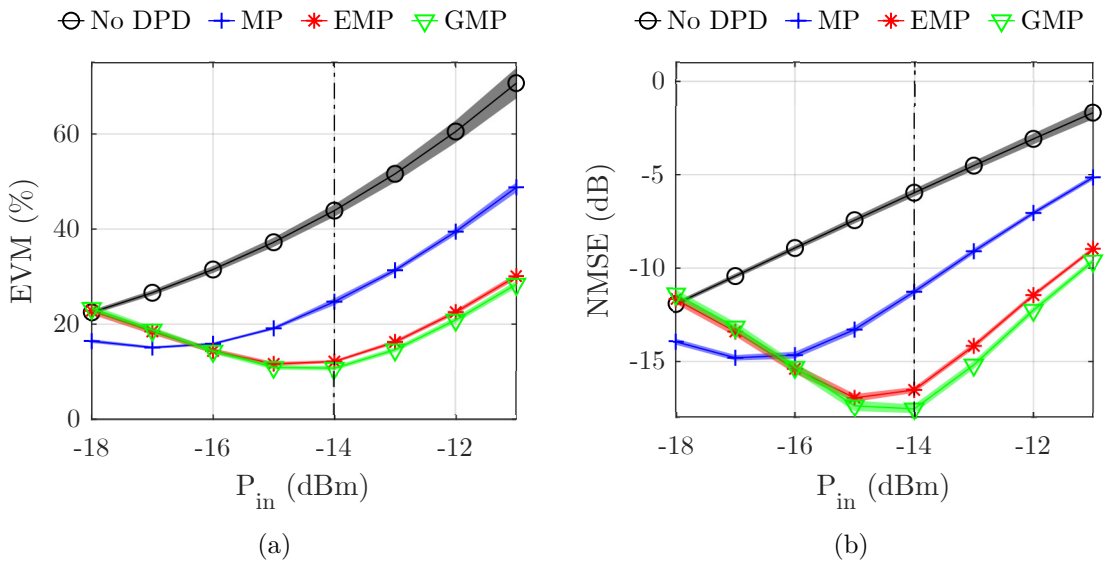


Figure 2.15 – (a) EVM and (b) NMSE for different predistorters with $P_{ref} = -14$ dBm and at different P_{in}

One might wonder if the MP, EMP, and GMP (Equations 1.24, 1.25, and 1.26) char-

acterize the SOA's nonlinear behavior well enough since they are predistorters with simplified structures. The value of these predistorters is tested against a Volterra predistorter which gives a more complete depiction of the SOA nonlinear behavior. We implemented a discrete time formulation of the baseband Volterra structure (Equation 1.22) which allowed us to generate Table 2.2. This table shows the performance and complexity of all predistorters. Here, we can gather that the complexity of Volterra is quite high for its subpar performance compared to that of GMP for instance. According to Figure 2.9, increasing the number of coefficients is sometimes counter productive in terms of quality. High order nonlinearities are difficult to capture with a very high sensitivity to noise , signal variation, and with a limited oversampling factor ($\rho_{os} = 4$). Also, long term effects like large memory depths are difficult to identify especially with relatively short frames. It makes sense in our case to stick to the simplified versions of Volterra as they have good performances and much lower parametric complexity.

Table 2.2 – Performance and complexity of different predistorters

	N_{coeff}	EVM (%)
No DPD	N/A	39.6
MP	6	24.2
EMP	15	11.8
GMP	16	10.3
Volterra	258	11.1

Since the GMP tends to give slightly better results than the EMP and the MP while staying low in terms of complexity, we choose to conduct the rest of the analyses by focusing on this predistorter structure.

2.4.2 Generalized memory polynomials for different bandwidths, number of subcarriers, and modulation order

As data rate is an important aspect of a transmission, we choose to venture into an analysis of DPD in the context of higher data rates. Looking at the parameters influencing data rate (Equation 1.3), we can choose to raise the data rate by increasing the number of subcarriers N_{sc}, the order of modulation M or the bandwidth of the OFDM signal ΔB. Although it might not be intuitive to increase the data rate by increasing the number of subcarriers, if one follows the 802.11 OFDM convention, when going from 128 to 256 subcarriers, the ratio of data subcarriers vs. the total number of subcarriers

slightly increases ($\frac{108}{128} = 0.84$ to $\frac{234}{256} = 0.91$). Additionally, for a constant OFDM symbol bandwidth ΔB, the symbol duration T_{symb} will be longer if the number of subcarriers is higher. It means that if the length of the cyclic prefix is kept constant when increasing the number of subcarriers, the cyclic prefix factor ρ_{cp} can be decreased from $\frac{1}{8}$ to $\frac{1}{16}$ which can further increase data rate.

When doing this analysis, it could prove useful to investigate the effect of these parameters (N_{sc}, ΔB and M) on the EVM. Note that increasing the bandwidth ΔB in our case has the side effect of also raising the sampling rate which is not always possible in practice. It can also possibly increase the memory depth required by the predistorter.

The first step in this analysis is to look at the behavior of the SOA when no DPD is involved. The base signal that is used for this analysis has the following parameters: ΔB=5 GHz, $N_{sc} = 128$, and $M = 4$. When changing one parameter, the others are those of the base signal. That is, if the bandwidth goes from 5 GHz to 10 GHz, the number of subcarriers is 128 and the modulation order is 4. Figure 2.16a shows the EVM at different input powers into the SOA for different bandwidths. It can be seen that as the bandwidth increases, the EVM drops for higher input powers. This makes sense since with a larger bandwidth, the spacing between subcarriers increases; the four wave mixing (FWM) effects are progressively reduced as soon as the subcarriers are spaced by more than 1 GHz which is the case for a greater number of subcarriers for larger bandwidths. On the other hand, when going into lower input powers into the SOA, the EVM rises with an increasing bandwidth. Indeed, when operating at the linear region of the SOA, the effect of amplified spontaneous emission (ASE) will be more apparent; a larger bandwidth equates to more noise.

Figure 2.16b shows the EVM at different input powers into the SOA for different orders of QAM. The requirements for 4-QAM and 16-QAM are quite different. A higher order modulation in the optical communication system requires a higher optical signal to noise ratio (OSNR) at the receiver, more complex DSP to keep the same BER, and higher DAC resolutions [Idl+17]. Seeing the trend in the behavior of the SOA still proves useful. Here, it can be seen that as the order of modulation increases, so does the EVM. It is not surprising as a higher order of modulation accentuates the changes in instantaneous power which was referenced to be a cause of nonlinearities in the SOA.

Figure 2.16c shows the EVM at different input powers into the SOA for different numbers of subcarriers. The EVM seems to be almost immune to changes in the number of subcarriers. This notion might seem counter intuitive at first since one would assume

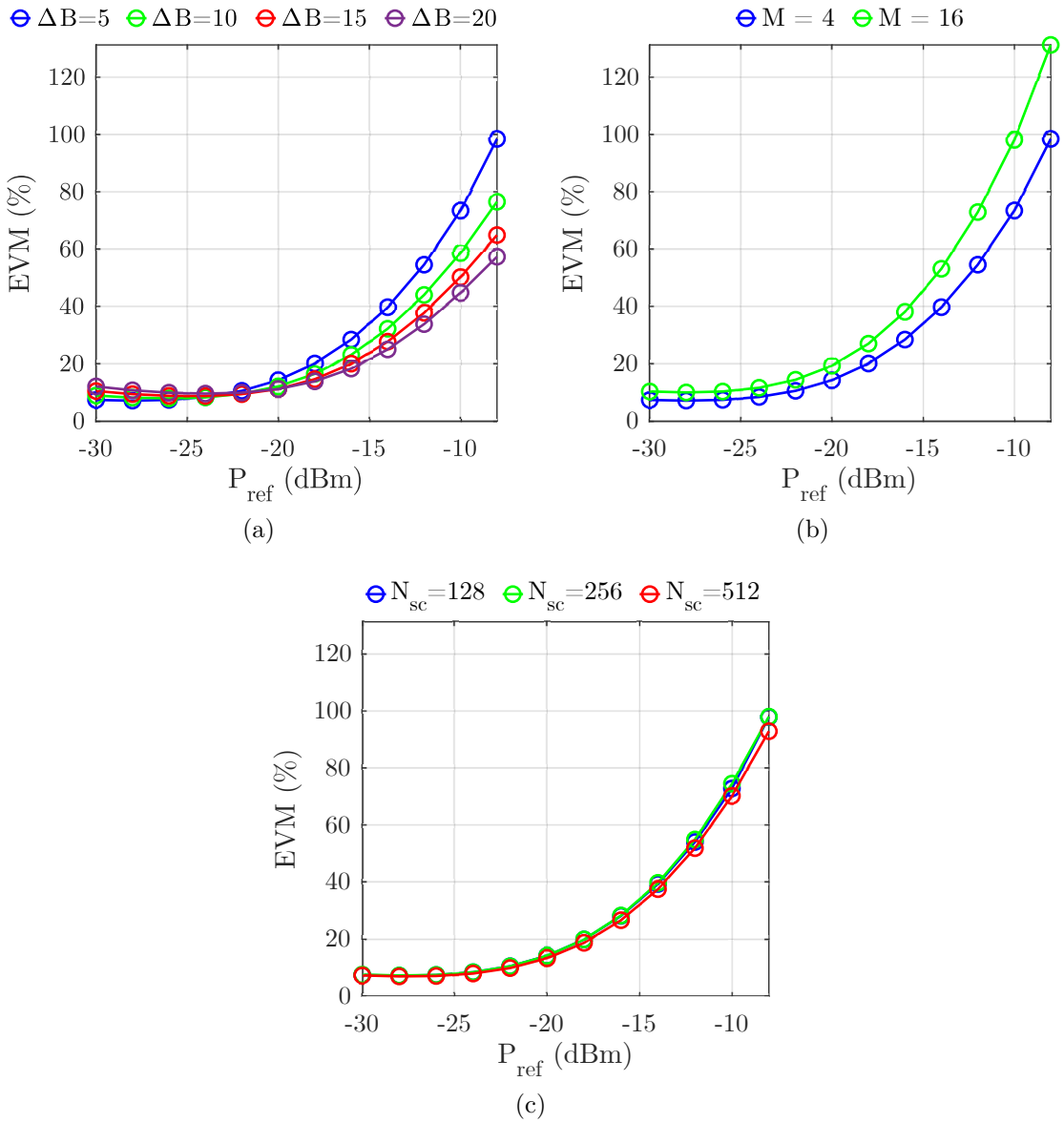


Figure 2.16 – EVM vs. P_{ref} for the no DPD case at (a) different bandwidths ΔB in GHz (b) at different modulation orders M (c) at different number of subcarriers N_{sc}

that an increase in the number of subcarriers leads to a decrease in subcarrier spacing and therefore an increase in the effects caused by FWM. However, it is important to note that for the same bandwidth, as the number of subcarriers increases, the power associated with each subcarrier decreases. This counter-effect could be an explanation as to why the EVM is unchanged since the lower power per subcarrier reduces the impact of FWM.

Now that the behavior of the SOA has been established with respect to N_{sc}, M, and ΔB, an analysis of the GMP should be undertaken. We start with a comparison with the no DPD case for different bandwidths (c.f. Figure 2.17a and Figure 2.17b). The GMP is able to linearize the SOA as the input power into the SOA for the worst case can be increased from -16 dBm to -10 dBm and still meet the EVM requirement of 30%. The improvements added by GMP are still visible but are lower for higher bandwidths as there are less nonlinearities to compensate and more noise. Also note that here a larger bandwidth entails a higher sampling frequency meaning the predistorters would generally require larger memory depths and thus more predistorter coefficients (with the identification problems that go along with that requirement).

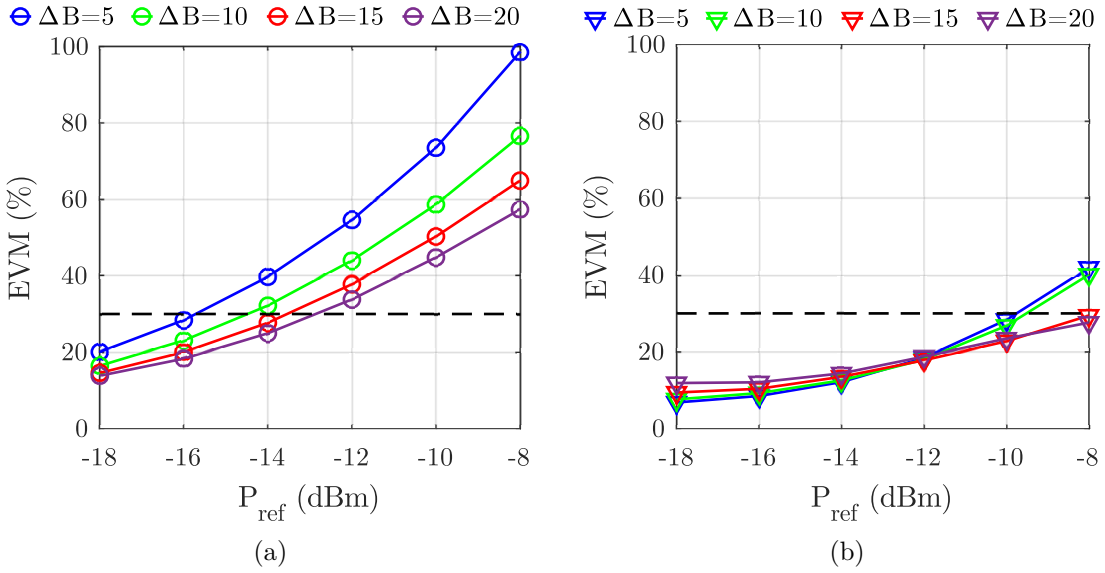


Figure 2.17 – EVM vs. P_{ref} at different bandwidths ΔB in GHz (a) for the no DPD case (b) for the GMP case

Table 2.3 shows the different data rates and increase in reach achieved with the use of GMP. It seems that operating with a bandwidth of 15 GHz allows us to increase data rate by three-fold while having a similar reach increase.

Comparing the 4-QAM and 16-QAM configuration is tricky as they have different requirements in terms of EVM. For the same BER of 10^{-3}, 4-QAM has an EVM limit of 30% whereas 16-QAM has an EVM restriction of 16% [Sch+11]. It is first more appropriate to compare the no DPD cases with their respective GMP results. Figure 2.18a shows the comparison for 4-QAM where the power can be increased from -16 dBm to -10 dBm

Table 2.3 – Data rate R_b and reach increase when changing ΔB The reach increase is calculated using the power gain between the no DPD and the GMP case

ΔB (GHz)	R_b ($\times 10^8$) (Gbit/s)	Reach increase (km)
5	2.31	20.89
10	4.61	13.08
15	6.92	18.27
20	9.22	14.79

while staying below the EVM limit of 30%. Figure 2.18b looks into the 16-QAM case where the input power can be increased from -23 dBm (cf. Figure 2.16b) to -14 dBm.

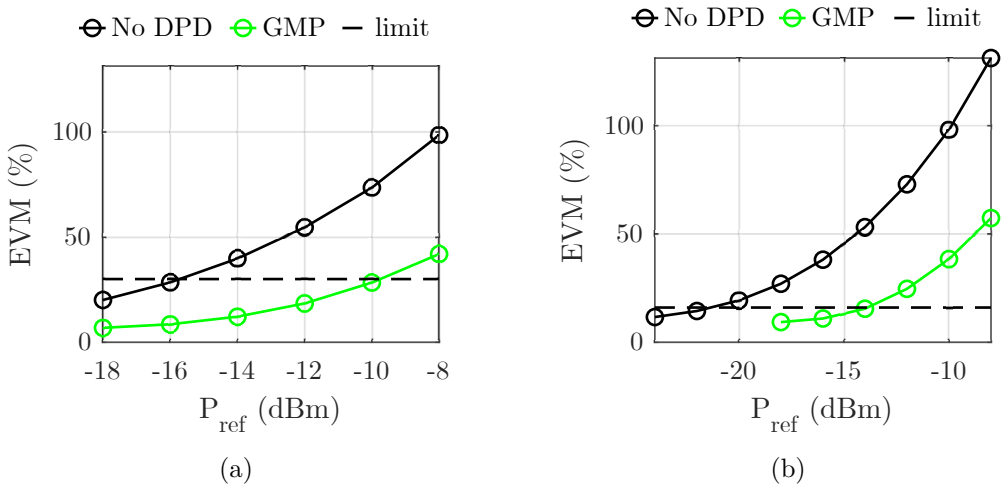


Figure 2.18 – EVM vs. P_{ref} for the no DPD and GMP cases for (a) 4-QAM (b) 16-QAM

Table 2.4 shows the data rate and reach increase for different order of modulations. It can be seen that the reach increase is higher for the 16-QAM case which is in favor of choosing preferably a 16-QAM configuration. Nevertheless, if one is restricted in terms of receiver sensitivity for example, choosing the 4-QAM configuration seems more appropriate as it allows for higher powers going into the receiver. Indeed, the difference in reach between 4-QAM and 16-QAM is about 31 km for the no DPD case and around 13 km for the GMP case. These values which were calculated according to the gain of the SOA show that GMP has more visible improvements in the 16-QAM case, but not enough to surpass the 4-QAM case in terms of overall reach. Furthermore, Figure 2.18a and 2.18b convey that the 4-QAM configuration does give more leeway with respect to the upper EVM limit. This can lead to think that the 4-QAM configuration is a safer

choice than the 16-QAM configuration especially since we will discover in Chapter 3 that when looking at the EVM variance when taking into consideration some of the system's uncertainties, the requirements might not be always met.

Table 2.4 – Data rate R_b and reach increase when changing order of modulation M. The reach increase is calculated using the power gain between the no DPD and the GMP case

M	R_b ($\times 10^8$) (Gbit/s)	Reach increase (km)
4	2.31	20.89
16	4.62	39.19

The last bit of analysis to perform is with an increasing number of subcarriers as shown in Figure 2.19a and Figure 2.19b. All cases here are similar whether DPD is applied or not. In the case where GMP is used, the input power into the SOA is increased from roughly -16 dBm to about -10 dBm for the same EVM limit of 30% for all cases.

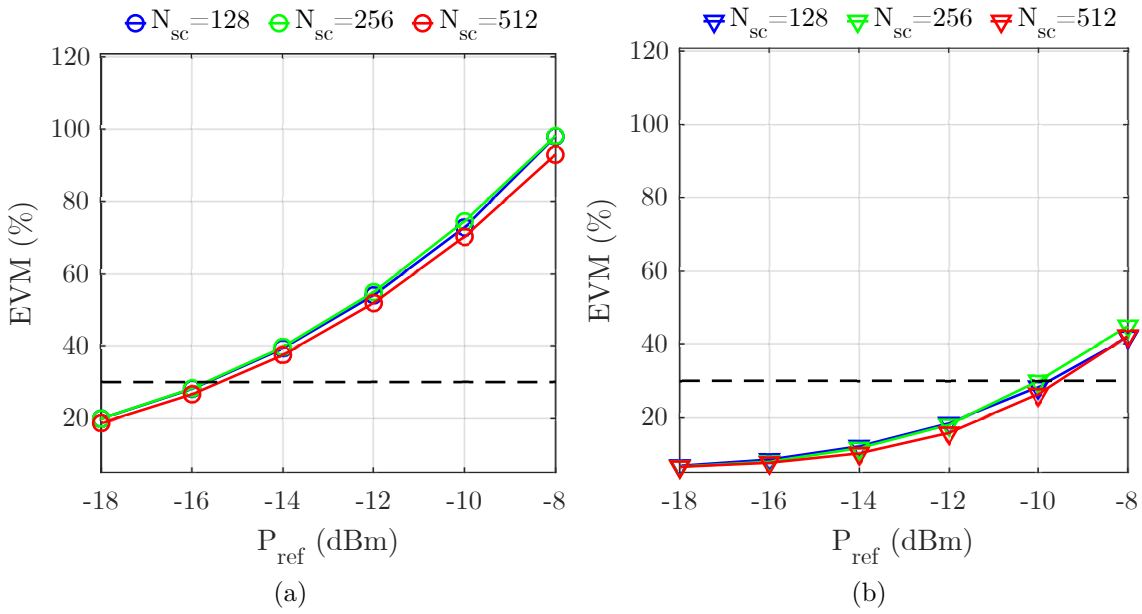


Figure 2.19 – EVM vs. P_{ref} at different N_{sc} (a) for the no DPD case (b) for the GMP case

If this study is extended to the data rate and reach increase (Table 2.5), it looks like it is favorable to choose a setup where the number of subcarriers is 512 since it yields slightly higher data rates for a similar reach increase.

Table 2.5 – Data rate R_b and reach increase when changing the number of subcarriers N_{sc}. The reach increase is calculated using the power gain between the no DPD and the GMP case

N_{sc}	R_b ($\times 10^8$) (Gbit/s)	Reach increase (km)
128	2.31	20.89
256	2.68	20.89
512	2.80	20.89

As was seen in Figure 2.16c, the number of subcarriers does not seem to change the behavior of the SOA for a constant bandwidth. It is then pertinent to verify that a common predistorter can be used for all cases lowering the amount of calculations made to create multiple predistorters. Figure 2.20 shows the EVM for different numbers of subcarriers using the predistorter calculated at $N_{sc} = 128$.

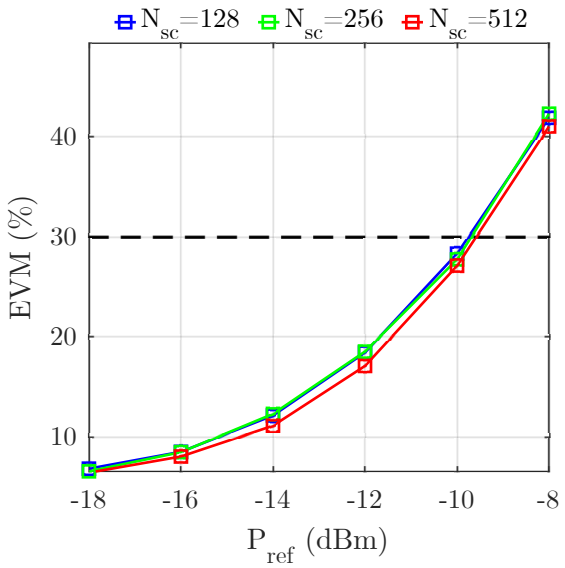


Figure 2.20 – EVM vs. P_{ref} for the GMP case at different N_{sc} using the GMP predistorter computed with $N_{sc} = 128$

It is seen in this case that it is possible to use the same predistorter in all cases which therefore means that it is also feasible to simplify the physical layer as only one set of coefficients need to be stored.

2.5 Digital predistortion for an in-line SOA

The previous section dealt with several aspects of DPD in the context of a booster SOA based CO-OFDM system in an optical B2B configuration. This was done in order to verify the ability of DPD to linearize the emitter. This section will cover the case where the SOA is put in-line instead. To do so, the setup shown in Figure 2.21 will be used.

The transmission is similar to that of Figure 2.1 used for the booster SOA case with a few exceptions. An input bit stream is used to construct the OFDM signal. Here, the bit stream used for the synchronization preamble is different from the one used previously. The fiber will introduce imperfections and particularly chromatic dispersion. This can have a great influence on the synchronization at the receiver side. A more realistic synchronization algorithm therefore will be implemented. More details on the new synchronization algorithm will be given in the next section. As fiber is added and the SOA in-line is used now, some channel compensation will be used. We will explore the case where the channel is pre-compensated but also post-compensated. Finally, a study of digital post-distortion (DPoD) will be done, hence the activating of that block in the chain.

Figure 2.21 – In-line SOA implemented CO-OFDM transmission. Mod: modulation, equal.: equalization, sync.: synchronization, P/S: parallel to serial, DPD: digital predistortion, PAPR: peak to average power ratio, CD: chromatic dispersion, comp.: compensation, DAC: digital to analog conversion, opt.: optical, att.: attenuation, SOA: semiconductor optical amplifier, demod.: demodulation, ADC: analog to digital conversion, imb.: imbalance, CG/CP: common gain/common phase, DPoD: digital post-distortion, CFO: carrier frequency offset, S/P: serial to parallel, freq.: frequency, s_k and r_k are the sent and received QAM symbols, and $x[n]$ and $y[n]$ are the time signals used to estimate the predistorter

2.5.1 Synchronization

The synchronization scheme can be divided into two parts: the synchronization sequence design at the emitter and the synchronization algorithm at the receiver. For the numerical results shown in section 2.4.1, the time synchronization is made at the receiver by using a simple amplitude based cross-correlation between the received signal and the known synchronization preamble. At the emitter, the synchronization preamble is designed by relying on pattern repetition. To achieve this, a subcarrier associated with a binary phase shift keying (BPSK) symbol is exclusively turned on every L subcarrier making the pattern repeat L times. An example of that is pictured in Figure 2.22a with $L = 4$. The synchronization preamble pattern is apparent in the time domain as seen in Figure 2.22b compared to, for instance, the data OFDM symbol in Figure 2.22c which does not show any signs of motif. Note that to maintain the same mean power as the rest of the frame, the synchronization preamble's amplitude is multiplied by a factor of 2.

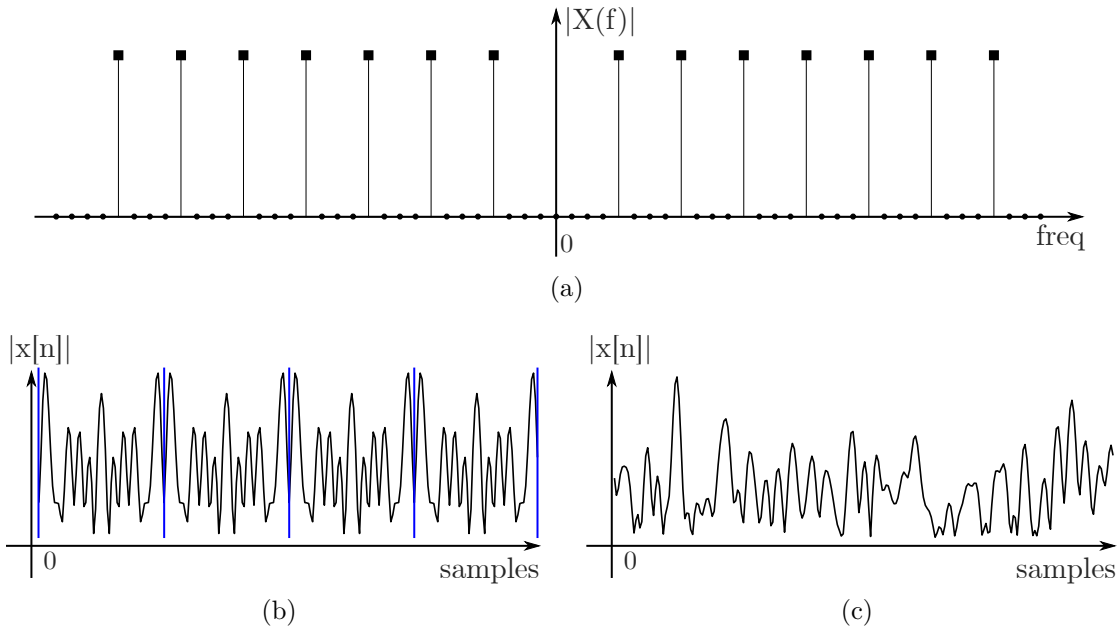


Figure 2.22 – (a) Synchronization preamble in the frequency domain with $N_{sc} = 64$ and $L = 4$. The null subcarriers used for the oversampling are not shown. (b) Synchronization preamble in the time domain without cyclic prefix and with $L = 4$ (c) a typical OFDM data symbol in the time domain without cyclic prefix

While optimizing synchronization for OFDM is not the aim of this thesis work, we eventually implemented a more robust approach for the in-line SOA numerical results.

That method proposed by *Minn et al.* [Min+03] which is considered robust for OFDM systems.

As far as preamble design is concerned, *Minn et al.* relies on pattern repetition as was the case in the previous synchronization scheme but it uses a different data stream and assigns different weights to each pattern segment. The *Minn et al.* sequence construction creates a specific bit stream using Golay sequences [Gol61]. These sequences' sole purpose in this thesis is to restrain nonlinearities at the transmitter by ensuring a low PAPR OFDM frame [Van96]. The sequence is created such that two complimentary series are used. We will use the simplest pair example with $\mathbf{a}_1 = [0\,0]$ and $\mathbf{b}_1 = [0\,1]$, the literature does provide algorithms for the search of binary sequence pairs of a certain length [Bor+04]. The pair is deemed complimentary since the number of pairs of like elements in one series is the same as the number of pairs of unlike elements in the other series. Longer length sequences can be created by combining $\mathbf{a}_n\mathbf{b}_n$ and using the following properties: $\mathbf{a}_n = \mathbf{a}_{n-1}\mathbf{b}_{n-1}$ and $\mathbf{b}_n = \mathbf{a}_{n-1}\bar{\mathbf{b}}_{n-1}$ where $\bar{}$ indicates a bitwise complement. For example, $\mathbf{a}_2\mathbf{b}_2 = \left[\mathbf{a}_1\ \mathbf{b}_1\ \mathbf{a}_1\ \bar{\mathbf{b}}_1\right] = [0\,0\,0\,1\,0\,0\,1\,0]$. The Golay bit stream is then modulated into BPSK symbols.

The higher the number of repeated sequences, the better the synchronization [Min+03]. Considering an example where $N_{sc} = 512$ and $L = 16$ repeated sequences, we would need to construct a baseline Golay sequence of length $M_{sc} = 32$. We construct a Golay sequence of length 32 for the bit stream which can be modulated into BPSK symbols that are assigned to 32 subcarriers. Contrary to Figure 2.22a, all subcarriers are active. An oversampling of 4 is done by adding null subcarriers in the frequency domain. The OFDM synchronization symbol is then defined as

$$s_{sync}[n] = \sqrt{M_{sc}\,\rho_{os}} \sum_{k=0}^{M_{sc}-1} s_{k,sync} e^{i2\pi \frac{k}{M_{sc}\rho_{os}} n} \tag{2.5}$$

with $s_{k,sync}$ being the BPSK symbol put on the subcarrier, and with n going from 1 to $M_{sc}\,\rho_{os}$. That sequence is depicted in the time domain in Figure 2.23a. The OFDM synchronization symbol is then a 16-periodic signal in the time domain (Figure 2.23b). To be more discriminatory, the repeated sequence can follow a specific pattern. The chosen pattern which determines the sign of the repeated sequence is the following: [1 -1 -1 1 1 1 -1 -1 1 -1 1 1 -1 1 -1 -1]. This pattern is considered one of the best in terms of synchronization with $L = 16$ according to *Minn et al.* after a computer search. A cyclic prefix is then added onto the signal.

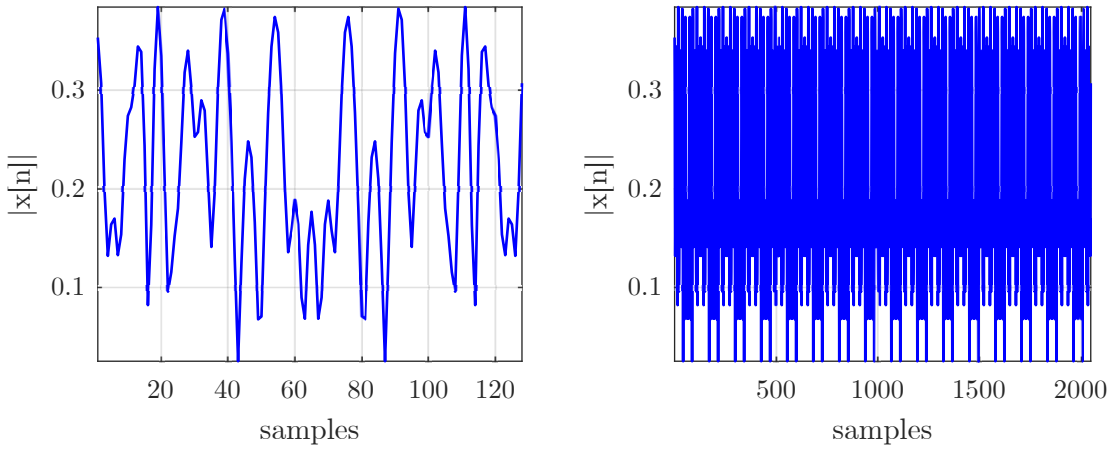


Figure 2.23 – (a) One sequence in the time domain (b) the sequence repeated $L = 16$ times to form the entire synchronization symbol (here without cyclic prefix)

At the receiver, the metric used for synchronization is the following:

$$\begin{aligned} \Lambda &= \left(\frac{L}{L-1} \frac{\mid P(d) \mid}{E(d)} \right)^2 \\ \text{with} \quad P\left[d\right] &= \sum_{k=0}^{L-2} b(k) \sum_{m=0}^{M_{sc}-1} r^*\left[d + kM_{sc} + m\right] \; r\left[d + (k+1)\, M_{sc} + m\right] \\ \text{where} \quad b(k) &= p(k)p(k+1) \\ \text{and} \quad \text{with} \quad E\left[d\right] &= \sum_{m=0}^{M_{sc}-1} \sum_{k=0}^{L-1} \mid r(d + m + kM_{sc}) \mid^2 \end{aligned} \tag{2.6}$$

Λ is the metric to be maximized, L is the number of repeated sequences of length M_{sc}, r is the received signal, d designates the starting point of the signals used for the auto-correlation of the received signal, and $p(k)$ represents the respective signs of the repeated sequences.

A visual depiction of the metric is portrayed in Figure 2.24 where we can discern that P gives the bulk of the contribution to the metric Λ. The received signal here is of the following form: $N_{sc} = 512$, $\rho_{os} = 4$, $\rho_{cp} = \frac{1}{8}$, and the total number of OFDM symbols in the frame is $N_{ofdm} = 32$ including the syncronization and equalization preamble with a total frame length of $N_{sc}\, \rho_{os}\, (1 + \rho_{cp}) N_{ofdm} = 73\,728$ samples. The timing search is done over the span of one OFDM frame and one OFDM symbol meaning the search is over

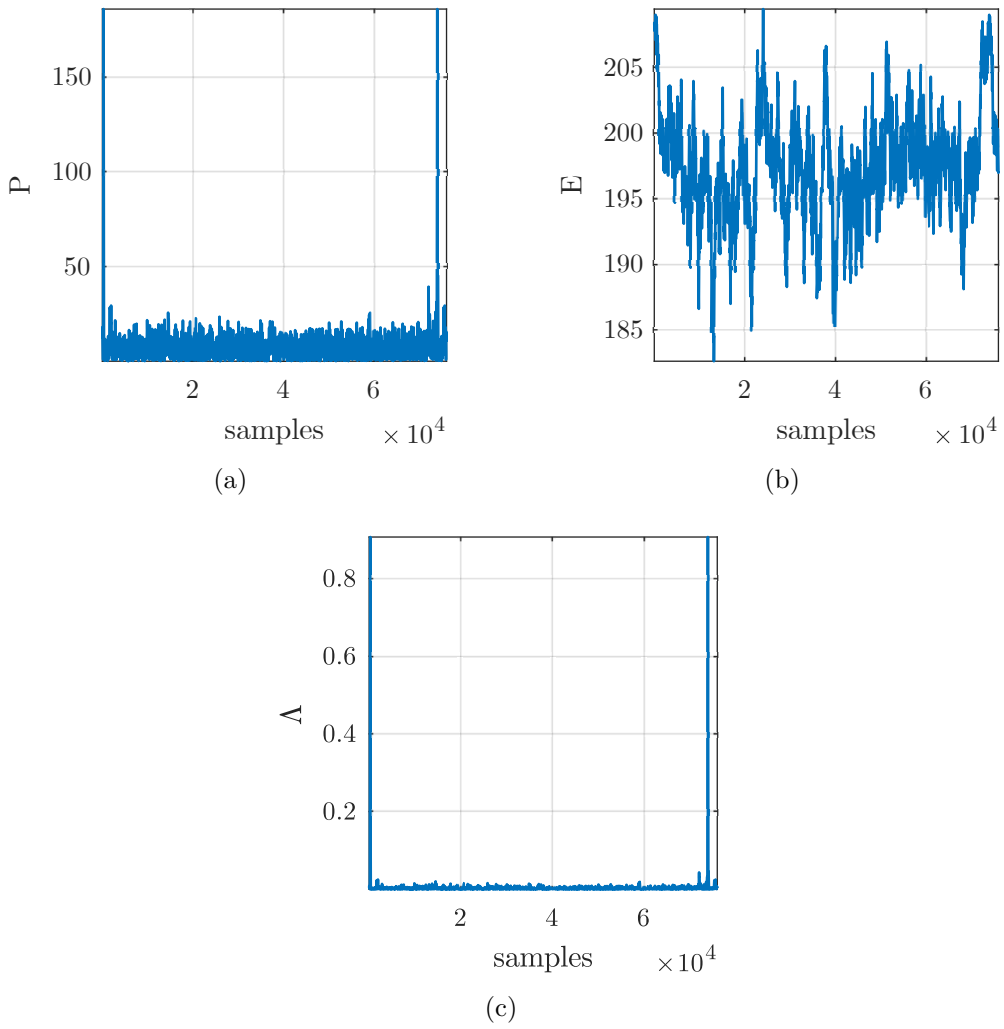


Figure 2.24 – (a) P (b) E (c) Λ over the length of one OFDM frame plus one OFDM symbol

76 032 samples. Figure 2.24c allows us to find two values $d = 258$ and $d = 73\,986$. The $d = 73\,986$ peak corresponds to the start of a second frame. The estimated time delay d_{max} is computed such that $d_{max} = d - N_{cp} - 1$ with N_{cp} the length of the cyclic prefix.

After synchronizing the signal at the sample level, it is then possible to perform a fractional synchronization. To achieve that, it is possible to use an estimated impulse

channel response as such:

$$\mathbf{S} = \begin{bmatrix} s_{sync}[1] & s_{sync}[0] & \cdots & s_{sync}[-N_{cp}+1] \\ s_{sync}[2] & s_{sync}[1] & \cdots & s_{sync}[-N_{cp}+2] \\ \vdots & \vdots & \cdots & \vdots \\ s_{sync}[N] & s_{sync}[N-1] & \cdots & s_{sync}[N-N_{cp}] \end{bmatrix} \tag{2.7}$$

where $s_{sync}[n]$ is the sent synchronization preamble with $s[1]$ being the starting point of the preamble without the cyclic prefix, and N_{cp} is the length of the cyclic prefix.

An impulse channel response can then be estimated to be:

$$\hat{\mathbf{h}} = \mathbf{S}^{\dagger}\mathbf{r} \tag{2.8}$$

with $\mathbf{r}$ being the received preamble without cyclic prefix. The fractional delay corresponds to $\frac{T_s \Delta\Theta}{2\pi}$ with $\Delta\Theta$ being the phase change obtained from the frequency channel response $\hat{\mathbf{H}}$ in the passband, and T_s the sample period.

Figure 2.25 illustrates the change in the constellation after sample and fractional synchronization when a 20 GHz signal is sent through the SOA at a reference power of -18 dBm. Contrary to the earlier sections, the SOA now introduces a more visible fractional delay as the sampling rate is high (80 GHz).

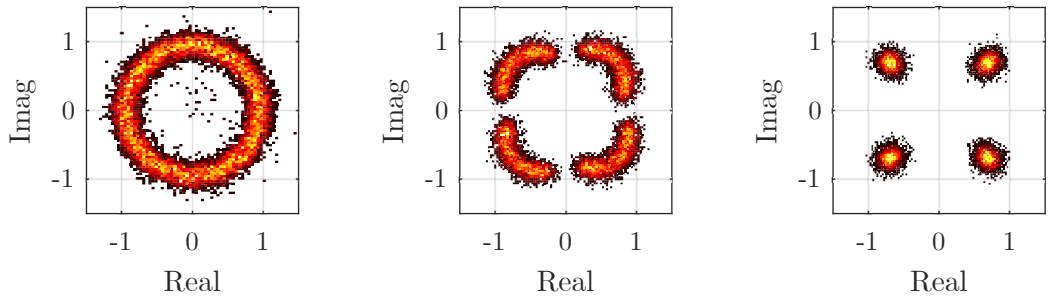


Figure 2.25 – Constellation for a booster SOA at $P_{ref} = -18$ dBm with a modeled channel delay (a) without synchronization (b) after sample synchronization (c) after fractional synchronization. The brightest areas have the highest concentration of QAM symbols

Now that the synchronization process has been described, the next stage to look into is the equalization to compensate the undesirable effects of the channel, most notably those of the fiber.

2.5.2 Chromatic dispersion mitigation

The added value of OFDM is that it enables us to perform simple equalization. In a similar fashion to that of the synchronization scheme, the channel equalization scheme can be divided into two phases. First, the preamble is constructed at the emitter and then the compensation is done at the receiver. In this thesis, the equalization preamble is designed in a straightforward way using pseudo-random rotated 4-QAM symbols according to the 802.11 convention.

At the receiver, the channel response can be easily obtained by doing a simple division of the received and sent equalization preamble in the frequency domain.

Before looking into what the equalization at the receiver looks like, a set of parameters for the simulations should be chosen. Using the same system parameters as in Table 2.1 from the previous simulations, we can take notice of the minor effects of chromatic dispersion displayed in Figure 2.26.

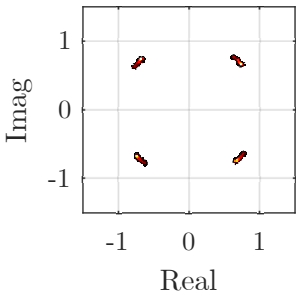


Figure 2.26 – Constellation of a 5 GHz OFDM signal with N_{sc}=128 with 80 km worth of chromatic dispersion. The brightest areas have the highest concentration of QAM symbols

For this study, we choose to accentuate these effects by increasing the signal bandwidth (excluding oversampling) to ΔB=20 GHz. Since the bandwidth is increased by a factor of 4, the number of subcarriers is also increased from 128 to 512 to stay consistent. Figure 2.27 demonstrates that as expected, chromatic dispersion is more pronounced in that new scenario.

With that new scenario, we can start making observations on the equalization scheme. Usually, class B+ passive optical network (PON) transceivers' minimum and maximum output optical powers are 1.50 dBm and 5 dBm respectively. Taking some precautions, we choose to pursue two potential in-line SOA scenarios where the power into the fiber at the emitter is either 2 dBm or 4 dBm in which case the power into the SOA after 80 km would be either -14 dBm or -12 dBm for a single-mode fiber (SMF) with an attenuation

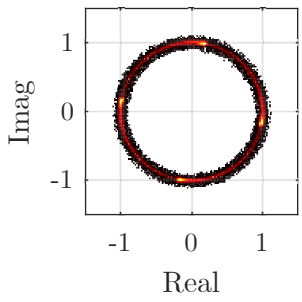


Figure 2.27 – Constellation of a 20 GHz OFDM signal with N_{sc}=512 with 80 km worth of chromatic dispersion. The brightest areas have the highest concentration of QAM symbols

factor of 0.2 dB/s. The SOA being in-line for this study, another 80 km of fiber would be inserted after the amplifier.

Using the configuration of 80 km of fiber → SOA → 80 km of fiber for two SOA input powers of -14 dBm and -12 dBm, we attempt to showcase the behavior of chromatic dispersion by calculating the channel response using the whole received signal and equalization preamble with the oversampling included. The left-hand side of Figure 2.28 illustrates the calculated channel response (amplitude and phase) for a 20 GHz (passband) OFDM signal. The two figures on the left-hand side demonstrate the overall appearance of chromatic dispersion as an allpass filter since the absolute value tends to be flat with the phase of the signal being the only element that changes significantly as described by Equation 1.12. The fluctuations around the module and the phase are due to the SOA. Note that here, the SOA gain is removed at the receiver with the common gain compensation block which explains why the $P_{ref} = -14$ dBm and $P_{ref} = -12$ dBm cases are at the same level. Also note that the gap in the middle is due to the subcarriers near DC being null by construction.

For the channel equalization process at the receiver, the zero-forcing method is used where the received subcarriers are divided by the non-null subcarriers of the known equalization preamble in the frequency domain [Shi+08]. Note that this method is simple, but it intrinsically causes signal to noise ratio (SNR) to decrease for subcarriers with low amplitudes [Mar+09]. The obtained channel frequency response links back to the two figures on the right-hand side of Figure 2.28. Looking at the 20 GHz passband region, we can comment on the fact that the observed behavior is quite similar regardless of the input power into the SOA.

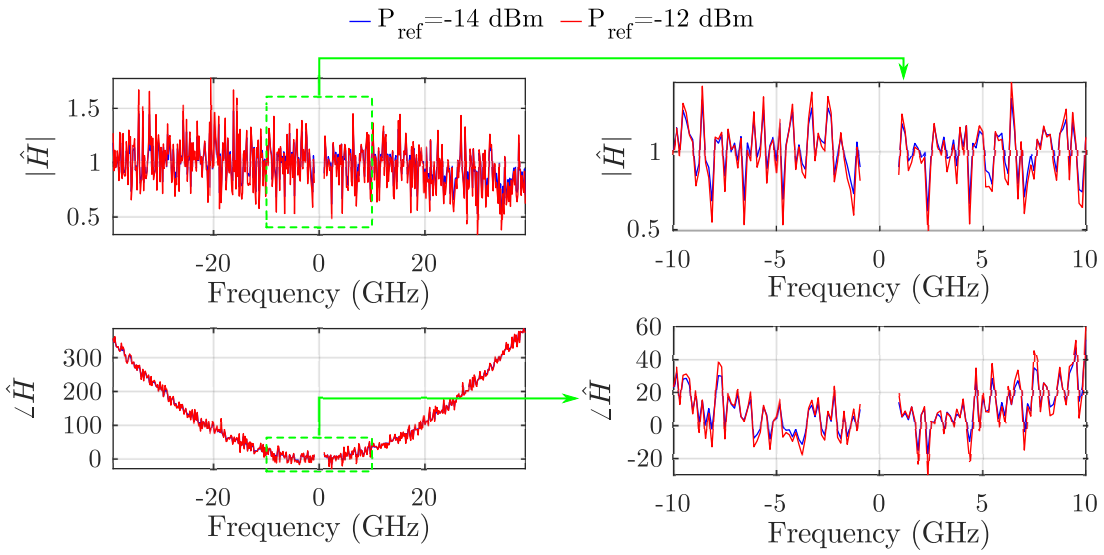


Figure 2.28 – Channel response of a 160 km transmission with an in-line SOA at 80 km. | | is the module and ∠ is the phase. The gain of the SOA is removed by the common gain compensation block

As a first step on this study, we choose to focus on the -14 dBm case. The equalization is done by applying the inverse of the channel response in the passband region seen on the right-hand side of Figure 2.28. The constellations pictured on Figure 2.29 demonstrate that doing the equalization in the frequency domain helps with the added value of implementing simple calculations and not needing to know the fiber length.

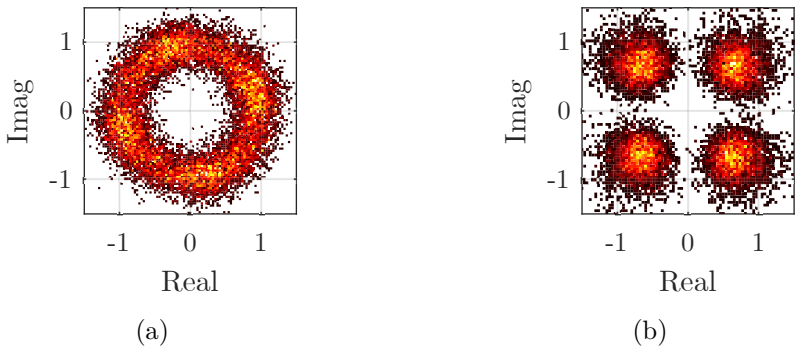


Figure 2.29 – (a) Constellation with an in-line SOA at $P_{ref} = -14$ dBm with 160 km of fiber before equalization (b) Constellation with an in-line SOA at $P_{ref} = -14$ dBm with 160 km of fiber after equalization. The brightest area have the highest concentration of QAM symbols

It might be interesting to contemplate compensating the fiber segment that is placed before the SOA. The approach [Egh+14] taken in this thesis is based on a complex impulse response derived from the inverse frequency response of the chromatic dispersion defined in Equation 1.12, i.e.

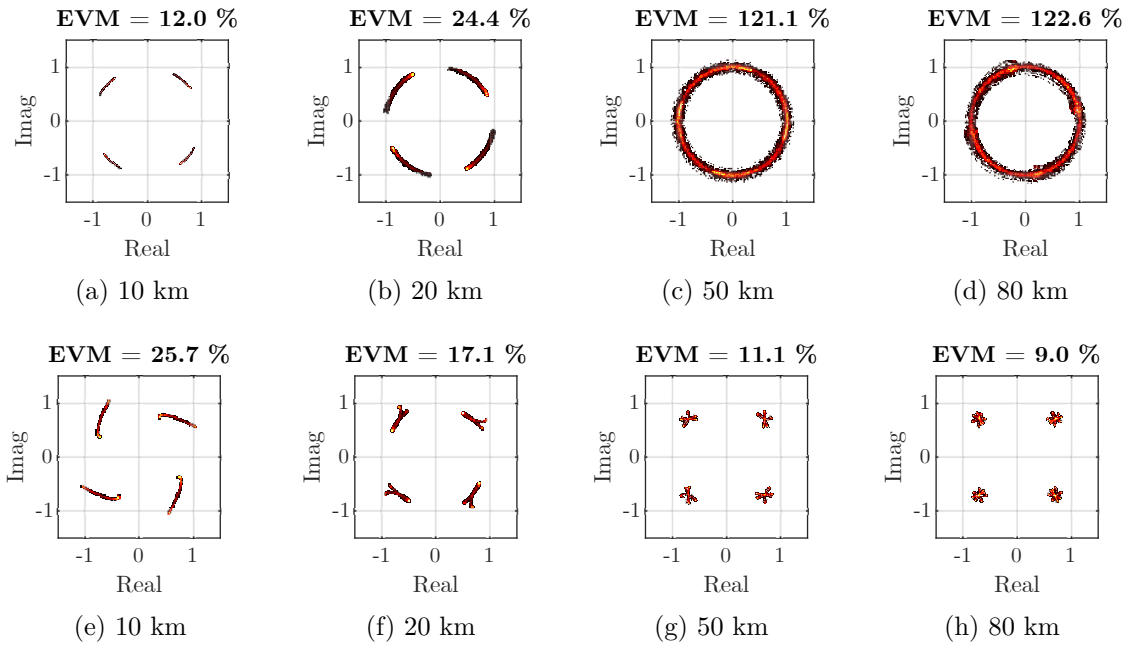
$$\begin{aligned} h[n] &= \sqrt{\frac{i}{4K\pi}} e^{-i\frac{n^2}{4K}}, \quad -\lfloor \frac{N}{2} \rfloor \leq n \leq \lfloor \frac{N}{2} \rfloor \\ \text{where} \quad N &= 2\lfloor 2K\pi \rfloor + 1 \\ \text{and} \quad K &= \frac{D\lambda_{LO}^2 z}{4\pi c T_s^2} \end{aligned} \tag{2.9}$$

where N is the length of the finite impulse response (FIR) filter which is always an odd number, D is the fiber dispersion parameter, λ_{LO} is the wavelength, z is the considered fiber length, and c is the speed of light.

A quick preliminary study of that method of compensation is done simulating only the fiber chromatic dispersion as an undesirable effect. The pre-compensation block is placed right before the DAC as shown in Figure 2.21. Figure 2.30 displays the constellations of 10, 20, 50, and 80 km of fiber without compensation on the first row, and the same length of fiber on the second row with compensation using the corresponding filter.

It is possible to discern from these constellations that the implemented scheme works best for longer lengths of fiber. It makes sense since the filter length is directly related to the length of the fiber. It is also important to note that in cases where the fiber length is quite low, the compensation starts being detrimental to the quality of transmission. In these cases, it seems preferable to not activate the pre-compensation block.

With the equalization and synchronization methods having both been set in stone, we can now add DPD into the mix to mitigate the nonlinear effects of the SOA and the IQ modulator. The goal will be to not only evaluate the effectiveness of DPD, but also develop a method of identification that is compatible with the in-line SOA scenario.



(a) 10 km (b) 20 km (c) 50 km (d) 80 km

(e) 10 km (f) 20 km (g) 50 km (h) 80 km

Figure 2.30 – Constellations of different fiber lengths. The first row designates the constellations without compensation and the second row the constellations with compensation. The brightest areas have the highest concentration of symbols

2.5.3 Generalized memory polynomials for an in-line SOA

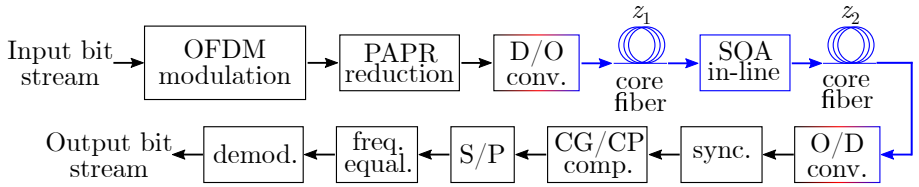
The parameters used for the next two sections are shown in Table 2.6. These parameters are almost the same as the ones used in Section 2.4 apart from the number of subcarriers and the bandwidth which were changed to induce more chromatic dispersion. Take note that this section only deals with the GMP predistorter.

The different scenarios that are explored in this section are represented in Figure 2.31. Four scenarios are looked into: a transmission without DPD, two different transmissions with DPD, and one last transmission with both DPD applied as well as pre-compensation of the fiber. The total fiber length, $z_1 + z_2$, is 200 km.

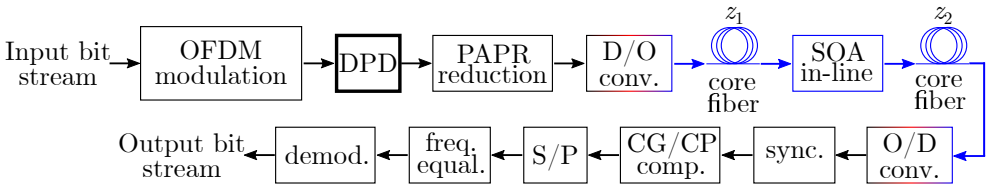
The identification phase of the DPD block is done using either one of the procedures displayed in Figure 2.32. One of the scenarios from Figure 2.31b and the scenario in Figure 2.31c essentially use a preditorder learned in the optical B2B setup, i.e. without any fiber present. The other scenario in Figure 2.31b uses a predistorter learned with a fiber segment positioned before the in-line SOA of length z_1. A new predistorter is identified every time that z_1 changes. The fiber length z_2 is inserted at the validation

Table 2.6 – Parameters used for the in-line SOA scenario

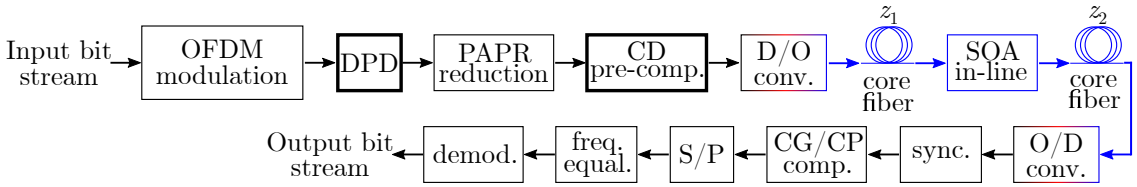
Parameter	Symbol	Value	Units
Modulation order	M	4	-
Number of subcarriers	N_{sc}	512	-
Bandwidth	ΔB	20	GHz
Cyclic prefix factor	ρ_{cp}	$\frac{1}{8}$	-
Total number of QAM data symbols sent	N_{data}	3180	-
Oversampling factor	ρ_{os}	4	-
Clipping ratio	γ	8	dB
MZM peak-to-peak voltage	V_{pp}	$1.25 \times V_{\pi}$	volts
Laser wavelength	λ_{LO}	1540	nm
SOA bias current	I_{bias}	150	mA



(a) No DPD



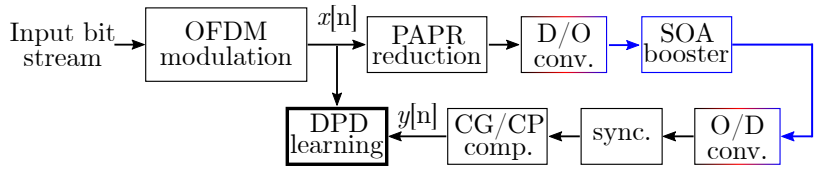
(b) DPD learned in the presence of fiber (z_1 only) or without fiber. The predistorter is then inserted into the transmission chain for validation (with both z_1 and z_2)



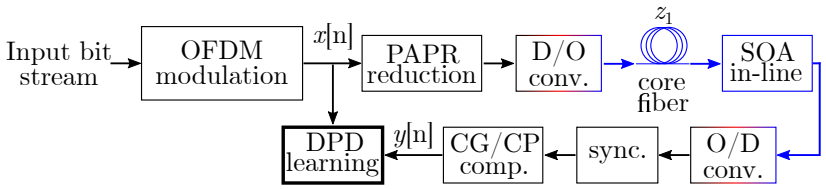
(c) DPD learned without fiber. Validation phase done with chromatic dispersion (CD) pre-compensation

Figure 2.31 – In-line scenarios validation phase. D/O conv.: digital to optical conversion, sync: synchronization, CG/CP comp.: common gain/phase compensation, S/P: serial to parallel

phase of DPD and is to be compensated at the receiver using the OFDM equalization preamble along with any residual channel dispersion effects.



(a) DPD learned without fiber



(b) DPD learned with z_1 length of fiber before the in-line SOA

Figure 2.32 – In-line scenarios DPD learning phase options. D/O conv.: digital to optical conversion, sync: synchronization, CG/CP comp.: common gain/phase compensation

Figure 2.33 shows the EVM obtained from the scenarios presented earlier for different fiber lengths placed before the SOA. Note that the total length of fiber is always 200 km meaning if, for example, $z_1 = 20$ km, then $z_2 = 180$ km.

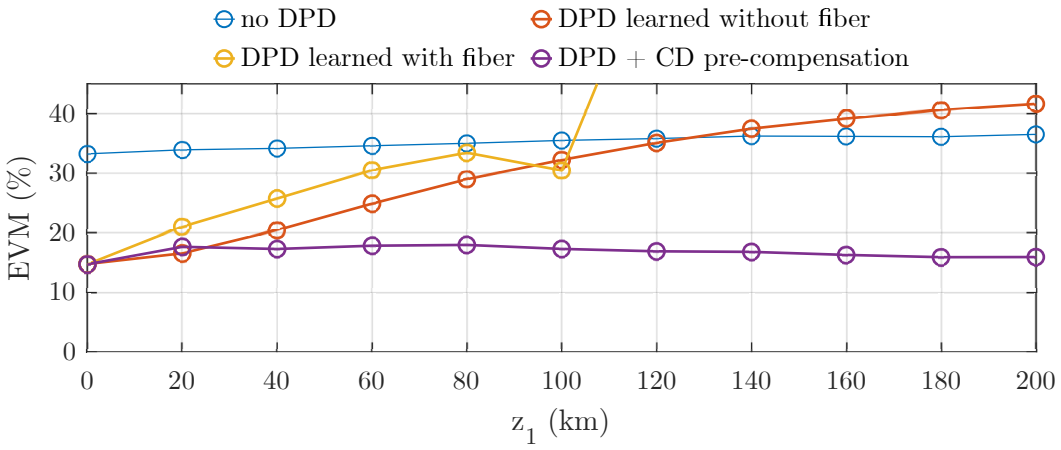


Figure 2.33 – In-line scenarios with the SOA moving along 200 km of fiber. The predistorter used with the chromatic dispersion (CD) pre-compensation is learned without fiber

An important notion to ascertain from Figure 2.33 is that even without performing pre-compensation of the fiber or DPD (blue curve), it seems that the channel frequency equalization scheme at the receiver is able to partly compensate the combined effects of chromatic dispersion and SOA nonlinearities regardless of the SOA's position in the transmission chain. Indeed there is not much difference in EVM whether $z_1 = 0$ km (where

the SOA acts as a booster) or $z_1 = 200$ km (where the SOA acts as a pre-amplifier). The predistorter learned with z_1 length of fiber (yellow curve) has better results than the no DPD case up to 100 km of fiber before the SOA. Note that once $z_1 > 100$ km, the EVM explodes; the yellow curve is thus not shown in full to keep the rest of the plot legible. It is still possible to meet the EVM requirements of 30% if the length of fiber before the SOA is below 60 km.

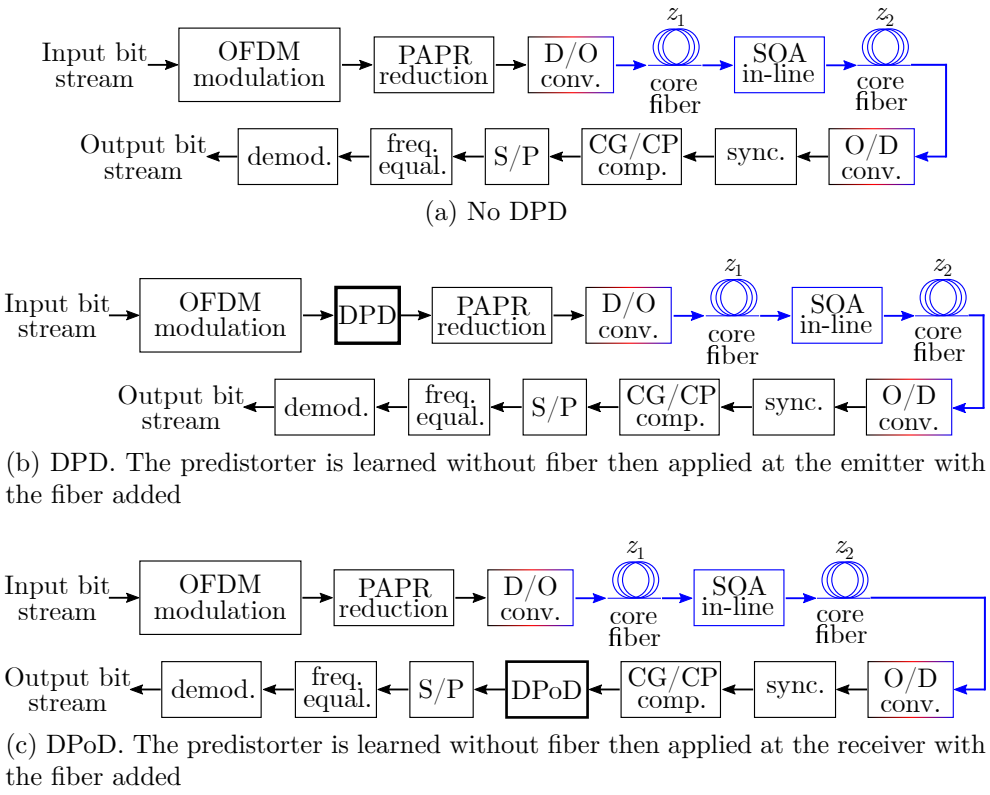
The predistorter identified in optical B2B (orange curve) actually performs better than the predistorter learned with fiber (yellow curve). This could be attributed to the fact that DPD uses the NMSE as a metric to find the proper coefficients even though the NMSE is not suitable in the presence of chromatic dispersion. The predistorter learned without fiber is also able to outperform the case with no DPD so long as the length of fiber before the SOA is not too high ($z_1 < 120$ km). Finally, the predistorter learned without fiber is valid up to 86 km before the SOA as it meets the EVM requirement of 30% for 4-QAM.

Applying chromatic dispersion pre-compensation of the fiber of length z_1 located before the SOA (purple curve) allows for room to spare in terms of EVM as it gives values well below the desired limit of 30%. It is therefore preferable to pre-compensate chromatic dispersion so that the DPD can fully play its role. Otherwise, it will be necessary to restrict the length of fiber before the SOA. Needless to say, these results are contingent on the OFDM signal's bandwidth. Indeed, for a smaller bandwidth, the chromatic dispersion is less important and its pre-compensation is therefore less useful in that case unless the length of fiber z_1 before the SOA is larger.

2.5.4 Digital predistortion vs. digital post-distortion

In section 2.5.3, the GMP was evaluated as a predistorter in different settings. The current section serves to draw a comparison between DPD applied at the emitter, and DPoD applied at the receiver. Figure 2.34 shows the scenarios that are explored. The DPD scenario uses a GMP predistorter learned in the optical B2B setup (without fiber) then applied at the emitter. The DPoD scenario uses the same GMP learned without fiber but applies it at the receiver instead.

In Figure 2.35, the EVM is displayed at different lengths of fiber placed before the SOA for the three scenarios of no DPD, DPD, and DPoD. Two aspects seem worthy of being mentioned. The first one is that DPD and DPoD seem to work in a somewhat inverted manner; that is, DPD works best when the SOA is positioned close to the emitter and



(a) No DPD

(b) DPD. The predistorter is learned without fiber then applied at the emitter with the fiber added

(c) DPoD. The predistorter is learned without fiber then applied at the receiver with the fiber added

Figure 2.34 – In-line scenarios for the no DPD, DPD, and DPoD cases. The predistorters used for DPD and DPoD are the same but applied at different locations of the transmission chain. D/O conv.: digital to optical conversion, sync: synchronization, CG/CP comp.: common gain/phase compensation

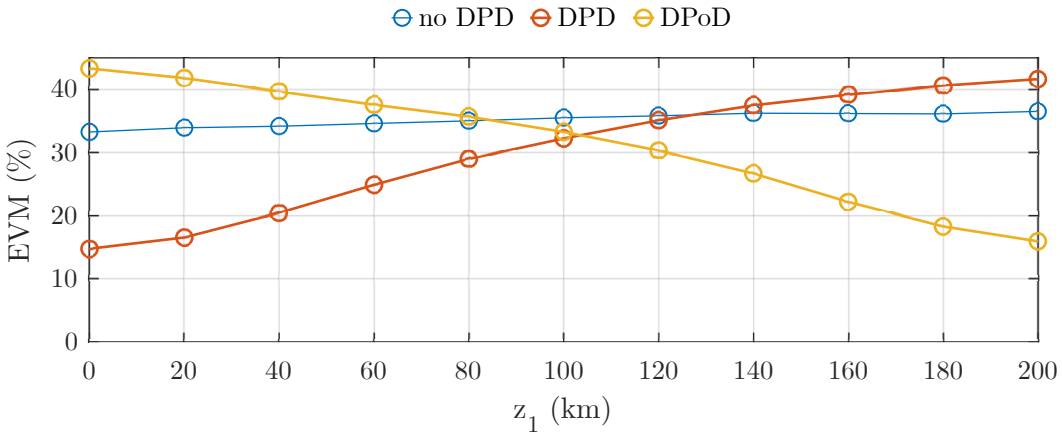


Figure 2.35 – In-line scenario with the SOA moving along 200 km of fiber. Predistortion vs. post-distortion

DPoD is better when the SOA is closer to the receiver until about the halfway point of 100 km where they cross. It makes sense intuitively, but Figure 2.35 provides a visual confirmation of that expected behavior. With that in mind, the choice of where to apply the mitigation can be done by directly looking at the anticipated position of the amplifier. The second aspect to take notice of is the slightly asymmetrical relationship between DPD and DPoD in terms of EVM. In fact, DPD tends to perform slightly better overall. In theory, if the channel is linear (up to a certain order of nonlinearity [Sch76]), DPD = DPoD and therefore DPD $\rightarrow$ SOA = SOA $\rightarrow$ DPoD. The reality does not perfectly abide by this relationship as can be seen in Figure 2.35. Knowing the opto-electronic conversion is considered ideal in our scenario, it is possible that this asymmetry could be even more noticeable in a real transmission. This could possibly give more reason to focusing on optical transmitter compensation in order to launch a more desirable signal into the optic fiber instead of dealing with a distorted and noisy signal.

2.6 Conclusion

This chapter covered simulation results obtained in the context of a CO-OFDM transmission with a booster SOA then an in-line SOA. Some of these results have been presented at the ICCE 2020 [Sim+21c]. To set up the analyses, a relationship is first established between the EVM and BER to justify using the EVM as the main metric of measure since it requires lower computational times. Then the process for PAPR reduction is presented as a way to help restrict the nonlinearities of the system although it is noted that it is not enough to mitigate these effects. The Hill-Climbing algorithm is introduced as a means to rapidly find the structural parameters of predistorters, especially in the event that the predistorter has many parameters like the GMP. It was shown that better results are obtained when the predistorters are kept quite simple therefore more weight should be put on the complexity of the predistorters (the complexity here being the number of coefficients). The high order nonlinearities of the SOA and the slow behavior of the SOA are difficult to capture and to compensate by the DPD (at least with the open loop method and the relatively low oversampling factor we use). Moreover, for high speed optical communications it is preferred to have very simple DPD to limit latency. Section 2.3.2 justified the choice to include even orders in the predistorters contrary to custom as they tend to reduce the EVM especially at higher input powers into the SOA. The numerical results for the booster SOA case demonstrated that the GMP usually has slightly better

results with little added complexity compared to other known predistorters such as the MP and EMP. Afterwards, a study on the use of a GMP predistorter when increasing data rate was explored. It was found that increasing bandwidth while using DPD could prove beneficial, and the same predistorter can be used to compensate the SOA regardless of the number of subcarriers when the same bandwidth was maintained.

The remaining of the chapter was spent exploring the possibilities for a CO-OFDM transmission with an in-line SOA. First, the different methods of synchronization and equalization were presented. Then, simulation results were presented where it was found that chromatic dispersion does not seem to pose significant additional issues in the eyes of the SOA and a GMP predistorter identified for an optical back to back setup could suffice to meet EVM requirements. Finally, DPD and DPoD were compared and it was determined that DPD works best when the SOA is close to the transmitter, and inversely, DPoD when the SOA is closest to the receiver. DPD tended to give a slightly better quality of transmission overall. The weighted combination of DPD and DPoD according to the fiber length is a possible solution which has not yet been studied and could be an interesting perspective to this work.

Chapter 3

Digital predistortion for semiconductor optical amplifier based coherent optical orthogonal frequency division multiplexing systems: stochastic analysis

---

Chapter 2 served as proof that digital predistortion (DPD) can be used to not only improve the quality of transmission of a coherent optical orthogonal frequency division multiplexing (CO-OFDM) system, but also to increase the immunity to the variability of the input waveform. In this context thus far, the natural randomness of the communication system is not taken into account making the evaluation of the effectiveness of DPD incomplete. Indeed, one cannot ignore the fact that any communication system is affected by some degree of uncertainty due to manufacturing tolerances, environmental conditions, and calibration imperfections which can have a significant impact on overall system performances. Previous work done in the Laboratoire des Sciences et Techniques de l'Information de la Communication et de la Connaissance (Lab-STICC) consisted in conducting a robustness analysis of different predistorters including the memory polynomials (MP), envelope memory polynomials (EMP) [Dio+17]. In that context, the robustness was defined as the ability of the predistorter to still meet the error vector magnitude (EVM) requirements of the transmission when one parameter in the system was changed, the predistorter in this case remaining the same. That study showed that even when one of the parameters related to the semiconductor optical amplifier (SOA) or the IQ modulator drift from its nominal value, DPD continues to be beneficial with some

DPD algorithms performing better than others.

It is possible to model the randomness of the system even more fully via probabilistic models. Although Monte Carlo (MC) simulations are typically used for that purpose [Rub+16], some preliminary work done in the Lab-STICC and in other fields has shown that using polynomial chaos expansion (PCE) helps reduce the computational cost of MC [You19][Man+17a][Kai+18]. This chapter investigates predistorter sensitivity by performing a full stochastic analysis, thus accounting for the simultaneous variation of several system parameters. Contrary to previous work, this analysis will include the study of the generalized memory polynomials (GMP) algorithm, variations from other parameters of the system, and will account for the presence of IQ imbalances, digital to analog conversion (DAC) imperfections, and chromatic dispersion. First, the reliability of PCE in the context of an SOA based CO-OFDM system will be established. Then, two sets of cases will be explored specifically. The first sets of results will deal with an optical back to back (B2B) scenario with a booster SOA. The second sets of results will take the first scenario one step further by including the fiber and DAC imperfections. In both cases, the component uncertainties will be inflicted on the IQ modulator and the SOA.

3.1 Polynomial chaos expansion

As previously mentioned, MC simulations are universally used for statistical analyses. The idea is to sample realizations of an input vector according to its distribution and a computational model is run for each sample to then estimate some statistical property of a random response. To obtain an accurate depiction of these estimates, one would usually need $10^{3–4}$ samples [Sud14]. Knowing that the SOA modeled on Advanced Design System (ADS) used throughout this thesis can take a couple of minutes to run, this would entail extremely long computational times. To avoid this computationally demanding models, some alternative methods need to be employed. The following paragraphs describe the chosen method in this thesis: the PCE.

PCE aims to take a random vector $\boldsymbol{X} \in \mathbb{R}^N$ composed of independent components with specific probability distributions and model their outcome, a random variable Y, assumed to have finite variance through the polynomial expansion [Sud14]

$$Y = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^A} z_{\boldsymbol{\alpha}} \Psi_{\boldsymbol{\alpha}}(\boldsymbol{X}) \tag{3.1}$$

where A designates the number of considered input variables, $z_{\boldsymbol{\alpha}}$ are the coefficients related to the multivariate orthonormal polynomials $\Psi_{\boldsymbol{\alpha}}(\boldsymbol{X})$ which form a basis of the Hilbert space associated with the ordered lists of integers $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_A)$ through

$$\Psi_{\boldsymbol{\alpha}}(\boldsymbol{X}) = \prod_{j=1}^{A} \psi_{\alpha_j}(X_j) \tag{3.2}$$

where ψ_{α_j} are the univariate polynomials chosen according to the distribution of X_j. Table 3.1 shows the advised polynomials for a variable with different types of classical distributions.

Table 3.1 – Input distributions with their associated orthogonal polynomials

Distribution type	Orthogonal polynomials
Uniform	Legendre
Gaussian	Hermite
Gamma	Laguerre
Beta	Jacobi

To have an exact representation of the random response, infinite series would be used. However, in real-life applications, it is necessary to reduce the space $\mathbb{N}^A$ to a set $\mathcal{A}$. The computational model $\mathcal{M}$ then yields

$$Y \approx \mathcal{M}(\boldsymbol{X}) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}} z_{\boldsymbol{\alpha}} \Psi_{\boldsymbol{\alpha}}(\boldsymbol{X}) \tag{3.3}$$

$\mathbb{N}^A$ is reduced to a set $\mathcal{A}$ by limiting the total degree of a multivariate polynomial $|\alpha| = \sum_{j=1}^{M} \alpha_j$ to a maximum degree p (usually of degree 3 to 5). Examples of the ordered lists if $p = 2$ and $A = 1$, and if $p = 3$ and $A = 2$ are shown in Table 3.2.

The standard truncation scheme thus yields

$$\boldsymbol{\mathcal{A}} = \left\{ \boldsymbol{\alpha} \in \mathbb{N}^A : |\alpha| \leq p \right\} \tag{3.4}$$

The number of terms belonging to this set $\mathcal{A}$ is as follows:

$$\operatorname{card} \mathcal{A} = \frac{(A+p)!}{A!p!} \tag{3.5}$$

Other truncation schemes serve to reduce the computational cost further such as the

Table 3.2 – Ordered lists for $p = 2$ and $A = 1$, and $p = 3$ and $A = 2$. p is the maximum polynomial degree and A is the number of uncertain parameters

j	$p = 2,\ A = 1$	$p = 3,\ A = 2$
0	$\boldsymbol{\alpha_0} = (0)$	$\boldsymbol{\alpha}_0 = (0,\ 0)$
1	$\boldsymbol{\alpha_1} = (1)$	$\boldsymbol{\alpha}_1 = (1,\ 0)$
2	$\boldsymbol{\alpha_2} = (2)$	$\boldsymbol{\alpha}_2 = (0,\ 1)$
3	-	$\boldsymbol{\alpha}_3 = (2,\ 0)$
4	-	$\boldsymbol{\alpha}_4 = (1,\ 1)$
5	-	$\boldsymbol{\alpha}_5 = (0,\ 2)$
6	-	$\boldsymbol{\alpha}_6 = (3,\ 0)$
7	-	$\boldsymbol{\alpha}_7 = (2,\ 1)$
8	-	$\boldsymbol{\alpha}_8 = (1,\ 2)$
9	-	$\boldsymbol{\alpha}_9 = (0,\ 3)$

hyperbolic truncation scheme [Bla+11]. The idea is to select $\mathcal{A}$ such that the multi-indices are of q-norm less than or equal to p with $0 < q \leq 1$. A visual illustration of that truncation scheme is shown in Figure 3.1 for $A = 2$ hence the 2-dimensional graphs. As can be seen from Figure 3.1, the hyperbolic truncation scheme with $q = 1$ is equivalent to the standard truncation scheme.

The coefficients can be calculated either by means of orthogonal projection or least-square minimization. The latter is chosen for this thesis; its definition is given by:

$$\hat{\boldsymbol{z}} = \underset{\boldsymbol{z} \in \mathbb{R}^{\mathrm{card}\,\mathcal{A}}}{\operatorname{argmin}}\, \mathbb{E}\left[\left(Y - \sum_{\boldsymbol{\alpha} \in \mathcal{A}} z_{\boldsymbol{\alpha}} \Psi_{\boldsymbol{\alpha}}(\boldsymbol{X})\right)^2\right] \tag{3.6}$$

Experimentally, $\boldsymbol{\mathcal{X}} = \left\{x^{(1)}, \ldots, x^{(n)}\right\}$ is a sample set of points (particular realizations of X) that can be obtained through different sampling schemes. In order to estimate the coefficients, the communication chain is simulated for each point in $\boldsymbol{\mathcal{X}}$ and stored in a vector $\boldsymbol{\mathcal{Y}} = \left\{y^{(1)}, \ldots, y^{(n)}\right\}$. Note that as a rule, the size of $\boldsymbol{\mathcal{X}}$ should be at least two to three times bigger than the number of terms in the series of Equation 3.3.

It is also possible to add sparsity to the problem notably in the case of high degree and high number of parameters of Equation 3.1 and Equation 3.2. A penalty term $\zeta||z||_1$, where $||z||_1 = \sum_{\boldsymbol{\alpha} \in \mathcal{A}} |z_{\boldsymbol{\alpha}}|$, can then be added to Equation 3.6 to give priority to low rank solutions.

Once the coefficients $z_{\boldsymbol{\alpha}}$ are found and the PCE model $\mathcal{M}$ is constructed, it is possible to get the probabilistic content of the quantity of interest (QoI). The PCE presented so

far made the simplifying assumptions that the input variables are independent and that the QoI is a scalar, but the algorithm has been successfully tested with correlated input variables and a vector as the QoI [Sud14][Waq+21].

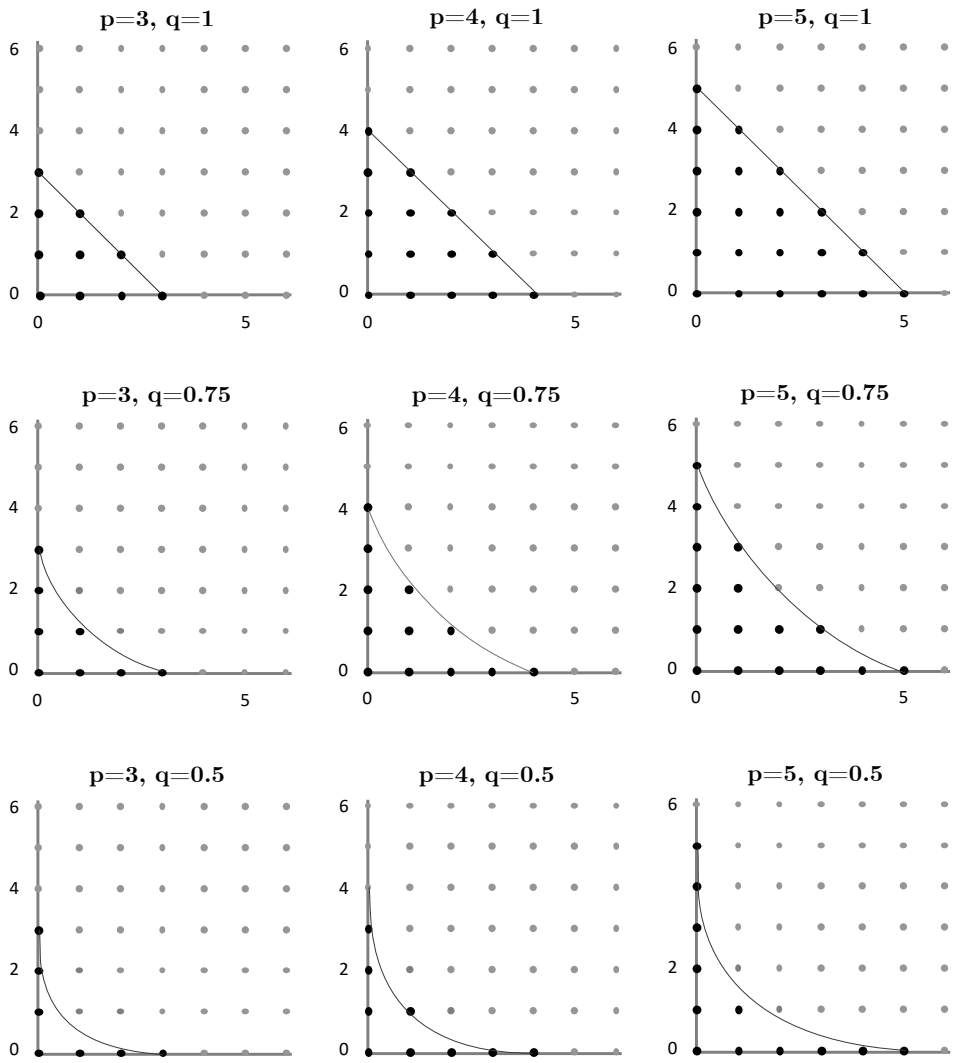


Figure 3.1 – Example of a hyperbolic truncation scheme for $A = 2$, degree $p = 3, 4, 5$, and $q = 0.5, 0.75, 1$. The black dots are the indices of the polynomials kept for the model

3.2 Polynomial chaos expansion for optical transmission applications

Throughout this thesis work, the PCE was implemented using the MATLAB UQLab toolbox [Mar+14]. Compared to other methods, this process is considered non-intrusive

i.e. MC simulations and the PCE model estimation are introduced as an outside layer which envelops the already existing simulation code. An example of that non-intrusive process is depicted in Figure 3.2 for $A = 3$. First, the three uncertain variables will

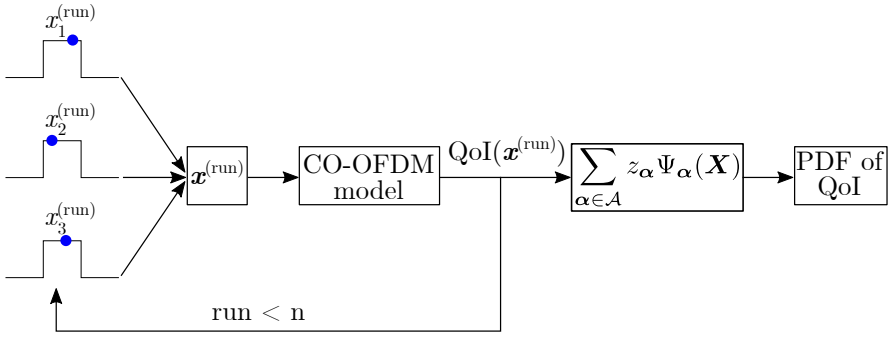


Figure 3.2 – Example of a PCE setup for optical communications. n is the total number of runs

be assumed to have uniform distribution since there is no a priori knowledge on them. According to Table 3.1, Legendre polynomial bases should be used in that case. Table 3.3 shows the advised polynomials for the first uncertain variable x_1. To obtain the multivariate polynomials, one would simply need to multiply the univariate polynomials together.

Table 3.3 – Polynomials for an uncertain variable with uniform distribution with $p = 3$

ψ_{α_1}	$\psi_0 = 1$	$\psi_1 = \frac{X_1}{\sqrt{\frac{1}{3}}}$	$\psi_2 = \frac{3X_1^2-1}{2\sqrt{\frac{1}{5}}}$	$\psi_3 = \frac{5X_1^3-3x_1}{2\sqrt{\frac{1}{7}}}$

As an example, an optical communication system could have a chosen uncertain variable be the IQ modulator's peak-to-peak voltage V_{pp} or the SOA's bias current I_{bias}. Note that introducing uncertainties in these variables goes beyond their external uncertainty and can also mimic variability induced in component manufacturing or a change in the environment. For instance, a change in the SOA's active region length or a change in temperature affects the SOA gain; similarly so, a change in I_{bias} will also affect the SOA gain.

The experimental design points are obtained through a sampling scheme. There are different types of sampling including MC, latin hypercube sampling (LHS) [McK+00], and quasi-random sequences like Sobol or Halton sequences [Nie92]. The LHS is a statistical method found to be computationally cost efficient and representative of the real variability

of a system as it attempts to visit all regions of a cumulative density function that has been divided into equiprobable regions. Figure 3.3 illustrates that point by comparing LHS to a random sampling process for a two-dimensional problem.

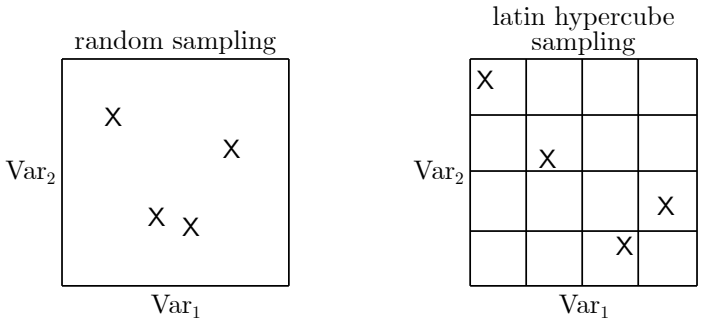


Figure 3.3 – Example of LHS vs. random sampling in a 2-D scenario

The CO-OFDM model is run using the sampled values in $\boldsymbol{x}^{(run)}$ then a QoI that is associated to these sampled values is outputted. The QoI could be the bit error rate (BER) or the EVM for example. After the n^{th} run, the PCE model can be estimated through sparse least angle regression (LARS) [Efr+04]. LARS is an algorithm designed to minimize

$$\hat{\boldsymbol{z}} = \underset{\boldsymbol{z}\in\mathbb{R}^{\text{card}\mathcal{A}}}{\operatorname{argmin}}\, \mathbb{E}\left[\left(Y - \sum_{\boldsymbol{\alpha}\in\mathcal{A}} z_{\boldsymbol{\alpha}}\Psi_{\boldsymbol{\alpha}}(\boldsymbol{X})\right)^2\right] + \zeta||z||_1 \tag{3.7}$$

The LARS algorithm is described as such:

1. Initialization
   - $z_{\boldsymbol{\alpha}} = 0,\ \forall \boldsymbol{\alpha} \in \mathcal{A}$
   - Candidate set which contains all of the regressors Ψ_α
   - Active set which contains the regressors whose coefficients will be updated in the current iteration: $\varnothing$
   - Residual: $\boldsymbol{r} = \boldsymbol{\mathcal{Y}} - \sum_{\boldsymbol{\alpha}\in\mathcal{A}} z_{\boldsymbol{\alpha}}\Psi_{\boldsymbol{\alpha}}(\boldsymbol{\mathcal{X}}) = \boldsymbol{\mathcal{Y}}$
2. Find the regressor $\Psi_{\alpha_j}\,(X_j)$ that is most correlated with the current residual $\boldsymbol{r}$ and move it to the active set
3. Move the coefficients of the active set towards their least-square value until another regressor from the candidate set is as correlated to the residual $\boldsymbol{r}$ as the ones in the active set. That regressor will be the most correlated to the residual in the next iteration

4. Repeat steps 2-3 until the size of the active set is equal to $\min(P, n)$, $P = \operatorname{card} \mathcal{A}$ being the polynomial basis size and n being the experimental design size

Once the model has been successfully estimated, the probabilistic content of the QoI can be studied. In this thesis, we looked into observing the probability density function (PDF) of the QoI. The following sections go over the reliability of the PCE in the context of CO-OFDM systems and what can be inferred from DPD after having computed the PDF of the EVM.

3.3 Stochastic analysis of digital predistortion for a booster SOA

The scenario considered in this section is the same as the one in Section 2.4.1. Figure 3.4 is a more visually compact version of that scenario which also showcases the chosen uncertain parameters in bold. The input bit stream undergoes orthogonal frequency divi-

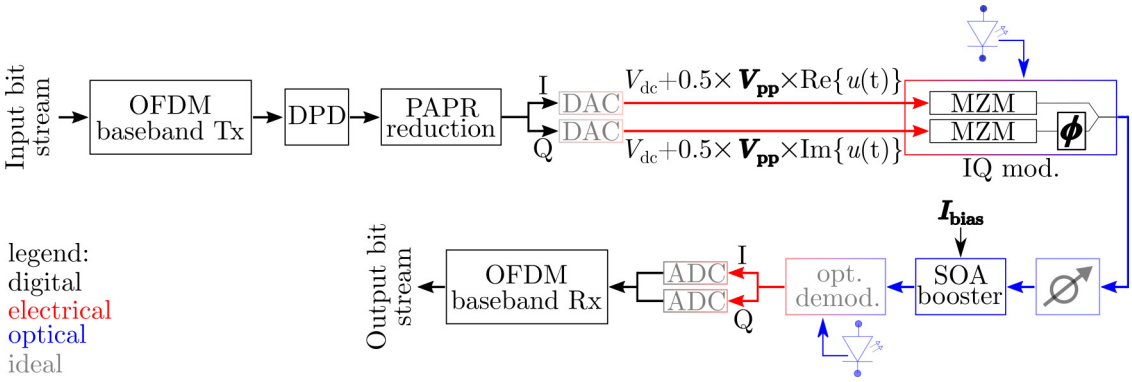


Figure 3.4 – Optical B2B CO-OFDM transmission system with a booster SOA with uncertain variables V_{pp}, ϕ, and I_{bias}. The DACs, opto-electronic conversion, and lasers are considered ideal

sion multiplexing (OFDM) modulation then DPD is applied. The predistorders are those learned in the optical B2B scenario when the input power into the SOA is equivalent to its value when there are no uncertainties applied. After DPD, the signal is hard-clipped at a 6 dB ratio before getting converted to the electrical domain by an ideal DAC (computer's precision). The IQ modulator used for the electro-optic conversion is one of the components which will have uncertain parameters: both Mach-Zehnder modulators (MZMs) peak-to-peak amplitude V_{pp} and the phase difference ϕ between each MZM branch. The

other component which will have uncertainties is the booster SOA with its bias current I_{bias} being the uncertain parameter. The conversion from the optical domain to the digital domain is considered ideal for now as well. Then the signal is demodulated. The OFDM baseband receiver includes the synchronization and the common gain and phase compensation. In this transmission both the fiber and the laser are ideal thus there will not be any chromatic dispersion (CD) or phase noise to take into account. Also note that the input power into the SOA may vary when the uncertain parameters change.

Table 3.4 shows the parameters used for this analysis. The parameters are the same as the ones used in Section 2.4.1. One new case with a modulation order of $M = 16$ is looked into. The uncertain parameters are of the following nominal values: $V_{pp} = 1.25 \times V_\pi$, $\phi = \frac{\pi}{2}$, and $I_{bias} = 150$ mA. The following cases are scrutinized for a fluctuation in V_{pp}, ϕ, I_{bias}: no DPD, MP, EMP, and GMP DPDs. The percent uncertainty applied depends on the modulation order. For the 4-quadrature amplitude modulation (QAM) case, the percent uncertainty σ_{PCE} will take values 5%, 10%, and 15 %. For the 16-QAM case which calls for more stringent material requirements, σ_{PCE} will be 5%.

Table 3.4 – Parameters used in the CO-OFDM simulation

Parameter	Symbol	Value	Units
Modulation order	M	4 or 16	-
Number of subcarriers	N_{sc}	128	-
Bandwidth	ΔB	5	GHz
Cyclic prefix factor	ρ_{cp}	$\frac{1}{8}$	-
Total number of QAM data symbols sent	N_{data}	3180	-
Oversampling factor	ρ_{os}	4	-
MZM peak-to-peak voltage	V_{pp}	$1.25 \times V_\pi \pm \sigma_{PCE}$	volts
IQ modulator phase difference	ϕ	$\frac{\pi}{2} \pm \sigma_{PCE}$	radians
Laser wavelength	λ_{LO}	1540	nm
SOA bias current	I_{bias}	$150 \pm \sigma_{PCE}$	mA

3.3.1 Reliability analysis of the polynomial chaos expansion

In Chapter 2, the simulation results showed that the variablity of the frame influences the EVM values. Therefore, the same frame constituted of 3 180 QAM data symbols is used throughout the whole analysis to keep the results meaningful and avoid introducing undesired uncertainty due to the OFDM signal itself. An investigation will be presented

later to confirm that the conclusions presented hereafter hold regardless of the emitted sequence.

A reliability analysis is first performed on the no DPD case for a 4-QAM signal sent at a reference input power for the SOA $P_{ref} = -14$ dBm. A PCE model for the EVM is first constructed from 50 simulation runs. Then, the EVM values of 10^4 MC runs and the values of the PCE model evaluated using the same experimental design are compared as depicted in Figure 3.5. That figure plots the EVM values obtained by PCE vs. the EVM values obtained by MC. The "ideal match" red line corresponds to a scenario where the obtained MC EVM values are exactly the same as the PCE EVM values. The blue points all congregating around the red line indicate that the PCE estimates the model extremely well. It can be seen from Table 3.5 that PCE saves 200 times more simulation time

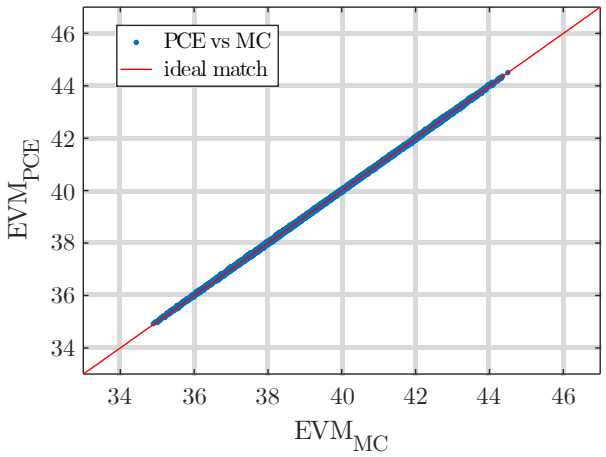


Figure 3.5 – Polynomial chaos expansion constructed from 50 simulations comparison with 10^4 Monte Carlo simulations for the no DPD case 4-QAM configuration at $P_{ref} = -14$ dBm.

compared to MC and once the PCE model has been constructed, the time to evaluate 10^4 points is negligible.

Before moving onward, a comprehensive analysis with a range of SOA operating conditions and a different OFDM configuration in order to stress the system sufficiently is carried out to further solidify the effectiveness of PCE compared to MC in optical transmissions. Figure 3.6a and Figure 3.6b compare PCE with MC at -17 dBm and -11 dBm respectively for a 5% uncertainty on all parameters. As was shown with the case at -14 dBm, there is an excellent congruence between the EVM values obtained from PCE and those obtained from MC.

Table 3.5 – PCE vs. MC simulation times

	No of runs	Total time
Single run	1	141.5 s
MC analysis	10^4	1,415,000 s (16.4 days)
PCE model construction	50	7075 s (2 h)
PCE analysis	10^4	0.5 s
PCE vs MC speedup ratio	200	

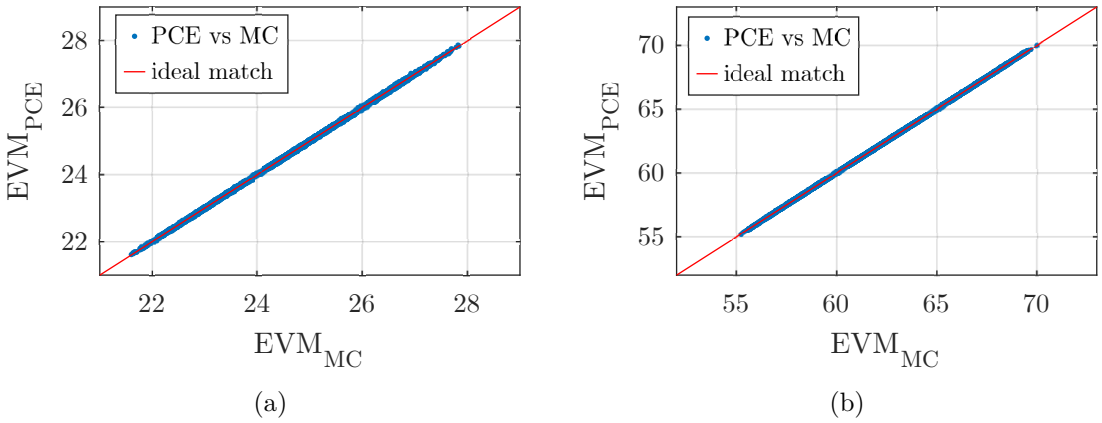


Figure 3.6 – Polynomial chaos expansion constructed from 50 simulations comparison with Monte Carlo simulations for the no DPD case 4-QAM configuration at (a) $P_{ref} = -17$ dBm (b) at $P_{ref} = -11$ dBm

Through Figure 3.7, it is demonstrated that even with a different OFDM configuration using $N_{sc} = 512$ and $\Delta B = 20$ GHz instead of $N_{sc} = 128$ and $\Delta B = 5$ GHz, the model gives an accurate depiction of the EVM values for an uncertainty of 5% on all uncertain parameters. The scenario with $N_{sc} = 512$ and $\Delta B = 20$ GHz will be explored more thoroughly in section 3.4.

The last step of this study serves to show the validity of PCE by providing evidence that there is a good match with MC simulations at a wider range of EVM values. Figure 3.8 shows the following scenarios: GMP with 15% uncertainty, MP with 10% uncertainty, and no DPD with 5% uncertainty. Increased parameter tolerances intrinsically varies the operating conditions of the SOA. According to our simulations, inserted deviations of 5% to 15% can induce deviations in the input optical power of up to 10% to 25% respectively. Note that for higher percent uncertainties, it is necessary to increase the order of the PCE models. For instance, while the no DPD case with 5% uncertainty required a maximum

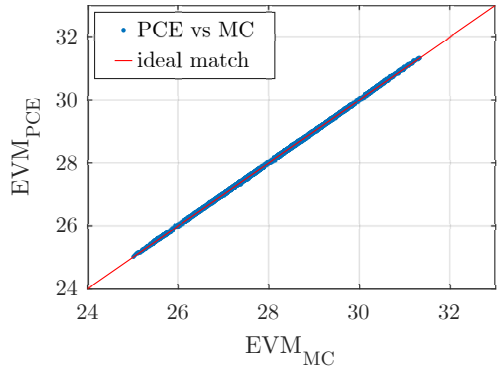


Figure 3.7 – Polynomial chaos expansion constructed from 50 simulations comparison with Monte Carlo simulations for the no DPD case 4-QAM configuration at $P_{ref} = -14$ dBm with a signal configuration of $N_{sc} = 512$ and $\Delta B = 20$ GHz

polynomial order of $p = 3$, a maximum polynomial order of $p = 5$ would be recommended for the GMP case at 15% uncertainty. Both of these values for the parameter p are considered low compared to other applications where PCE has been tested [Sud14].

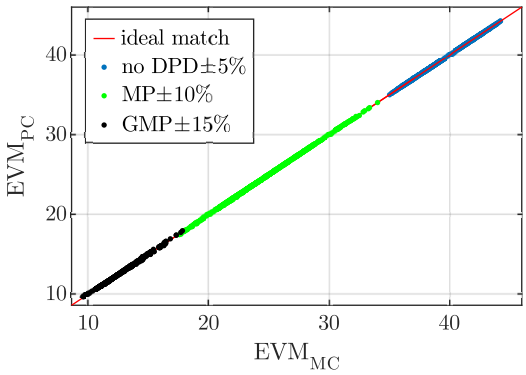


Figure 3.8 – Polynomial chaos expansion constructed from 50 simulations comparison with Monte Carlo simulations for the no DPD, MP (10% uncertainty), and GMP (15% uncertainty) cases with a 4-QAM configuration at $P_{ref} = -14$ dBm

As a sidenote, the PCE is constructed from 50 simulations to make sure the estimated model is as accurate as possible. However, it is possible to lower the number of simulations to 20 and still have a reliable model thus further increasing the 200 speedup ratio (c.f. Figure 3.9).

It has been shown that the PCE model mirrors the MC representation very well and is therefore reliable in further analysis. Thus, for the upcoming results, the PCE models are validated as well but on a smaller scale.

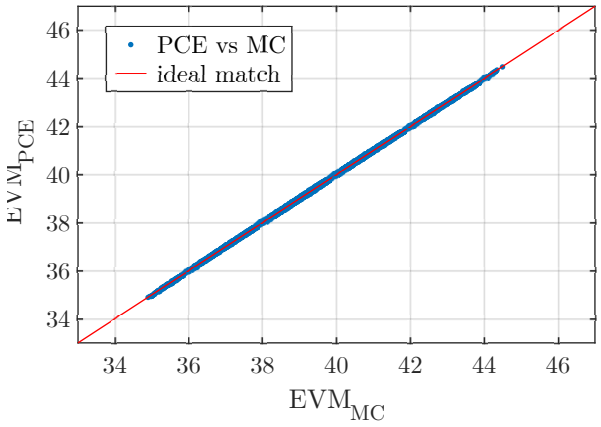


Figure 3.9 – Polynomial chaos expansion constructed from 20 simulations comparison with Monte Carlo simulations for the no DPD case 4-QAM configuration at $P_{ref} = -14$ dBm

3.3.2 Simulation results of a stochastic analysis of DPD for a booster semiconductor optical amplifier

Once the PCE scheme has been properly validated, PDFs are then computed with the PCE models of all cases (no DPD, MP, EMP, GMP) at different levels of uncertainties and are shown in Figure 3.10.

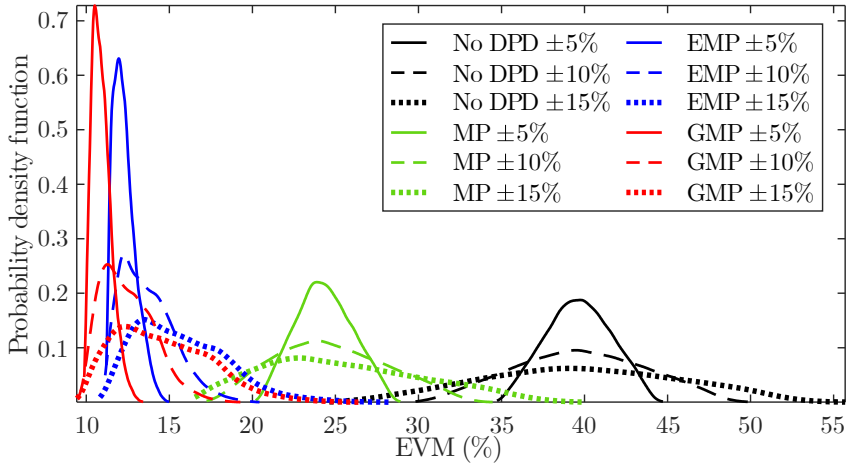


Figure 3.10 – 4-QAM probability density function of the EVM at $P_{ref} = -14$ dBm for different predistorters

There are three noteworthy observations to draw: a horizontal translation, a vertical stretch, and a horizontal stretch. The left horizontal translation proves that the aver-

age EVM improves with DPD with the EMP and GMP predistorters meeting an EVM objective of 30% regardless of the percent uncertainty. This entails that a higher input power to the SOA could be used signifying more reach. The vertical stretch illustrates the improvement in terms of variance with the steepest curve demonstrating more robustness to the uncertainties. All predistorters show improvements with respect to both. The horizontal stretch shows that as the percent uncertainty increases, there is more spreading on the PDFs. In some cases, most particularly the MP, this can become an issue because the 30% EVM limit is exceeded when the level of uncertainty becomes too great (from 10% uncertainty). In any case, the spreading tends to occur at the higher EVM values meaning the average EVM becomes higher as the level of uncertainty increases. Note that the better the DPD, the closer the output EVM is to a minimum threshold mostly caused by the SOA amplified spontaneous emission (ASE) noise and the hard-clipping; these effects are highlighted by the asymmetry seen for EMP and GMP.

Figure 3.11 confirms these observations by looking at a 2D-histrogram of the superposed QAM symbols received after 100 simulations this time by sending different frames. The population density of the QAM symbols is color coded from dark to bright (lowest to highest density). While the constellation is scattered at first, it is more concentrated in the GMP case.

It can be seen from the same figure that EMP shows some phase noise in its constellation. This phenomenon seems consistent with the EMP structure and will be explained next. Figure 3.12 shows the constellations for all scenarios in blue. In red and green are the QAM symbols corresponding to the two baseband high-frequency subcarrier extremities in the OFDM symbols (the red dots are for the negative frequency subcarriers and the green for the positive frequency subcarriers).

The SOA does not behave the same at all frequencies since there is a change in phase and amplitude away from the reference QAM symbol seen for the high-frequency subcarriers. However, in the no DPD case, this change in phase and amplitude is masked in the final constellation in blue since they end up looking Gaussian when combined.

As far as the MP is concerned, that same change in phase and amplitude is corrected but not fully and the effects combined look like a Gaussian as well. The GMP is able to fix both the phase and amplitude issue as all QAM symbols are centered around the point of reference.

It can be assumed that the EMP problem is intrinsic to its structure as it only takes into account the envelope of the signal compared to MP and GMP. In that regard, the

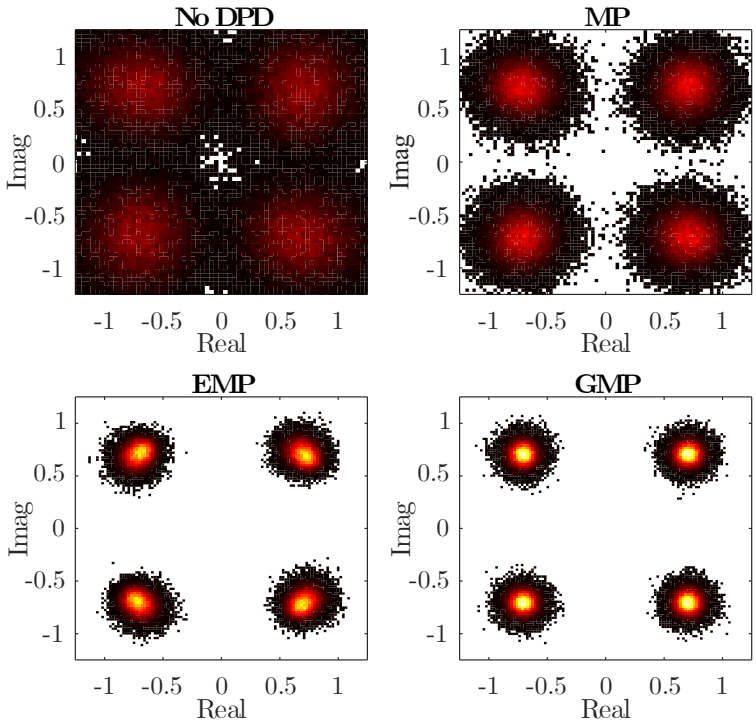


Figure 3.11 – 2D-histogram of a 4-QAM constellation at $P_{ref} = -14$ dBm from 100 simulation runs

EMP is designed to correct amplitude modulation and phase modulation caused by the module of $x[n]$; however, it does not made to properly correct amplitude modulation and phase modulation caused by the phase of $x[n]$. We therefore notice that the EMP has difficulty compensating the phase shift introduced by the SOA.

In this regard, it seems wiser to choose the GMP as a predistorter for this scenario as it provides the better improvement both in terms of global quality of transmission and robustness to uncertainties. It also takes into account the phase of $x[n]$ and better corrects the phase shift introduced by the SOA while staying relatively low in complexity (16 coefficients vs. 15 and 6 for the EMP and MP respectively).

Another uncertainty analysis is performed for a 16-QAM signal with $P_{ref} = -17$ dBm. An acceptable EVM for that scenario is in the 16% range [Sch+11] therefore it is safer to lower the SOA input power although it also lowers the reach. Higher modulation formats call for stricter material requirements so a $\pm 5\%$ uncertainty seems reasonable. Figure 3.13 shows the PDF computed for the no DPD, MP, EMP, and GMP predistorters. On Figure 3.13, it is observed that the MP is not able to consistently pass the threshold

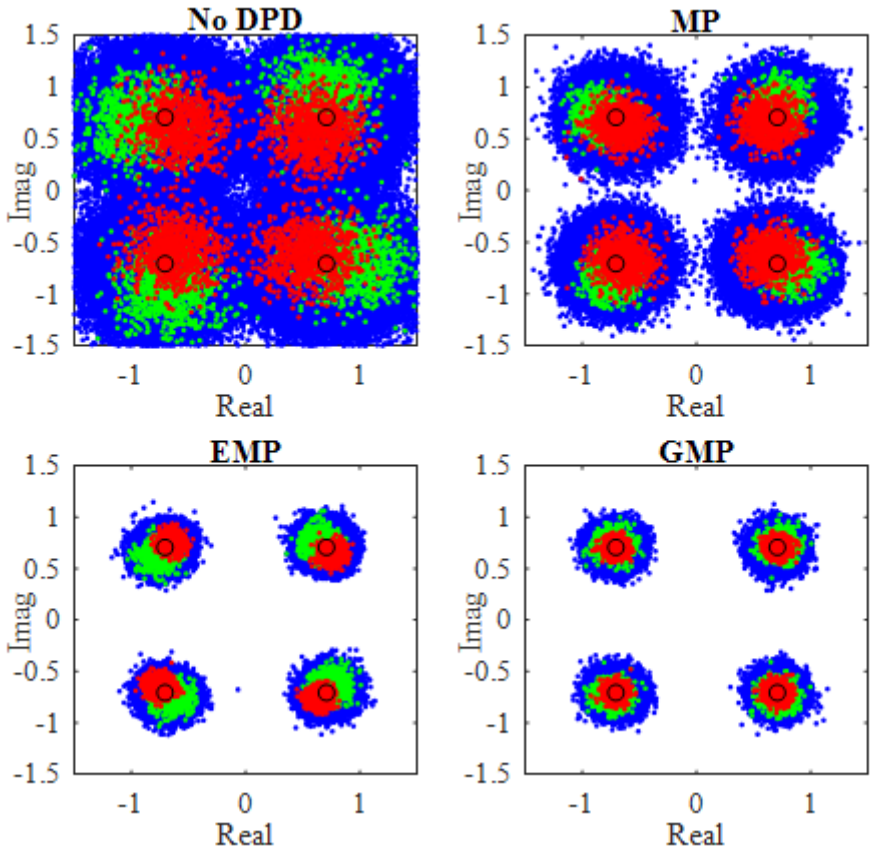


Figure 3.12 – 4-QAM constellation at $P_{ref} = -14$ dBm of 100 simulation runs. The blue shows all symbols. The red and green show the QAM symbols assigned to the negative and positive high frequency subcarriers respectively

set for an acceptable transmission, but the GMP and the EMP meet the requirements. Figure 3.13 also reiterates the results found in the 4-QAM case as the GMP improves both the average EVM and the robustness to the uncertainties to the best extent.

In Figure 3.14, the superposed constellation from 100 simulation runs is depicted for the 16-QAM case. The GMP is shown to have the better QAM symbol distribution and concentration.

The same phenomenon can be spotted in the EMP case where dephasing appears on most outer rings of the constellation. The most inner ring not displaying that same phase shift could be an indication in how much the intensity associated to a QAM symbol plays a role in the subsequent observed phase shift induced by the SOA. Going from the inner, to the middle, to the outer circle, it is seen that the phase shift appears in the midcircle

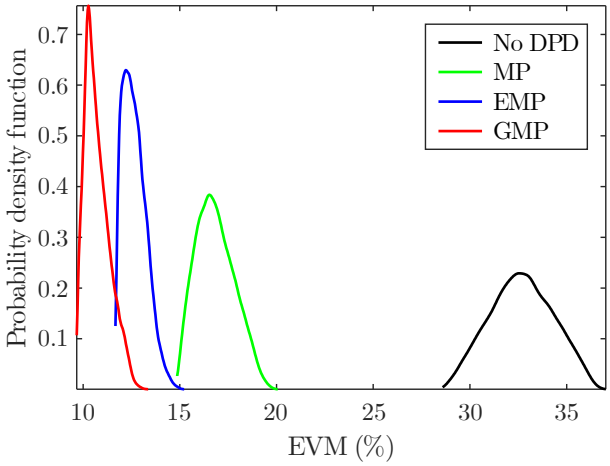


Figure 3.13 – 16-QAM probability density function of the EVM at $P_{ref} = -17$ dBm for different predistorters

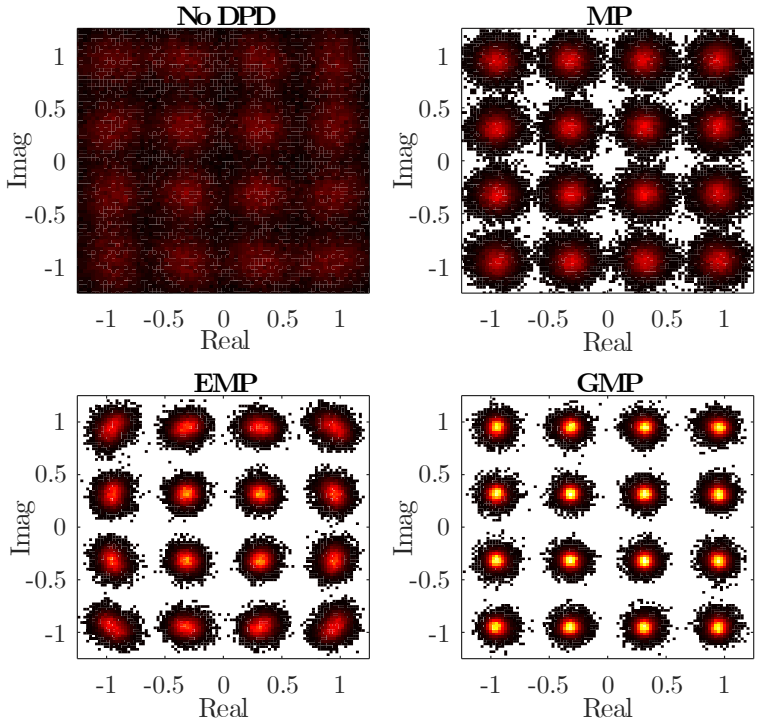


Figure 3.14 – 2D-histogram of a 16-QAM constellation at $P_{ref} = -17$ dBm from 100 simulation runs

then increases gradually.

Regardless, the GMP would be the predistorter of choice in this scenario as well using

14 coefficients and making it reasonably low in complexity for its better results compared to EMP (12 coefficients), and MP (6 coefficients).

At the beginning of the analysis, it was stated that the same frame was sent down the optical channel to avoid adding bias to the stochastic analysis due to the variability of the frame. Another analysis was performed to verify that the conclusions previously drawn from the stochastic analysis are still valid when using a different seed to construct the frame. Figure 3.15 depicts the PDF of the EVM for the four cases that have been studied (no DPD, MP, EMP, and GMP) using different frames. It can be deduced from that figure that although a different frame can yield slightly different EVM results, the trends in terms of average EVM and resistance to uncertainties still hold.

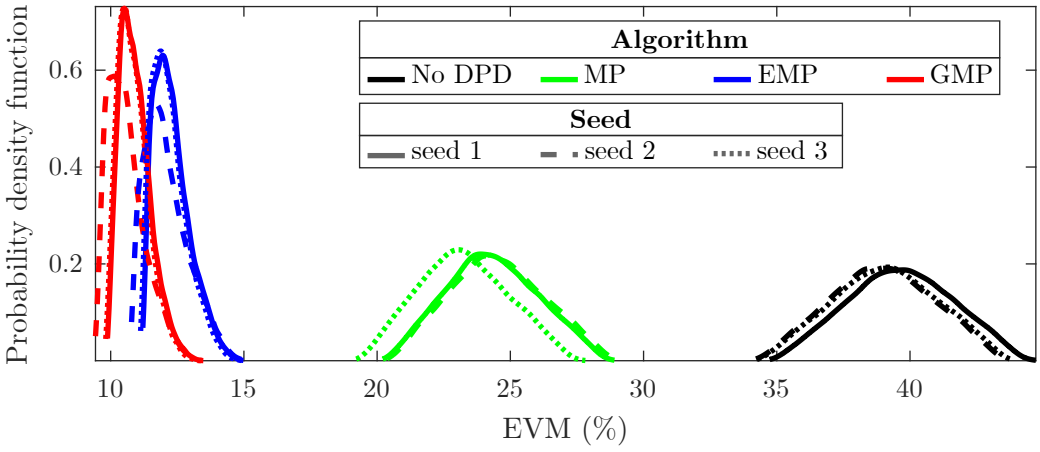


Figure 3.15 – 4-QAM probability density function of the EVM at $P_{ref} = -14$ dBm for different predistorters and seeds at 5 % uncertainty

3.4 Stochastic analysis of digital predistortion for a transmission with fiber

In section 3.3, it was proven that DPD is capable of improving the global quality of a system in addition to increasing the immunity to uncertainties intrinsic to the system. In this section, we will explore DPD subjected to the combined uncertainties of IQ imbalances (a current research focus [Far+17]) in the presence of CD and DAC imperfections.

The following simulations are based on the scenario depicted in Figure 3.16. A bit stream undergoes M-QAM modulation before being coded into 6 OFDM data symbols

(2 844 QAM data symbols) of bandwidth $\Delta B = 20$ GHz. The OFDM frame contains 8 OFDM symbols in total including the synchronization and equalization preamble compared to 32 OFDM symbols in one frame for the case with $N_{sc} = 128$ and $\Delta B = 5$ GHz of Section 3.3. That change in the number of OFDM symbols in the OFDM frame is done so that the length of the time signal stays the same (18432 samples). After accounting for the null subcarriers and the cyclic prefix, the data rate is 27.8 Gbit/s or 55.6 Gbit/s for a 4-QAM and 16-QAM configuration respectively. Some hard-clipping at a 8 dB ratio is done on the constant length time signal which is sent through a DAC with an effective resolution of 6 bits. This restriction in the quantification level of the DAC will add some quantization noise to the system. The electro-optic conversion includes an IQ modulator that has a peak-to-peak voltage of $1.25 \times V_\pi$. The channel configuration is 20 km of fiber followed by the SOA followed by 80 more km of fiber. At the receiver, the signal undergoes synchronization, common gain and phase compensation, equalization (using the equalization preamble in the frequency domain), and demodulation. Note that although the scenario that was just described is more complete, it still assumes an ideal laser and opto-electronic conversion. The nonlinearities of the fiber are not considered either although there is still an insight into the issue of four wave mixing (FWM) as it is a phenomenon that is present in the SOA as well. The predistorters used to linearize the SOA are identified in an optical B2B setup (fiber excluded) with no uncertainty then are applied with the fiber and uncertainties added.

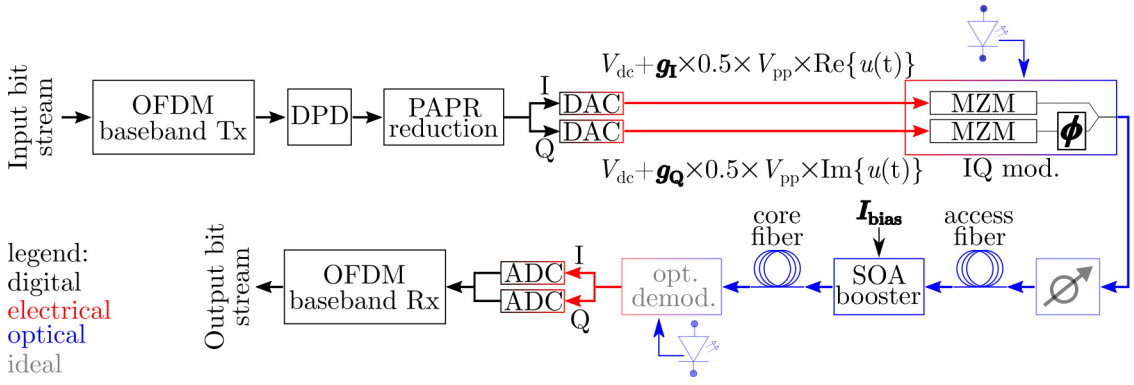


Figure 3.16 – CO-OFDM transmission with the SOA, fiber, and DAC imperfections. The uncertain variables are g_I, ϕ, and I_{bias} with $g_Q = 2 - g_I$. The opto-electronic conversion and the lasers are considered ideal

Figure 3.16 also displays the parameters featuring the performance-limiting uncertainties in the system: the modulator's gains (g_I and g_Q) and phase difference (ϕ), and the

SOA's bias current (I_{bias}). Their nominal values are 1 (where $g_Q = 1 - (g_I - 1)$), $\frac{\pi}{2}$, and 150 mA respectively.

The discussed scenario has parameters as summarized in Table 3.6. This new table adds in the uncertain parameter g_I and the DAC resolution.

Table 3.6 – Parameters used in the CO-OFDM simulation for a case with added fiber and DAC imperfections

Parameter	Symbol	Value	Units
Modulation order	M	4 or 16	-
Number of subcarriers	N_{sc}	512	-
Bandwidth	ΔB	20	GHz
Cyclic prefix factor	ρ_{cp}	$\frac{1}{8}$	-
Total number of QAM data symbols sent	N_{data}	2844	-
Oversampling factor	ρ_{os}	4	-
DAC resolution	b_{dac}	6	-
MZM gain	$g_I \pm \sigma_{PCE}$	1	-
MZM peak-to-peak voltage	V_{pp}	$1.25 \times V_\pi$	volts
IQ modulator phase difference	ϕ	$\frac{\pi}{2} \pm \sigma_{PCE}$	radians
Laser wavelength	λ_{LO}	1540	nm
SOA bias current	I_{bias}	$150 \pm \sigma_{PCE}$	mA

For this study, it was deemed valuable to examine cases where the uncertainties were set at 5% as previously done, but also explore cases where the assigned uncertainties were picked in such a way that the influence of each uncertain variable had a fairly equitable impact on the QoI (the EVM). To achieve this, a global sensitivity analysis (GSA) which is a way to assess parameter sensitivity is conducted. There are different methods available in the literature [Sal+08][Kuc+12]; the method of choice here is a variance based method called the Sobol' analysis also known as analysis of variance (ANOVA) [Sob93]. Assuming independent input variables with uniform distribution, the Sobol' decomposition of a response Y reads:

$$Y = f(X_1, \ldots, X_N) = f_0 + \sum_{j=1}^{N} f_j(X_j) + \sum_{1 \le j < k \le N} f_{jk}(X_j, X_k) + \ldots + f_{1,2,\ldots,N}(X_1, \ldots, X_N) \tag{3.8}$$

The first term f_0 is a constant that is equal to the expected value of $f(\boldsymbol{X})$. With the assumption of orthogonality, the variance of Y can be written as:

$$\text{Var}(Y) = \sum_{j=1}^{N} D_j + \sum_{1 \le j < k \le N} D_{jk} + \ldots + D_{1,2,\ldots,N} \tag{3.9}$$

where the first summation essentially designates the variance due to the effects of each uncertain variable X_j and the second summation represents the interaction between the uncertain variables X_j and X_k.

For a Sobol' analysis, two sets of measures can be obtained: the first order Sobol' index and the total Sobol' index. We will focus on the first order Sobol' index which gives a measure of the relative contribution of each uncertain variable and is defined like so:

$$S_j = \frac{D_j}{\mathrm{Var}\,(Y)} \tag{3.10}$$

The Sobol' indices are usually obtained through MC, but the UQLab toolbox allows to bypass that by computing the PCE based Sobol' indices during post-processing [Sud08]. Essentially, once the PCE model $\mathcal{M}$ and the coefficients $z_{\boldsymbol{\alpha}}$ are estimated, the coefficients are simply gathered according to the dependency of each basis polynomial, square-summed then divided by the total variance. For instance, let us take 3 uncertain parameters I_{bias}, g_I and ϕ of partial variance D_1, D_2 and D_3 respectively, and $p = 3$. If we want to obtain the first Sobol' index S_1 of I_{bias}, only the coefficients associated to the polynomial basis where $\boldsymbol{\alpha}$ is $(1,0,0)$, $(2,0,0)$, and $(3,0,0)$ will be taken, squared, and summed to obtain D_1. Then S_1 can be calculated such that $S_1 = \frac{D_1}{D_1+D_2+D_3}$.

For a 4-QAM scenario at $P_{ref} = -14$ dBm, Figure 3.17a looks at the influence of g_I, ϕ, and I_{bias} on the EVM by computing each uncertain variable's first order Sobol' index when imposing an uncertainty of 5% on all variables. It is interesting to note that I_{bias}

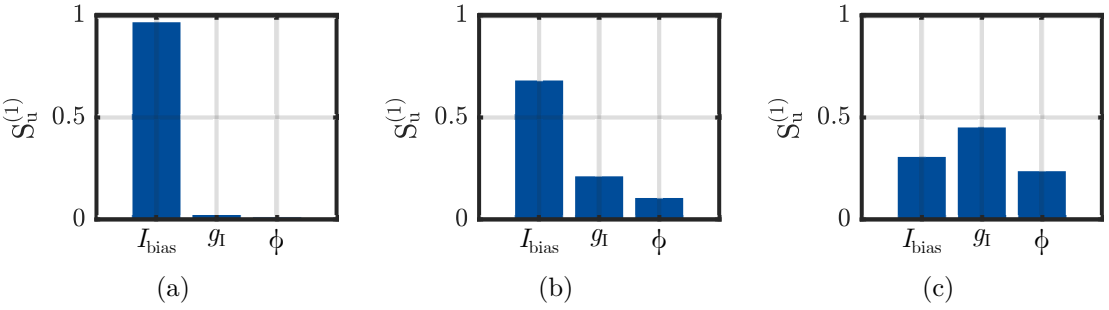


Figure 3.17 – 4-QAM configuration at $P_{ref} = -14$ dBm first order Sobol' index of uncertain variables I_{bias}, g_I, and ϕ at 5% uncertainty for I_{bias} and (a) at 5% uncertainty (b) at 10% uncertainty (c) at 15% uncertainty for g_I and ϕ

has the most influence by far in the performance of the CO-OFDM transmission when faced with uncertainties. In Chapter 2, it was discussed that the SOA nonlinearities have

the most impact on the degradation of the system in our case. Compared to a small variation of the gain and phase difference of the IQ modulator, a small change in I_{bias} can change the gain and therefore the nonlinearities brought about by the SOA. Raising the percent uncertainty of g_I and ϕ increases the variations of the input power P_{in} into the SOA which in turn can change the nonlinear characteristics of the SOA. Figure 3.17b and Figure 3.17c demonstrate the same analysis when the uncertainties of g_I and ϕ are raised to 10% then 15%. The case with I_{bias} ± 5%, g_I ± 15%, and ϕ ± 15% seems to show a scenario where all parameters have reached a similar level of impact on the transmission, thus the following analysis will encompass all cases up to 15% uncertainty for the IQ modulator's parameters.

The first set of simulations showcase some results for the 4-QAM at $P_{ref} = -14$ dBm configuration when looking at the four usual cases (no DPD, MP, EMP, and GMP predistorters). The predistorters are all identified taking into account all imperfections except the fiber's. The computed PDFs of the EVM is shown in Figure 3.18 where the explored uncertainties are at 5% for I_{bias} and 5%, 10% and 15% for g_I and ϕ.

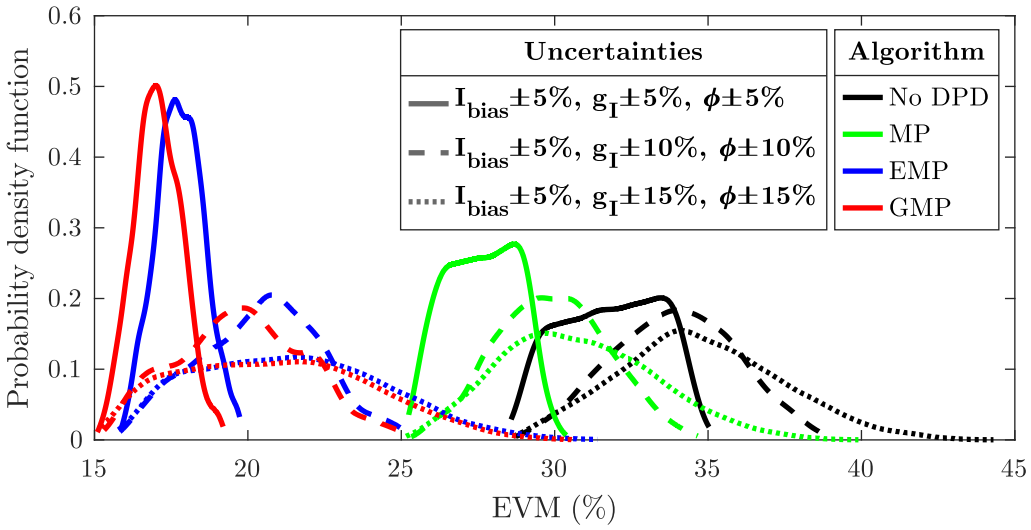


Figure 3.18 – 4-QAM probability density function at $P_{\text{ref}} = -14$ dBm

Like in the optical B2B case (Section 3.3.2), Figure 3.18 comes to support the claim that DPD is beneficial in two ways. This is apparent when looking at the range of EVM values becoming lower from no DPD to GMP. The more uncertainties added to the IQ modulator, the more the EMP and GMP PDFs start to merge (the GMP threshold still being lower than the EMP's). The IQ modulator's phase difference ϕ plays a role in the IQ

imbalance and introduces some phase uncertainty. The EMP is identified regardless of the phase thus might be blind to that change induced by ϕ. However, the GMP does take into account the phase of the signal and might be incorrectly compensating the nonlinearities related to the phase when ϕ starts becoming more influential. For the EMP and GMP, the values of EVM reflect a transmission that meets EVM requirements with its trailing edge for the case with 15% uncertainty having a low chance of occurrence. In the case of MP, it is only advisable to use it at 5% uncertainty. This configuration is also indicative of the fact DPD increases the immunity to the uncertainties in the system as the PDFs stretch upwards.

The constellations at the receiver constructed from 50 simulation runs (Figure 3.19) of the 15% case is displayed to show a visual representation of the results discussed previously. The GMP possess the better constellation with the brightest 2D-histogram.

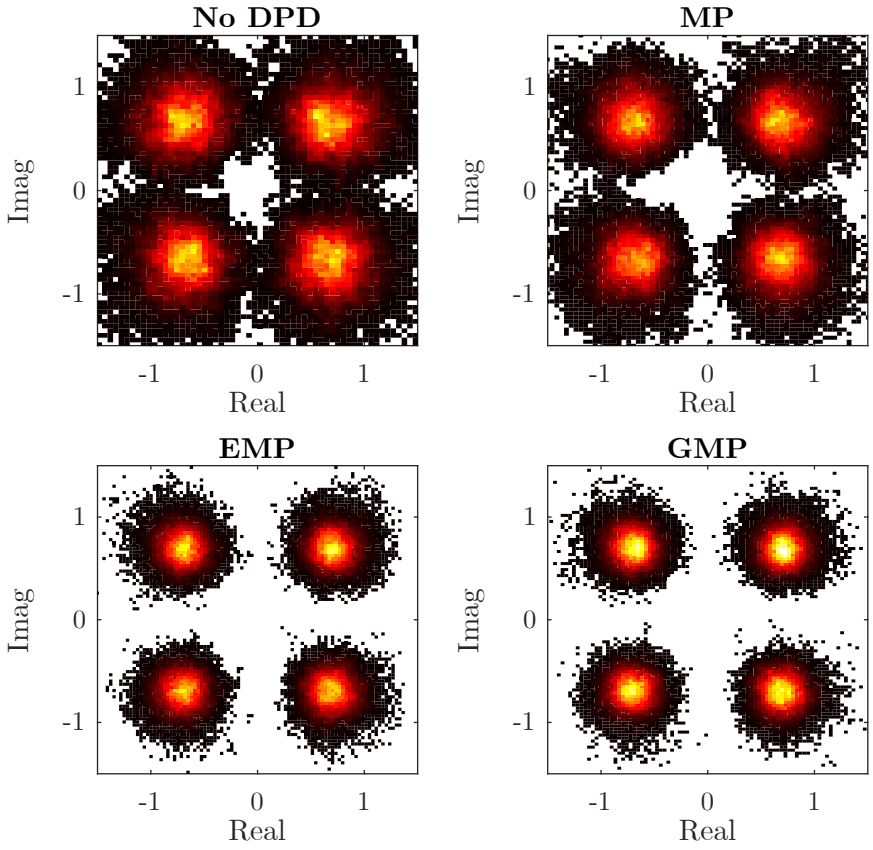


Figure 3.19 – 4-QAM constellations from 50 simulation runs at $P_{ref} = -14$ dBm with I_{bias} at 5% uncertainty and with g_I and ϕ at 15% uncertainty

Notice that here the EMP does not exhibit any phase shift as was the case in Section 3.3.

This could most likely be attributed to the equalization which would take care of not only the CD, but also any residual phase noise that the EMP was not able to compensate.

Taking into account the complexity of the predistorters (9 coefficients for MP, 33 for EMP, and 69 for GMP), the GMP could be a profitable predistorter to use as it has a more general structure and tends to give slightly better performances for a slightly higher complexity. When dealing with systems with quite nonlinear IQ modulators, EMP seems to also be a viable option with lower complexity.

Following is an assessment of the performances of the system for a 16-QAM configuration at $P_{ref} = -17$ dBm. Figure 3.20 demonstrates the estimated PDF from the PCE models of the same four cases (no DPD, MP, EMP, GMP) at 5% uncertainty. A

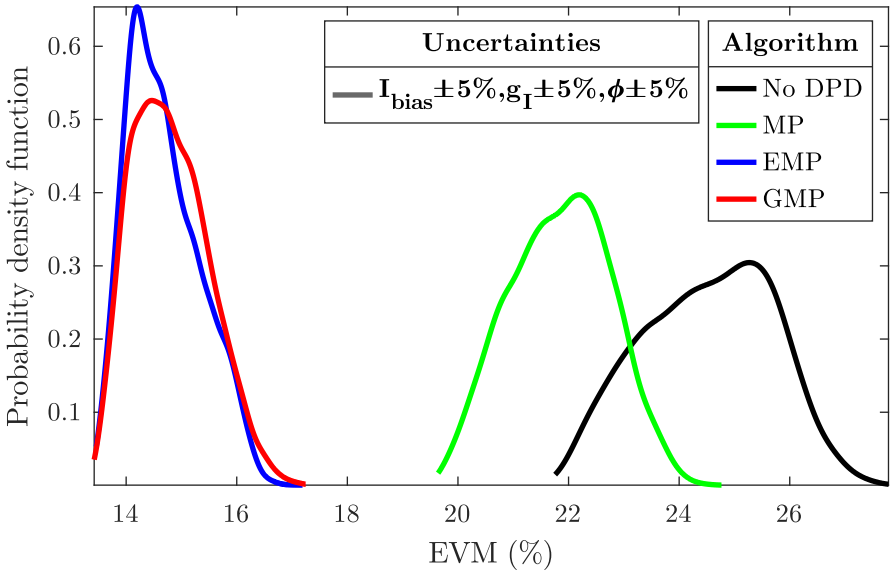


Figure 3.20 – 16-QAM probability density function at $P_{\text{ref}} = -17$ dBm

horizontal shift to the left signifies an improvement in EVM, hence a possibility for more reach. A vertical stretch upwards entails a higher robustness to uncertainties. The EMP and GMP comply with the EVM objective in most cases as less than 8.7% EVM values are above the threshold of 16%. In this regard, it can be said that PCE helps predict the failure of transmission as in 8.7% of cases, the transmission is deemed unsuccessful due to the values of EVM exceeding the designated limit. In the 16-QAM scenario, the EMP has a slightly higher immunity to component uncertainties than the GMP. A 16-QAM configuration having more high peak samples combined with the hard-clipping raises the the average power of the OFDM frame. The focus on the signal's envelope seems to work

rather well at high powers.

Figure 3.21 portrays 2D-histrograms of the 16-QAM symbols received. DPD shows brighter constellations with a tighter congregation especially for EMP and GMP.

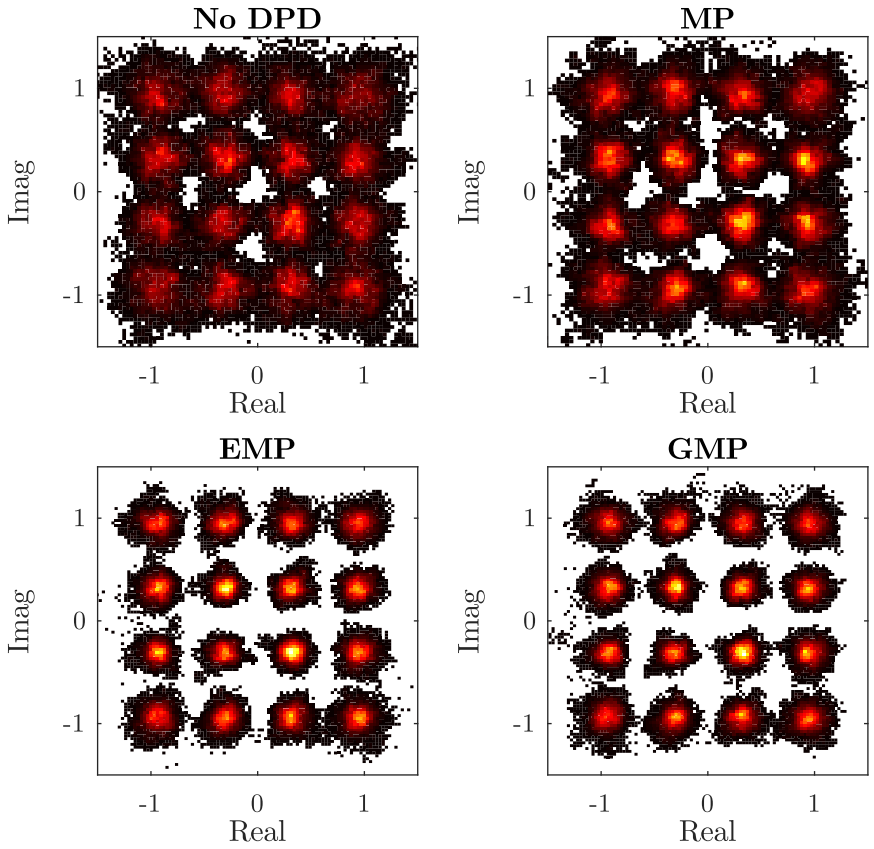


Figure 3.21 – 16-QAM constellations from 50 simulation runs for cases no DPD, MP, EMP, and GMP at $P_{ref} = -17$ dBm

The number of DPD coefficients for this scenario are the following: 9 for MP, 33 for EMP, and 55 for GMP. Being cognizant of the complexity of each predistorter, the better choice seems to lie on the EMP.

3.5 Conclusion

In this chapter, a stochastic analysis was performed for two cases: a case with a booster SOA in optical B2B and a case where the fiber and DAC imperfections are added. Joint parameter variations were applied on three parameters of the IQ modulator and the SOA.

For the booster case, the IQ modulator's peak-to-peak amplitude V_{pp} and phase difference between the branches ϕ were the chosen uncertain parameters as well as the SOA's bias current I_{bias}. For the case with fiber, the IQ modulator's V_{pp} was exchanged to introduce gain IQ imbalance by choosing the IQ modulator's branch gain g_I. The variability at the IQ modulator level not only allows for the sensitivity analysis of the impact of IQ imbalances, but also of the variation of the SOA's input power P_{in}. The variation of bias current I_{bias} has a direct impact on the SOA gain and can thus be equated to, for example, a change in the temperature or uncertainty on the length of the active region of the SOA which can have the same effect on the SOA gain. This means that choosing I_{bias} as an uncertain parameter can indirectly help understand the impact of environmental variations, component manufacturing variability, and sensitivity to operating parameter settings.

The PCE is shown to be a reliable source for stochastic analyses in the context of optical communication. The PDF of the EVM is computed for different DPD algorithms including the MP, EMP, and GMP. DPD is shown to decrease the EVM on average. More interestingly, it is discovered that DPD also serves to improve the robustness of the system to its uncertainties. The EMP and GMP come on top as DPD structures with the GMP being a more general structure showing slightly more promising results and manifesting more resilience to the phase shifts introduced by the SOA. The EMP still proves a reliable algorithm for its lower complexity and good performances. The results shown in this chapter have been published in IEEE *photonics technology letters* (PTL) [Sim+20] and have been presented at the CLEO 2021 and JNOG 2021 [Sim+21b][Sim+21a]. At the end of this chapter, it was discovered that the stochastic analysis could be used not only to see the probabilistic content of the EVM, but also to predict the failure of the system. This could lead into having a preview of predictions done for an experimental context. For instance, are the predistorters able to sustain the strain of uncertainties that are intrinsic to an experimental setup and are we able to accurately predict the range of EVM values that would be encountered in these conditions, thus predicting the rate of success of the global transmission?

Chapter 4

Digital predistortion for semiconductor optical amplifier based coherent optical orthogonal frequency division multiplexing systems: experimental results

In the two previous chapters digital predistortion (DPD) established itself as a valuable asset in reducing the undesirable effects of not only the semiconductor optical amplifier (SOA), but also those coming from the IQ modulator. Concurrently, DPD also proved to be efficient in diminishing the ramifications of signal variability and component uncertainties. These results were obtained after having made some simplifying assumptions about the transmission (e.g. the laser phase noise was not considered). The goal of this chapter is to reassert the benefits of DPD in an experimental setup where all of the components of the CO-OFDM transmission and their intrinsic uncertainties will be taken into account. First, the experimental testbed will be introduced. Then the basic setup of the different components of the transmission chain will be described. Finally, some experimental results with DPD will be presented. The memory polynomials (MP), envelope memory polynomials (EMP) and generalized memory polynomials (GMP) will be used with the addition of a novel linearization approach called the generalized envelope memory polynomials (GEMP).

4.1 Experimental setup

The experimental setup used in this chapter is depicted in Figure 4.1. As can be seen, the setup is similar to the one used for simulations. The "computer" portion is exactly the same implemented MATLAB code which now communicates with the components of the experimental setup. One of the objectives of this thesis was to be able to directly use our own algorithms for the baseband OFDM Tx and Rx and still be able to interact with the emitter and receiver components of the experimental setup, thus bypassing the use of the component maker's interface. Indeed, as the algorithms used by Keysight are a black box from our viewpoint, it makes it difficult to analyze what exactly is being done in the transmission and how Keysight's compensation schemes interact with each other. Doing our own signal processing on MATLAB also makes it possible for us to shape the signal in a more desirable way. For instance, different DPD or peak to average power ratio (PAPR) reduction schemes can be performed before sending the signal into the experimental testbed.

For the transmission shown in Figure 4.1, an OFDM frame including 1 synchronization symbol, 1 equalization symbol, and 30 data symbols is constructed. The 30 OFDM data symbols contain 8 pilots per symbol. For the OFDM data symbols, the input bit stream goes through a 4-QAM modulation. The 4-QAM symbols are coded into OFDM symbols with a bandwidth of 8 GHz consisting of 128 subcarriers. The bandwidth is chosen such that the oversampling factor of 4 is maintained knowing the waveform sampling frequency of the emitter is restricted to 32 GHz. The effective data rate for this configuration is 11.12 Gbit/s. Some hard-clipping at a 6 dB ratio is done on the time signal before being sent into the Keysight M8195A arbitrary waveform generator (AWG) through a virtual instrument systems architecture (VISA) connection. The 64 GSa/s digital to analog conversions (DACs) of the AWG have a theoretical resolution of 8 bits. Note that this resolution is not the effective number of bits (ENOB). The ENOB is an indicator of how the performance of DACs is degraded by noise and distortion and is dependent on the sampling rate and the signal's frequency. Considering [Key16], with an OFDM bandwidth of 8 GHz and the 32 GHz electrical bandwidth of the AWG, we assume the resolution would be closer to around 5 bits. The electro-optic conversion is done with the MXIQER-LN-30 optical modulator that has a V_π of 5.4 V and the Keysight N4391A optical modulation analyzer (OMA) local oscillator operating at 1540 nm with a laser linewidth of 100 kHz. The laser used at the emitter and at the receiver is the same hence the experimental setup

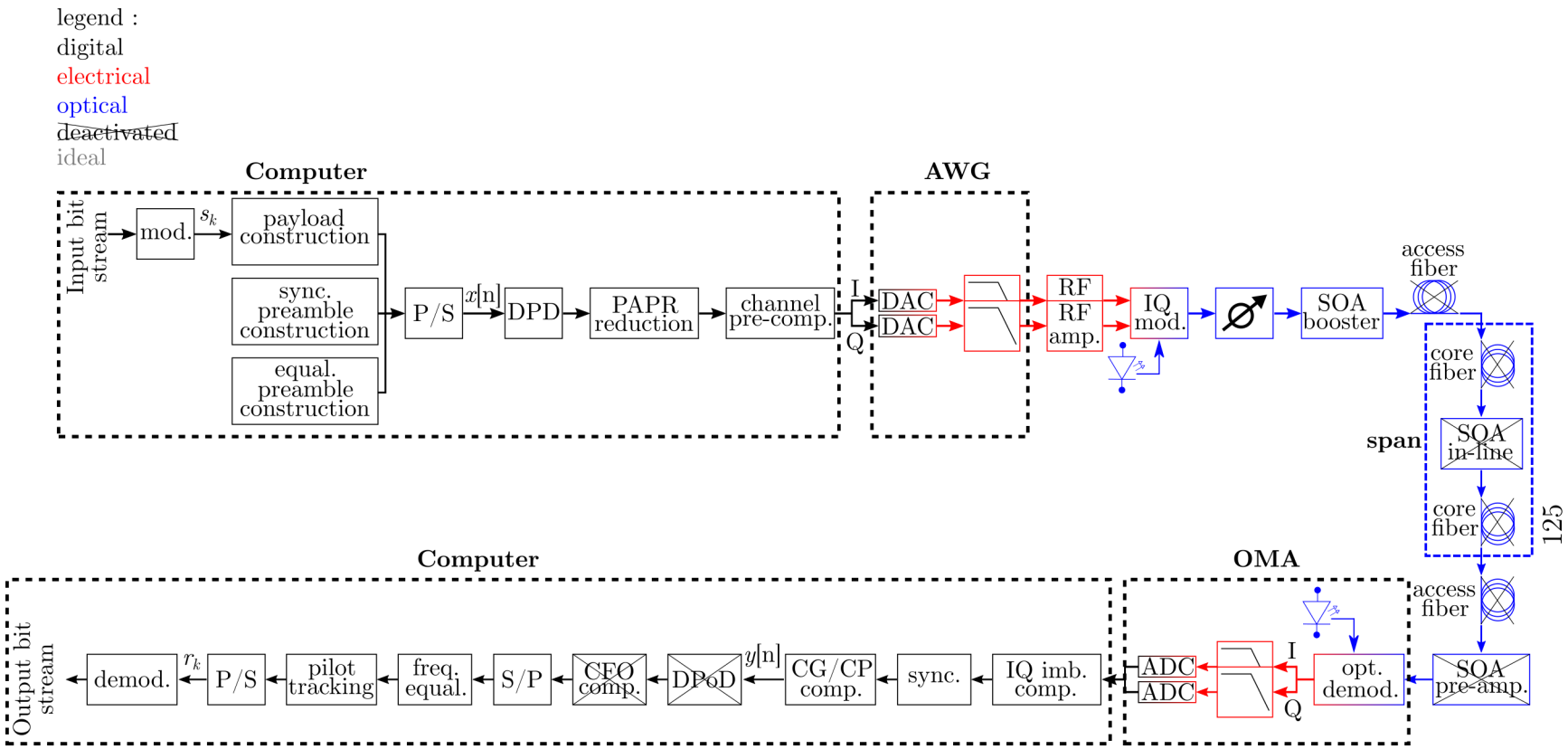


Figure 4.1 – Booster SOA experimental CO-OFDM transmission. Mod: modulation, equal.: equalization, sync.: synchronization, P/S: parallel to serial, DPD: digital predistortion, PAPR: peak to average power ratio, comp.: compensation, DAC: digital to analog conversion, opt.: optical, att.: attenuation, SOA: semiconductor optical amplifier, demod.: demodulation, ADC: analog to digital conversion, imb.: imbalance, CG/CP: common gain/common phase, DPoD: digital post-distortion, CFO: carrier frequency offset, S/P: serial to parallel, freq.: frequency. AWG: Arbitrary Waveform Generator, OMA: Optical Modulation Analyzer.

places us in slightly favorable conditions where the carrier frequency offset (CFO) is not an issue. The IQ modulator's operating point is stabilized via the MBC-IQ-LAB-A1 bias controller. The booster amplifier is a 750 μm INPHENIX-IPSAD1501 SOA with a gain of about 17 dB, a noise figure of 9 dB, and a supplied bias current that can go up to 300 mA. The OMA performs the opto-electronic conversion with the 40 GSa/s analog to digital conversions (ADCs) at a resolution of 8 bits (not in terms of ENOB), then sends the signal back into the computer by using the .NET interface. At the receiver, a low-pass filter with a 8.8 GHz bandwidth is applied to reduce out-band noise. Then, the signal goes through IQ imbalance compensation, synchronization, equalization, phase and amplitude tracking, and quadrature amplitude modulation (QAM) symbols demodulation. The IQ imbalance compensation, phase and amplitude tracking, and QAM symbols demodulation will be detailed in Section 4.3.1. The synchronization is an amplitude based cross-correlation between the received signal and the known synchronization preamble. The equalization is done by applying the inverse channel response between the received OFDM symbol and the known equalization preamble in the frequency domain. In this chapter, the fiber is not included as the focus is to reduce the nonlinearities caused by the SOA at the emitter side.

4.2 Setting up the experimental testbed

The experimental setup described in the previous section is a representation of the experimental testbed pictured in Figure 4.2. The following paragraphs discuss the most relevant components of that testbed.

The signal constructed on MATLAB is sent to the AWG in a dual channel configuration meaning the real part of the signal goes into channel 1 and the imaginary part into channel 4. The signal needs to be scaled to a peak amplitude of 1 (maximum between the real and imaginary part) in order to avoid clipping from the DAC. In the AWG, three specifications are of interest in this chapter: the sample time delay between I and Q (timing skew), finite impulse response (FIR) filter coefficients, and output amplitude. A sample time delay between the I and Q branches could be the result of having different electrical lengths between I and Q in the testbed. For example, if the I driving voltage has to travel through a longer path, there will be IQ imbalance which could hinder synchronization and leads to distortions in terms of constellation. Hence, there is a need to readjust the time delay through the AWG. One straightforward way to fix the delay is to look at the constellation or error vector magnitude (EVM) at the receiver and adjust it until there

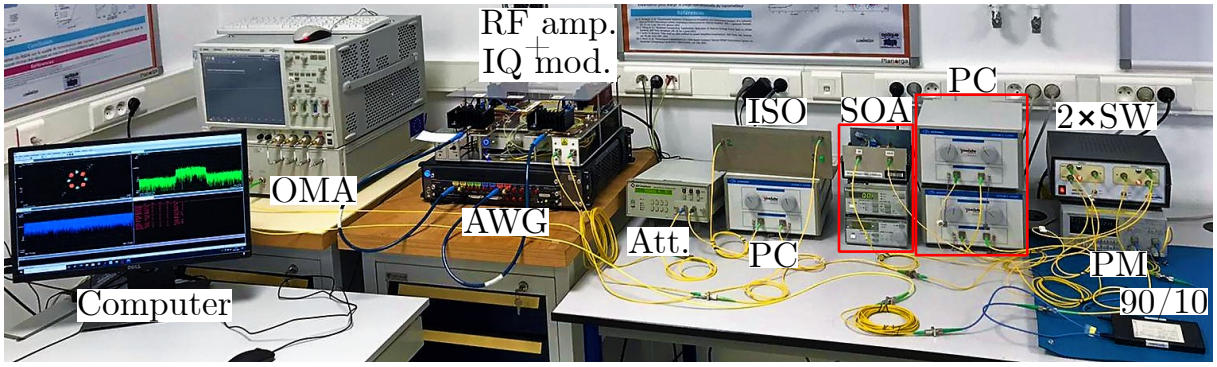
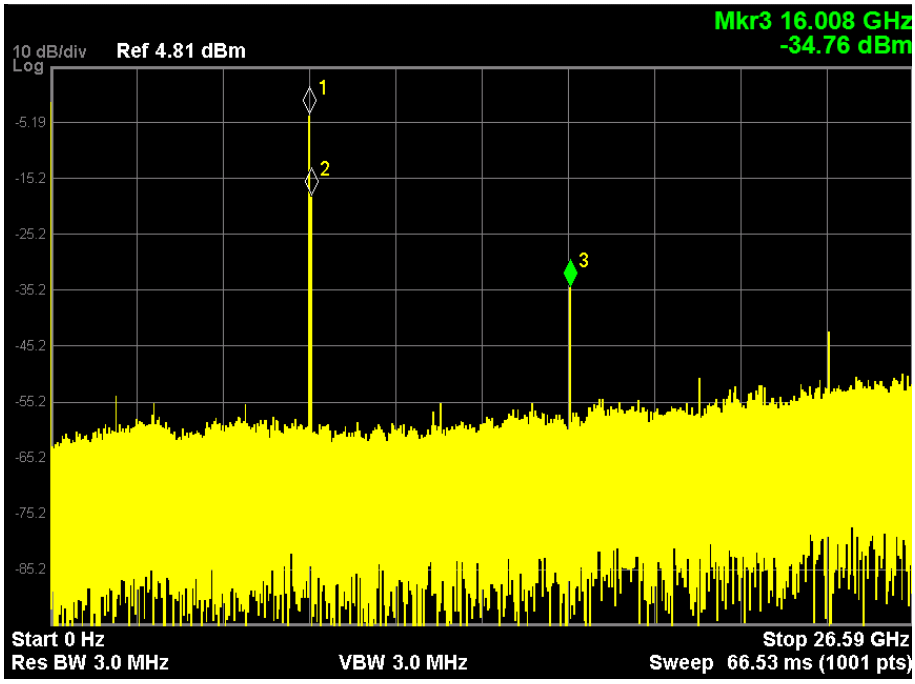


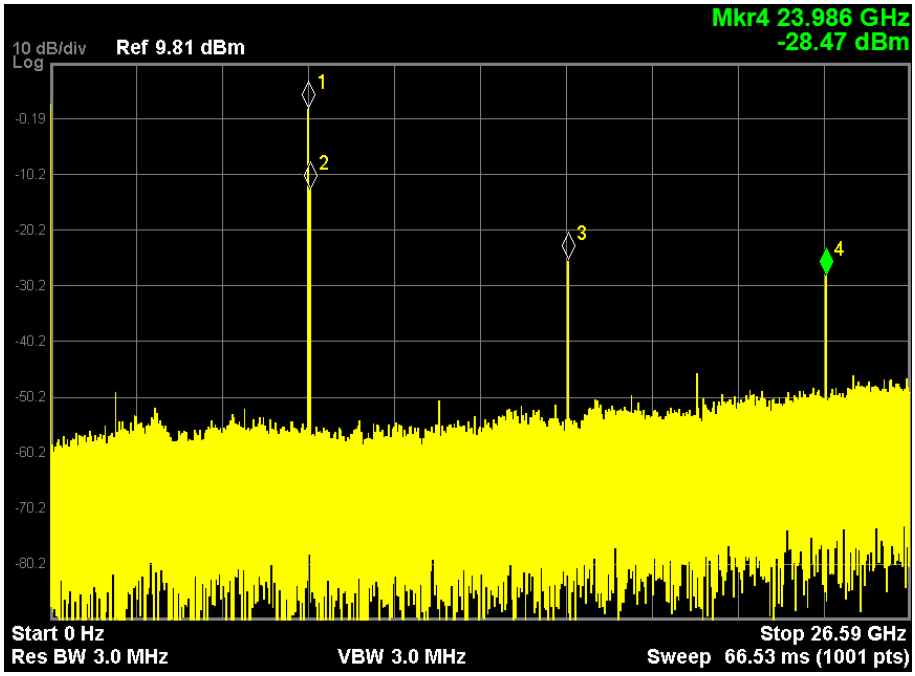
Figure 4.2 – CO-OFDM experimental testbed. 90/10: 90/10 splitter, Att.: Attenuator, AWG: Arbitrary Waveform Generator, IQ mod.: IQ modulator, ISO: Isolator, OMA: Optical Modulation Analyzer, PC: Polarization Controller, PM: Powermeter, SOA: Semiconductor Optical Amplifier, SW: Switch

is no visible improvement. In our case, the default configuration with no sample delay gives the best EVM thus none will be added throughout this experimental investigation. Before being converted to the electrical domain, a half-band Nyquist filter with a rollof factor of 0.2 and 32 taps (for the I and Q driving signals) is applied as an anti-alias filter. After having gone through the filter, the electrical signal exiting the AWG has a maximum peak-to-peak output that can be set to a maximum of 500 mV. It makes sense to think that the higher the output amplitude, the better since it entails a better signal to noise ratio (SNR) at the receiver. However, higher outputs could put the radio frequency (RF) amplifiers into their saturation region. A quick study is done to determine the state of linearity of the electrical signal at the emitter using the Agilent Technologies N9020A MXA signal analyzer. Figure 4.3 shows the spectrum of a 8 GHz sinusoid at the output of the AWG with peak-to-peak voltages of 250 mV and 500 mV. The first and second markers designate the fundamental frequency, the third marker the second harmonic, and the fourth marker the third harmonic. In both cases, there is at least approximately 30 dB between the fundamental and its harmonics so the AWG can be considered fairly linear at both 250 mV and 500 mV.

The electrical signal outputted by the AWG is amplified by an iXblue DR-AN-20-HO non-inverting single ended driver with a gain of 30 dB (according to the datasheet). Figure 4.4 shows the spectrum of a 8 GHz sinusoid at the output of the amplifier for different amplitudes coming out of the AWG (250 mV and 500 mV peak-to-peak). Note that attenuators are used to protect the signal analyzer so the values are to be viewed



(a) 250 mV peak-to-peak 8 GHz sinusoidal signal spectrum at the output of the AWG



(b) 500 mV peak-to-peak 8 GHz sinusoidal signal spectrum at the output of the AWG

Figure 4.3 – 8 GHz sinusoidal signal spectrum seen at the output of the AWG for different output amplitudes. Vertically: 10 dB/div, horizontally: 2.659 GHz/div

relatively. It can be seen from Figure 4.4 that at the 250 mV peak-to-peak AWG output, there is still about a 30 dB difference between the fundamental and the harmonics as was seen at the output of the AWG. However, at the 500 mV peak-to-peak amplitude, that difference goes down to about 25 dB meaning the amplifier is starting to saturate. This signifies that it would be preferable to limit the output amplitudes of the AWG closer to 250 mV peak-to-peak.

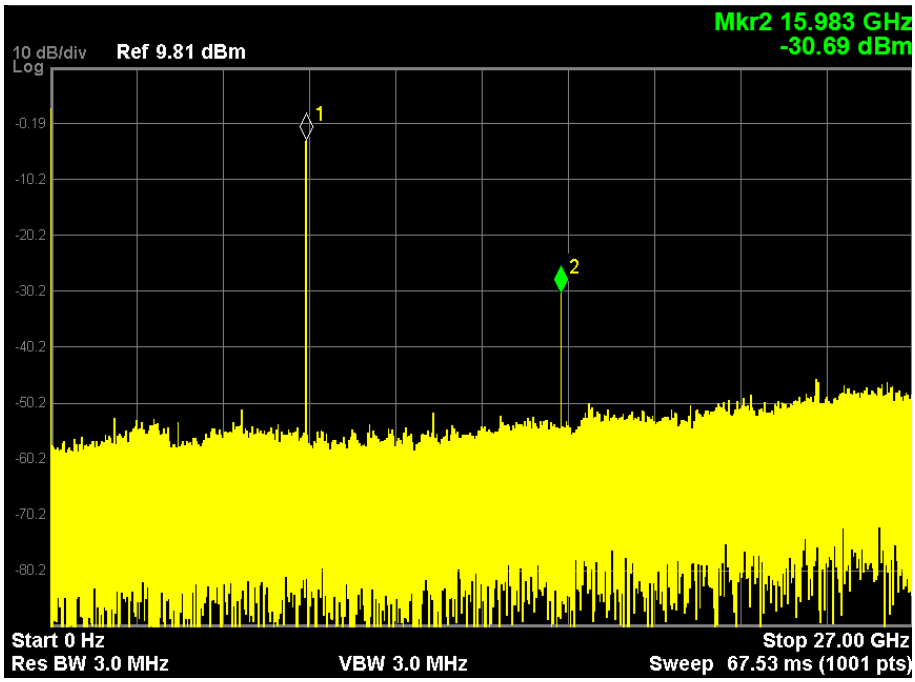
The nonlinearities of the RF amplifiers are not the only to watch out for as there could also be nonlinearities coming from the IQ modulator. Table 4.1 gives a glimpse into the role the IQ modulator could play in the transmission facing nonlinear effects:

Table 4.1 – Calculated MZM V_{pp} values according to the AWG output amplitude. The RF amplifiers have a gain of 30 dB and a set of 4 dB fixed attenuators are put at the RF amplifiers input for overvoltage protection

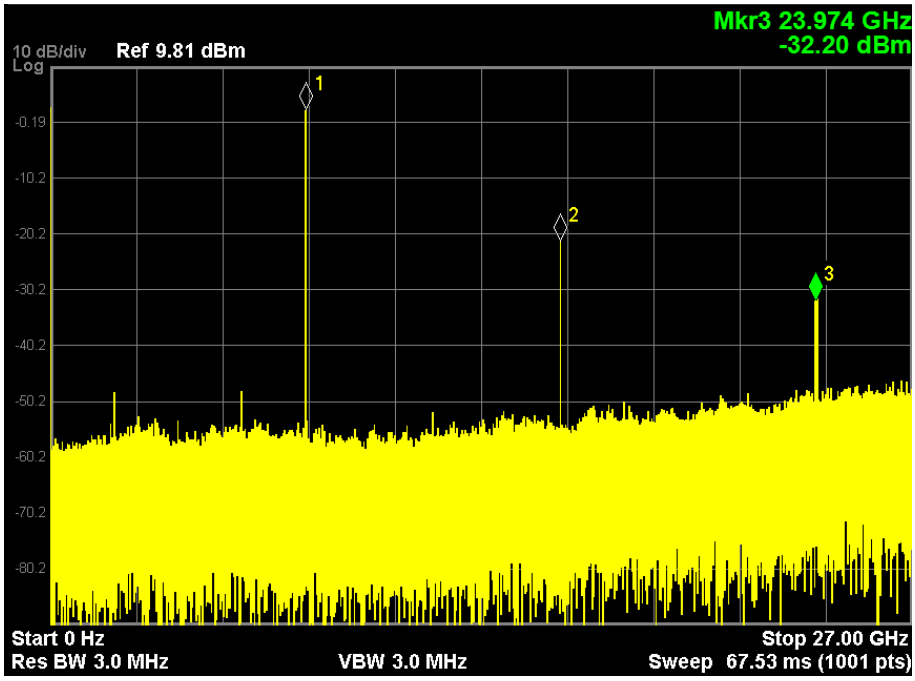
AWG output (mV)	MZM V_{pp} (V)
50	$0.18 \times V_\pi$
100	$0.37 \times V_\pi$
150	$0.55 \times V_\pi$
200	$0.74 \times V_\pi$
250	$0.92 \times V_\pi$
300	$1.11 \times V_\pi$
350	$1.29 \times V_\pi$
400	$1.48 \times V_\pi$
450	$1.66 \times V_\pi$
500	$1.85 \times V_\pi$

The MZM's transfer function (Figure 1.17) indicates that the IQ modulator has a nonlinear behavior. As V_{pp} increases, the signal pushes the IQ modulator into its nonlinear region. An IQ modulator has negligible nonlinearities so long as $V_{pp} \leq V_{pi}$ [Fre+13]. According to Table 4.1, the IQ modulator's nonlinearities thus begin to be non negligible starting at 300 mV at the output of the AWG. A compromise will need to be found between avoiding nonlinearities caused by the Mach-Zehnder modulator (MZM) and SNR.

Apart from the nonlinearities that can arise from the IQ modulator, some other undesirable effects can come about due to the IQ modulator when it is not set on the appropriate operating point. The proper operating point for our purposes was presented in Section 1.3.1. In an experimental setup, this operating point can be achieved by putting the two MZM branches at their null point corresponding to the point at the lowest power. This is done by adjusting a driving voltage in each MZM branch. Then, a third driving



(a) 250 mV peak-to-peak 8 GHz sinusoidal signal spectrum at the output of the RF amplifier



(b) 500 mV peak-to-peak 8 GHz sinusoidal signal spectrum at the output of the RF amplifier

Figure 4.4 – 8 GHz sinusoidal signal spectrum seen at the output of the RF amplifier for different output amplitudes. Vertically: 10 dB/div, horizontally: 2.7 GHz/div

voltage serves to adjust the phase difference between the two MZMs such that a phase difference of $\frac{\pi}{2}$ is obtained. This process is shown in Figure 4.5.

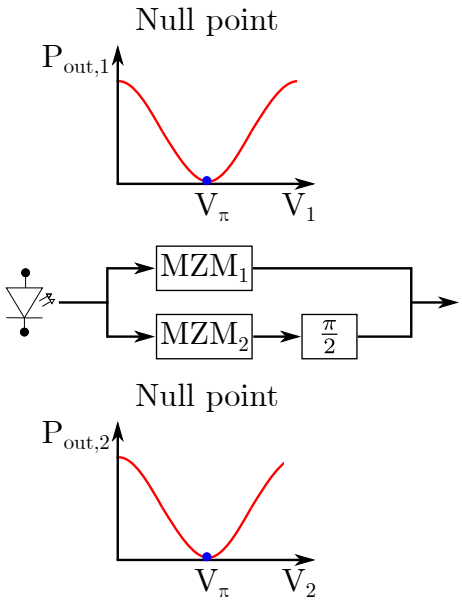
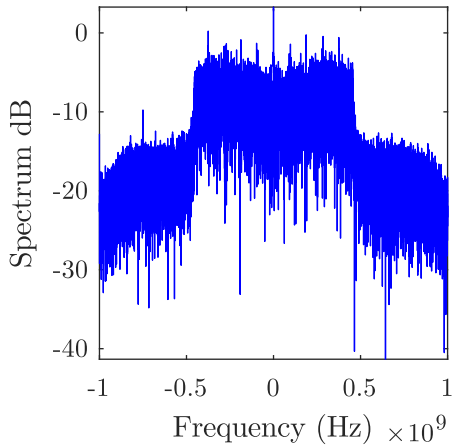


Figure 4.5 – Setting of the IQ modulator operating point. The two MZMs are operated at null point so that the IQ modulator can operate at the point where the optical field is most linear

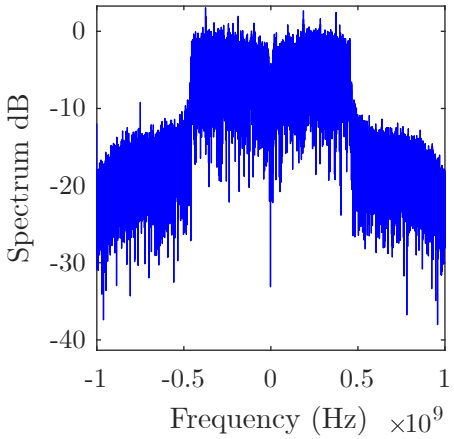
In this thesis work, a bias controller unit (iXblue MBC-IQ-LAB-A1 bias controller) has been used so as to automatically adjust these three driving voltages. Essentially, a small portion of the signal at the output of the MZM goes through a feedback loop after having been detected by a monitoring photodiode, then an algorithm uses a dither signal to readjust the DC bias voltage of the MZM when the IQ modulator has deviated from its operating point.

Figure 4.6 is a depiction of the spectrum obtained from a 1 GHz OFDM signal at the receiver when the IQ modulator is at its proper operating point (see the suppressed carrier at DC in Figure 4.6b) and when the IQ modulator is not at the right operating point (see the carrier at DC in Figure 4.6a). Their corresponding constellation below show that when the IQ modulator is not in its proper operating point, a significant imbalance occurs, resulting in a non Gaussian distribution of the received QAM symbols. Indeed, at the wrong operating point, the EVM is 32.9% compared to an EVM of 17.5% at the right operating point.

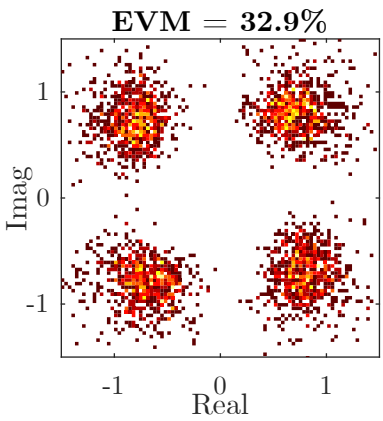
Figures 4.6b and 4.6d are examples showing that the algorithm in the bias controller



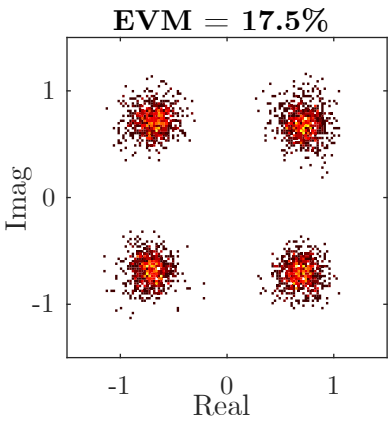
(a) IQ modulator at the wrong operating point



(b) IQ modulator at the right operating point



(c) IQ modulator at the wrong operating point



(d) IQ modulator at the right operating point

Figure 4.6 – Spectrum and constellation of a 1 GHz OFDM signal with an AWG output amplitude of 300 mV when the IQ modulator is at the wrong operating point ((a) and (c)) and at the proper operating point ((b) and (d))

unit is able to keep the correct operating point regardless of changes in the environment and polarization. Although the results presented in the next section relies on the iXblue device to perform an automatic bias control, some other work done in the Laboratoire des Sciences et Techniques de l'Information de la Communication et de la Connaissance (Lab-STICC) has shown that it may be possible to deactivate that device and still compensate the general problem of IQ imbalance in presence of laser phase noise by performing a joint parameter estimation [Fru+21].

As stated in the previous paragraph, the bias controller is used which allows us to focus on the undesirable effects of the SOA. The SOA unit is composed of the SOA, the Thorlabs TED 200 thermoelectric temperature controller using a control loop with a thermistor as a temperature sensor that is set to a resistance of 12.5 kΩ, and the Thorlabs LDC 210 C laser diode controller to control the bias current. A manual polarization controller is also put in front of the SOA as a change in polarization can lead to a change in gain of typically 1.5 dB, as well as an optical attenuator to control the input power into the SOA.

The other polarization controllers are placed in front of the IQ modulator and the OMA to maximize the power at the receiver. The polarization controller placed before the IQ modulator stems from the previous set up where the IQ modulator's operating point had to be set up manually; the new IQ modulator unit monitoring the polarization reduces the need for that controller. The polarization controller placed in front of the OMA is put in anticipation for a dual polarization configuration; with only one polarization, the OMA is able to combine the two modes at the receiver without necessarily needing to adjust the polarization.

The OMA receives 90% of the signal while either the Anritsu MT9810A powermeter or the APEX AP2683A optical complex spectrum analyzer (OCSA) takes in 10% of the signal using splitters. Using the OCSA, it is possible to see that the emitter's frequency response is not all-pass. Figure 4.7 suggests that the frequency response of the AWG and IQ modulator produces an uneven spectrum for a 8 GHz OFDM signal sent into an optical back to back (B2B) setup.

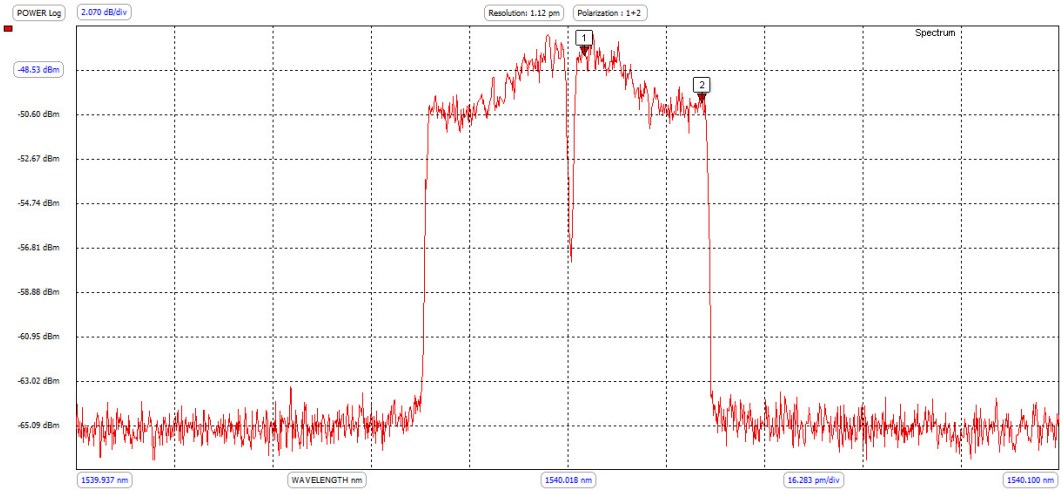


Figure 4.7 – 8 GHz OFDM signal spectrum in optical B2B without passband pre-compensation

This behavior is corrected by estimating the passband frequency response in the MATLAB OFDM baseband receiver using the equalization preamble, then constantly using the inverse of that response as the channel pre-compensation block. Figure 4.8 is a visual representation of the AWG being compensated. This correction is done to ensure the main changes seen in the signal's spectrum can be attributed to the SOA.

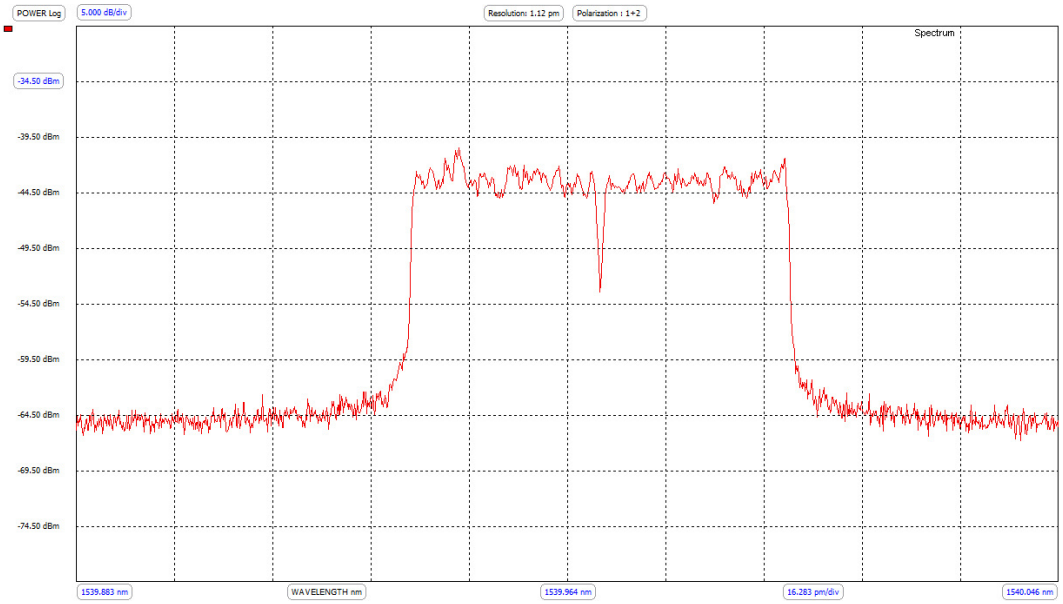


Figure 4.8 – 8 GHz OFDM signal spectrum in optical B2B with passband pre-compensation

Before presenting the experimental results, it is useful to know that the signal at the receiver is able to be retrieved with MATLAB by interacting with the Keysight 89600 vector signal analysis (VSA) software of the OMA where a trace of the signal can be saved onto the computer then reloaded on MATLAB for processing. When saving the trace, the OMA resamples the signal to a frequency equal to the frequency span multiplied by 1.28 for a complex signal (2.56 for a real signal). The frequency span is set to 25 GHz to keep the 32 GHz sampling frequency that is set by the emitter. Note that before saving a trace, it is necessary to tell the VSA software to automatically adjust to the input power range of the hardware to have meaningful results. Now that the experimental testbed has been set up accordingly, the experimental investigation can be pursued.

4.3 Experimental investigation of a CO-OFDM transmission with a booster SOA

The experimental setup now taking into account certain aspects that were bypassed when presenting the numerical results signifies a few baseband Rx blocks will appear: the IQ imbalance compensation, the amplitude and phase tracking, and a data-aided demodulation. The IQ imbalance compensation will be explained first.

4.3.1 Experimental baseband receiver

The IQ imbalance is defined as a quasi-static phase and amplitude mismatch between the real and imaginary part of the signal. These impairments can stem from the emitter or the receiver in the form of an improper biasing of the IQ modulator, imperfections in the transmission components, and timing mismatch due to the difference in cable lengths between I and Q. As the IQ modulator has a bias controller and the IQ delay skew (also called IQ time skew) is supposed to be corrected through the AWG, the IQ compensation here would be taking care of any residual imbalance. A well-known technique used to perform this compensation is the Gram-Schmidt orthogonalization procedure (GSOP) [Fat+08]. The GSOP aims to orthogonolize two nonorthogonal components of the received signal $r_I[n]$ and $r_Q[n]$, to then obtain a pair of orthonormal components $I^O[n]$ and $Q^O[n]$ like so:

$$I^O[n] = \frac{r_I[n]}{\sqrt{E\left[r_I^2[n]\right]}} \tag{4.1}$$

$$Q^O[n] = \frac{Q'[n]}{\sqrt{E\left[Q'^2[n]\right]}} \tag{4.2}$$

where

$$Q'[n] = r_Q[n] - \frac{E\left[r_I[n]\ r_Q[n]\right]\ r_I[n]}{E\left[r_I^2[n]\right]} \tag{4.3}$$

Figure 4.9 shows the compensation in action where the EVM can be reduced by roughly 2% and the constellation shows that the concentration of QAM symbols near the reference is higher. This slight improvement also shows that the IQ imbalance will not be the main issue in this analysis.

The next compensation that is now activated is the amplitude and phase tracking which uses the OFDM symbol's pilots to track the amplitude and the phase over time.

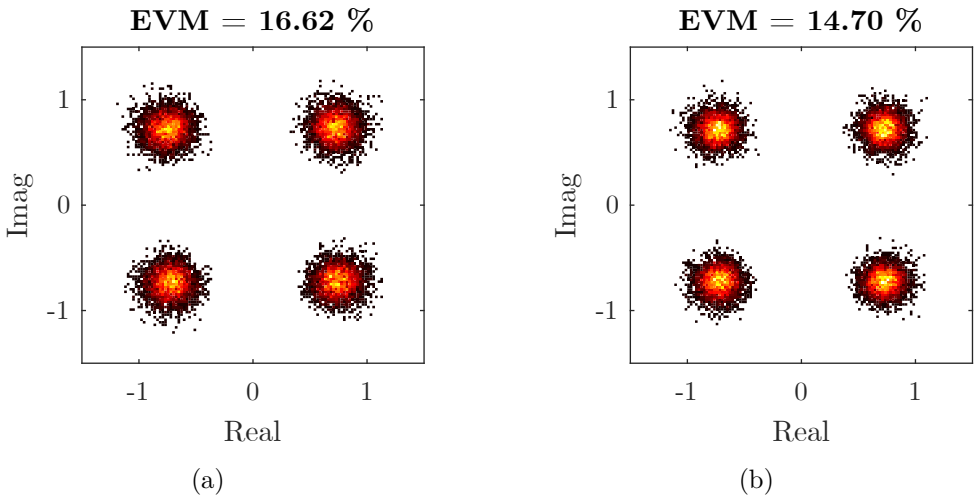


Figure 4.9 – Constellation of 150 OFDM (5 different measurements of a frame with 30 OFDM data symbols) with a bandwidth of 8 GHz OFDM signals (a) without IQ imbalance compensation (b) with IQ imbalance compensation

Contrary to the numerical results presented in Chapters 2 and 3, the experimental results will include the effects of laser phase noise. As stated in Section 1.3.1 and Equation 1.7, the laser phase noise process is a time varying process which is dependent on the laser linewidth $\Delta\nu$. Usually the phase noise of the emitter laser and the receiver laser add up together. However, considering our experimental setup uses one local oscillator, the OMA's 100 kHz of laser linewidth puts us in a favorable position. Note that the laser at the emitter beats with a delayed version of itself at the receiver meaning the impact of laser phase noise would increase with larger fiber lengths. A short investigation of the amplitude and phase tracking scheme will show that the laser phase noise in our situation is still nevertheless impactful.

Figure 4.10 represents the constellations from three different measurements of an 8 GHz OFDM frame. Below each set of constellation is their corresponding estimated phase at the OFDM symbol rate. The way the phase was estimated will be discussed later on. The change in phase per OFDM symbol seems to be coherent with the transient nature of the phase noise process. The phase observations also help understand the constellation shape. For example, in Figure 4.10a there is a clump of data shifted to the left of the reference QAM symbols by about 110° which corresponds to the phase shift shown in Figure 4.10d. Similarly so, juxtaposing Figures 4.10c and 4.10f it makes sense that the constellation has more data points shifted to the far right as 12 of the 30 OFDM symbols

are between 40-50° while the other symbols are spread out in terms of phase.

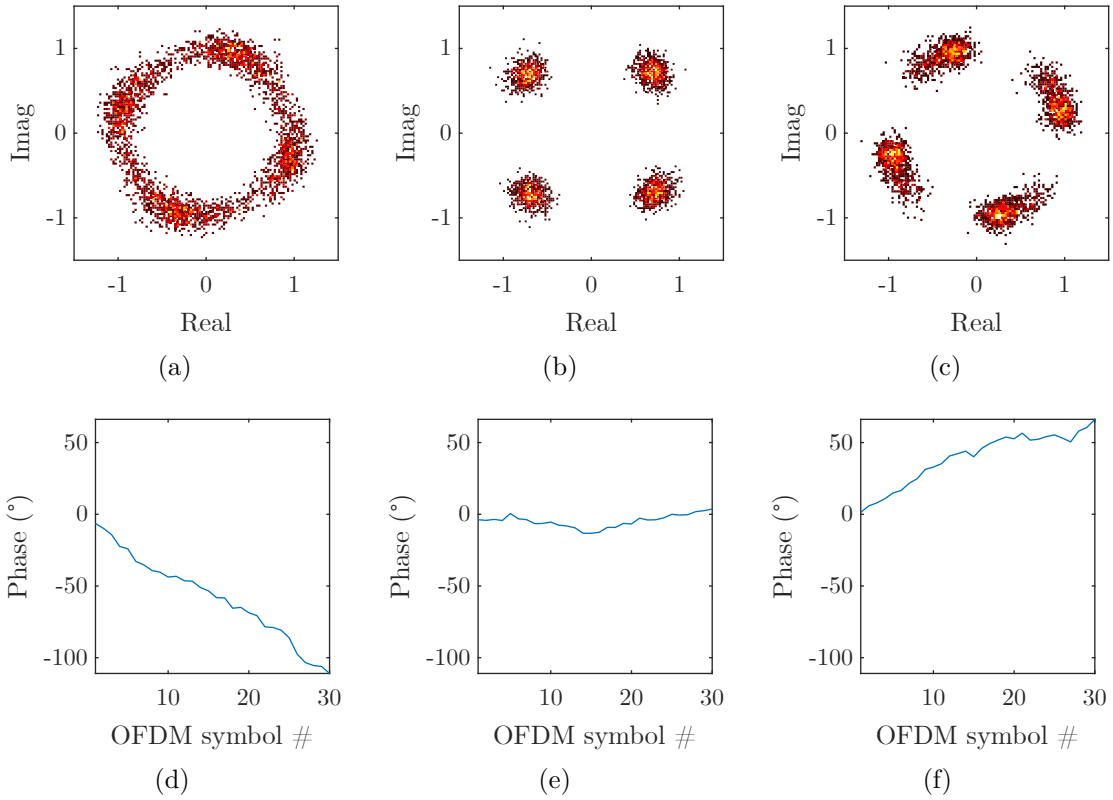


Figure 4.10 – Constellation and estimated phase per OFDM symbol from 3 different measurements of an 8 GHz OFDM frame subjected to laser phase noise. The estimated phase curves correspond to the same frame as the constellations above them

The amplitude and phase tracking is done in the frequency domain using a simple division between the sent OFDM symbol pilot subcarriers (8 pilots per orthogonal frequency division multiplexing (OFDM) symbol for our configuration) and the received OFDM symbol pilots to get a frequency response [Yi+07]. The complex coefficient that is used to do the compensation corresponds to the average of the inverse frequency response. Since the phase noise process is time varying, a new complex coefficient is calculated then applied per OFDM symbol. Note that the amplitude of this coefficient fluctuates around 1 in our setup which means this coefficient mainly corrects the phase of the signal.

The effectiveness of that compensation scheme is shown through Figure 4.11 which represents the amplitude modulation to phase modulation (AM/PM) from 5 different measurements of an 8 GHz OFDM signal with and without amplitude and phase tracking. The input in this case is the OFDM frame before hard-clipping is applied while the output

is the received OFDM frame after synchronization. In Figure 4.11a, phase spread shows that the laser phase noise has a consequential effect on the transmission. The goal is to have a 0 phase difference between the input and output regardless of the input amplitude. Figure 4.11b shows that the amplitude and phase tracking scheme is capable of greatly compensating the laser phase noise.

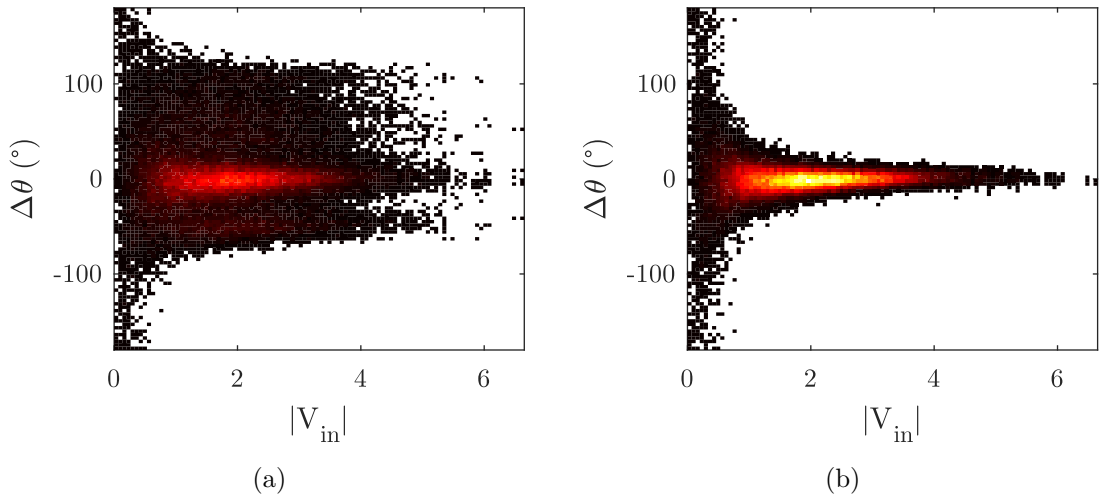


Figure 4.11 – Amplitude modulation to phase modulation curves of five 8 GHz OFDM frames (a) without amplitude and phase tracking (b) with amplitude and phase tracking

The last change in the MATLAB baseband Rx algorithms is in the demodulation which is now a data aided QAM demodulation instead of a simple one [Ozd+07]. Data aided QAM demodulation is done as depicted in Figure 4.12.

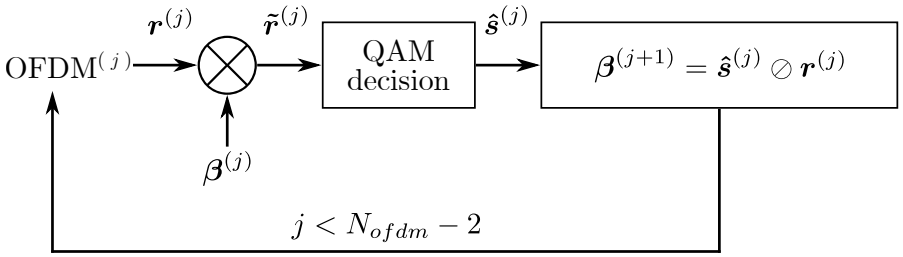


Figure 4.12 – Data aided QAM demodulation. $\boldsymbol{r}^{(j)}$ is a vector containing the received symbols of the j^{th} received OFDM symbol, $\tilde{\boldsymbol{r}}^{(j)}$ is a vector containing the received symbols corrected with $\boldsymbol{\beta}^{(j)}$, $\hat{\boldsymbol{s}}^{(j)}$ is the projected symbols obtained from the QAM symbol decision, $\oslash$ is an element-wise division, and N_{ofdm} is the total number of OFDM symbols in the frame

Data aided demodulation reads in the following way:

1. Initialize a complex vector $\boldsymbol{\beta}^{(j)}$ designating the channel frequency response to a set of $1 + 0i$ with length equal to the number of data QAM symbols per OFDM symbol
2. Multiply j^{th} OFDM symbol by $\boldsymbol{\beta}^{(j)}$
3. Obtain the projected symbols $\hat{\boldsymbol{s}}^{(j)}$ after doing QAM symbol decision
4. Update $\boldsymbol{\beta}^{(j)}$ at the rate of the symbol duration by taking the ratio between the projected symbols $\hat{\boldsymbol{s}}^{(j)}$ and the received QAM symbols $\boldsymbol{r}^{(j)}$
5. Increment j and repeat steps 2-6 until the end of the OFDM frame ($N_{ofdm} - 2$ with the -2 removing the synchronization and equalization preamble from the OFDM frame)

This allows better results as evidenced by Figure 4.13 which shows the constellation from five different measurements of an 8 GHz OFDM frame with and without data aided demodulation.

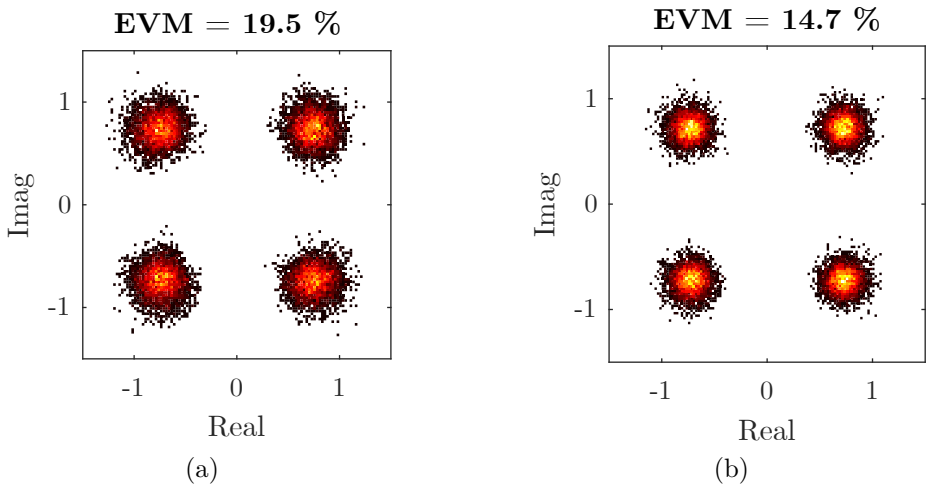


Figure 4.13 – Constellation of five 8 GHz OFDM signals (a) without data aided demodulation (b) with data aided demodulation

It is important to note that as mentioned before, the zero-forcing method albeit simple can accentuate the noise which can be seen in the constellation diagrams where some QAM symbols (in black) are farther from the reference symbol. A method other than zero-forcing could help improve the quality of transmission.

4.3.2 Experimental comparison of DPD algorithms

Using the experimental setup and algorithms presented in Sections 4.1 and 4.3.1, the behavior of the SOA is plotted with respect to its reference input power and its corresponding EVM (Figure 4.14). The input power into the SOA is measured by recording the power going into the OMA in an optical B2B setting. The optical B2B setting here means that the transmission is done without the SOA. As mentioned previously, the results are now a reflection of the algorithms we implemented on MATLAB instead of those from Keysight's VSA software. It therefore makes sense to first do a quick comparison of the two to verify the performance of the implemented MATLAB receiver. Figure 4.14 shows the EVM at different input powers into the SOA of five 8 GHz OFDM frames using the VSA receiver vs. the MATLAB receiver (without fiber). Figure 4.14 depicts that

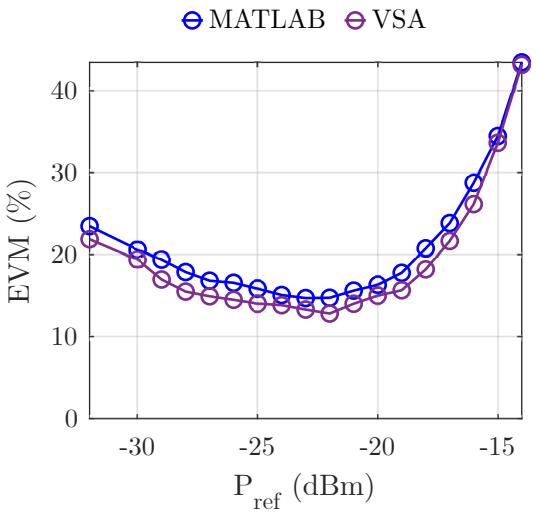


Figure 4.14 – Average EVM of the VSA and MATLAB receivers at different input powers into the SOA for five measurements of an 8 GHz OFDM frame. $I_{bias} = 200$ mA

the SOA behaves as expected with a slightly higher EVM at low input powers due to the amplified spontaneous emission (ASE) noise then a reduction in that EVM as power increases until the nonlinear region of the SOA is reached, forcing the EVM to rise. The values from VSA and MATLAB seem to be quite close to each other meaning the rest of the investigation can be pursued safely.

The second step of this investigation looks into the power spectral density (PSD) of the OFDM frames which gives an insight into the effects of the SNR and adjacent channel power ratio (ACPR), an aspect that was not touched upon as deeply in the numerical

results part of this thesis. The normalized PSDs of an 8 GHz OFDM frame are shown at different input powers into the SOA in Figure 4.15.

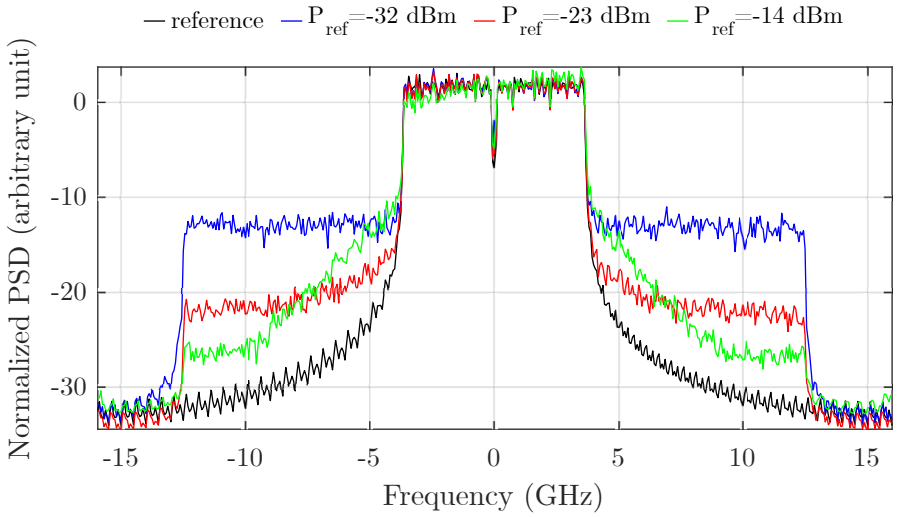


Figure 4.15 – PSD of an 8 GHz OFDM frame at different input powers into the SOA. I_{bias} = 200 mA, 300 mV peak-to-peak AWG output

It can be seen that there is a drop at a frequency of 12.5 GHz which is due to the OMA being set to a frequency span of 25 GHz (to have a sampling rate of 32 GHz). At $P_{ref} = -32$ dBm, the ASE noise is more present which explains the higher noise levels; there is a little more than a 10 dB difference with the passband. The $P_{ref} = -23$ dBm point is situated on what is considered the linear region of the SOA where there is now a little more than 20 dB difference with the passband. The region that will be of interest for this thesis includes $P_{ref} = -14$ dBm where the SNR is higher but some spectral regrowth can be seen on the adjacent channels as well as the effect of phase-amplitude coupling whose shape can be seen in the passband. This is in accordance to the numerical results shown in Figure 2.14 where the nonlinearities of the SOA manifested through some asymmetrical spectral regrowth in the out-band region, and the tilt in the passband due to the phase-amplitude coupling.

DPD will be used to mitigate these effects. With low complexity being a crucial aspect for photonics applications, this chapter chooses to deal with the same three simplified inspired versions of the Volterra Series: the MP, the EMP, and the GMP. A reminder of their mathematical expressions can be found in Table 4.2 where K_i are nonlinearity orders, L_i and M_i are the memory depths, and a_{kl}, b_{klm}, c_0 and c_{klm} are the coefficients

Table 4.2 – Known predistorter algorithms

$y_{MP}(n) =$	$\sum_{k=0}^{K-1}\sum_{l=0}^{L-1} a_{kl}x(n-l)\|x(n-l)\|^k$
$y_{EMP}(n) =$	$x(n)\left(c_0 + \sum_{k=1}^{K-1}\sum_{l=0}^{L-1} a_{kl}\|x(n-l)\|^k\right)$
$y_{GMP}(n) =$	$\sum_{k=0}^{K_a-1}\sum_{l=0}^{L_a-1} a_{kl}x(n-l)\|x(n-l)\|^k$ $+\sum_{k=1}^{K_b}\sum_{l=0}^{L_b-1}\sum_{m=1}^{M_b} b_{klm}x(n-l)\|x(n-l-m)\|^k$ $+\sum_{k=1}^{K_c}\sum_{l=0}^{L_c-1}\sum_{m=1}^{M_c} c_{klm}x(n-l)\|x(n-l+m)\|^k$

to be estimated.

Previous investigations suggested that predistorters based on the envelope of the signal tend to display very good performances [Dio+17]. Consequently, we propose a new memory polynomials DPD structure, the GEMP, with the aim to provide an extra degree of freedom as is the case for the GMP but focusing solely on the signal's magnitude. The GEMP's mathematical expression is shown below:

$$y_{GEMP}(n) = x(n)\left(c_0 + \sum_{k=1}^{K_a-1}\sum_{l=0}^{L_a-1} a_{kl}|x(n-l)|^k + \sum_{p=1}^{K_b}\sum_{q=1}^{K_c}\sum_{m=1}^{M_b}\sum_{l=0}^{L_b} b_{pqml}|x(n-l)|^p \cdot |x(n-l-m)|^q\right) \tag{4.4}$$

The added value of the GEMP is that its use can be extended to other scenarios such as intensity-modulation/direct-detection (IM/DD). To find the appropriate structural parameters while taking into account the impairments of the experimental setup, more emphasis was put on the normalized mean square error (NMSE) in the Hill-Climbing error criterion in Equation 2.3.

The results shown next were obtained by identifying the predistorters (MP, EMP, GMP, and GEMP) at different input powers of the SOA then validating the DPD by sending 32 OFDM symbols at that same input power. Figure 4.16 displays the mean EVM averaged from 5 different measurements at $I_{bias} = 200$ mA for input powers into the SOA ranging from $P_{ref} = -19$ dBm to $P_{ref} = -14$ dBm. Contrary to simulation analyses, here component limitations put restrictions on the maximum power that can

be sent into the SOA. For instance, the AWG output amplitudes are not able to go over 500 mV and the inputs of the RF amplifiers have to be controlled to not damage the component.

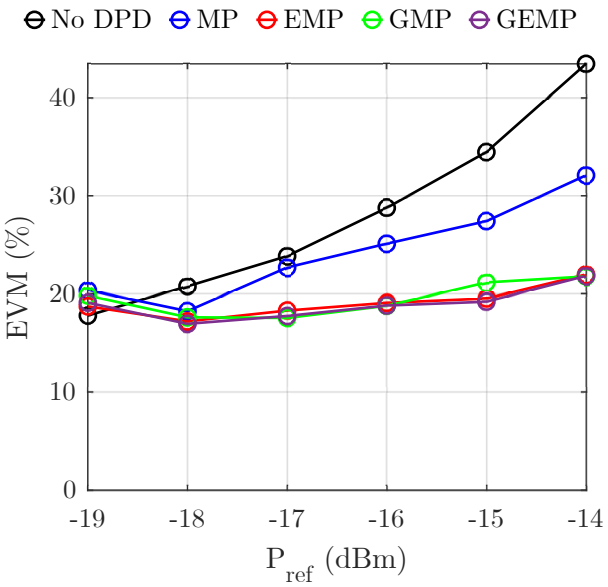


Figure 4.16 – Average EVM of multiple predistorters at different input powers into the SOA for five measurements of an 8 GHz OFDM frame. $I_{bias} = 200$ mA, 300 mV peak-to-peak AWG output

In Figure 4.16, it can be seen that starting at $P_{ref} = -19$ dBm, the system is too linear for DPD to be activated and stay beneficial. For higher input powers, looking at the 30% EVM limit, the MP allows for a 3.6 km reach increase going from $P_{ref} = -15.75$ dBm to $P_{ref} = -14.5$ dBm and accounting for the SOA gain behavior and the attenuation in a single-mode fiber (SMF). The EMP, GMP, and GEMP behave almost the same and are well below the required limit. For $P_{ref} > -18.5$ dBm, we see a relative improvement in the EVM that becomes all the more important when the power into the SOA increases. The EVM is seen to be reduced to about 20% for all input powers currently shown. Contrary to when the numerical results were presented, it is harder for the GMP to detach itself from the EMP in an experimental environment.

The constellation diagrams of Figure 4.17 comparing five measurements from all predistorters at $P_{ref} = -14$ dBm show a similar pattern where the MP has slight improvements compared to the case with no DPD while the EMP, GMP, and GEMP have the brightest constellations.

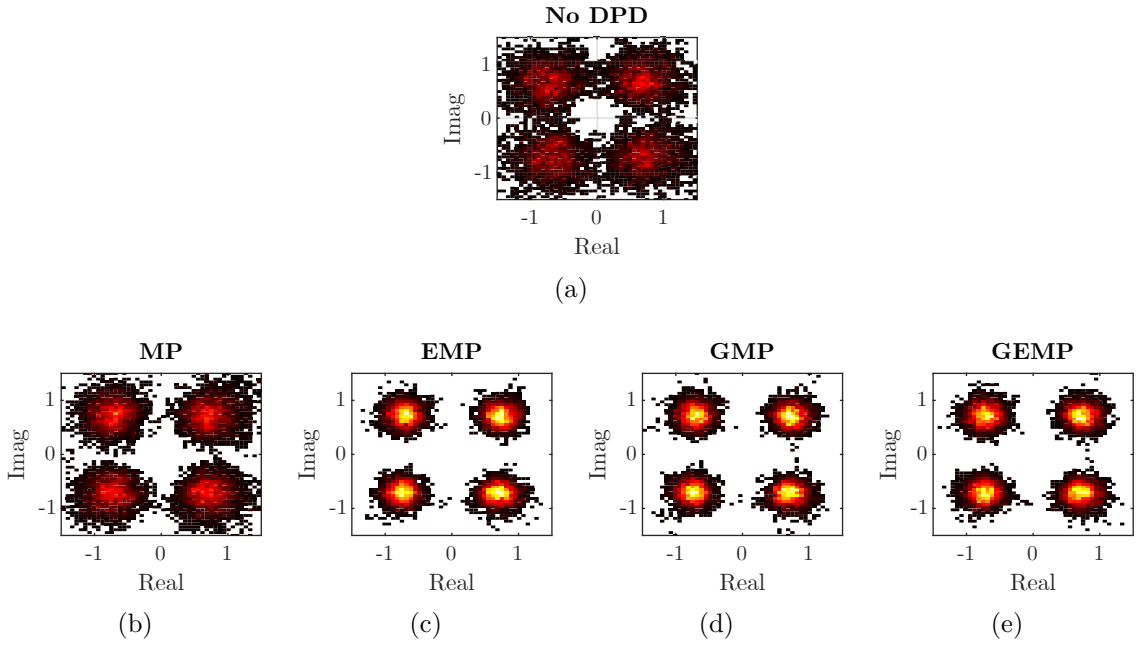


Figure 4.17 – Constellation diagrams of multiple predistorters at different input powers into the SOA for five measurements of an 8 GHz OFDM frame. $I_{bias} = 200$ mA, 300 mV peak-to-peak AWG output

One way to differentiate the three best predistorters is to look at their complexity. For the $P_{ref} = -14$ dBm case, the MP had 9 coefficients, the EMP 27, and the GMP and GEMP had 20 coefficients. With phase noise being the most prominent factor affecting the DPD learning phase, this could be an avenue to explore in order to explain that the EMP now requires more coefficients than the GMP to have similar performances.

The performance of the four predistorters is re-evaluated by increasing the bias current supplying the SOA to 240 mA and increasing the AWG output peak-to-peak amplitude from 300 mV to 425 mV in order to push the maximum SOA input power to -13 dBm instead of -14 dBm. This scenario should lead to more nonlinearities coming from not only the SOA, but also the RF drivers and the IQ modulator.

Before we delve into the case where $I_{bias} = 240$ mA, we will do an analysis of the effects increasing I_{bias} has on the transmission. A comparison is done between the case with $I_{bias} = 200$ mA and $I_{bias} = 240$ mA looking at the EVM over different input powers into the SOA in Figure 4.18. That figure displays that at first, the case where the bias current is higher yields higher values of EVM. This could be possibly attributed to two things: the ASE noise or the added nonlinearities. As stated earlier, a higher bias current

leads to more nonlinearities. A higher bias current for the same input power also means more ASE. Starting at $P_{ref} = -16$ dBm, there is a crossover between the two scenarios which could be connected to the higher SNR at the receiver for $I_{bias} = 240$ mA. Indeed, at $P_{ref} = -14$ dBm, the input power into the OMA is 2.50 dBm for $I_{bias} = 200$ mA compared to 3.86 dBm for $I_{bias} = 240$ mA.

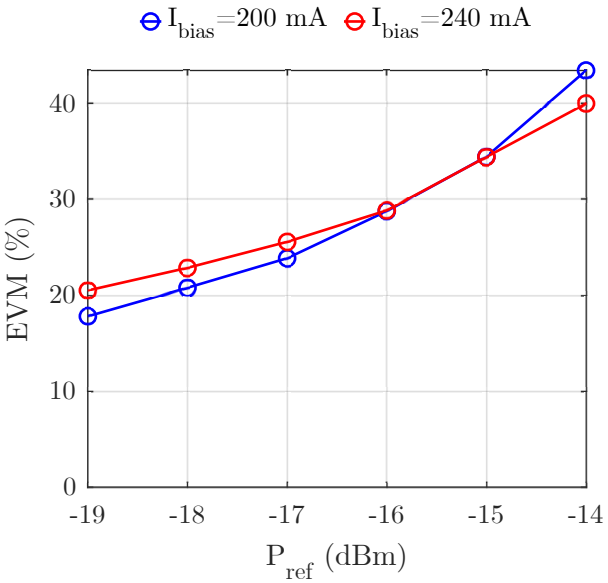


Figure 4.18 – Average EVM of $I_{bias} = 200$ mA (AWG amplitude of 300 mV) and $I_{bias} = 240$ mA (AWG amplitude of 425 mV) at different input powers into the SOA for five measurements of an 8 GHz OFDM frame (150 OFDM data symbols in total)

The remaining results that will be presented focus on the case where $I_{bias} = 240$ mA. First, the EVM of the different predistorters is shown for different input powers into the SOA in Figure 4.19. From Fig. 4.19, it can be inferred that by using DPD the input power into the SOA can be increased from -15.75 dBm (no DPD) to -14 dBm (MP) and -13 dBm (EMP, GMP, and GEMP) while still meeting the EVM requirement of 30%. This corresponds to an increase in reach of 4.5 km (MP) and 7.1 km (EMP, GMP, and GEMP) accounting for the change in gain in the SOA and the attenuation in the fiber.

Using the SOA with $I_{bias} = 200$ mA and $I_{bias} = 240$ mA would allow us to push the power into the SOA to -15.75 dBm and stay within the EVM requirement of 30%. Taking into account the output powers of the SOA for $I_{bias} = 200$ mA (1.5 dBm) and $I_{bias} = 240$ mA (2.95 dBm), there is difference in reach of about 7.25 km in favor of the $I_{bias} = 240$ mA case.

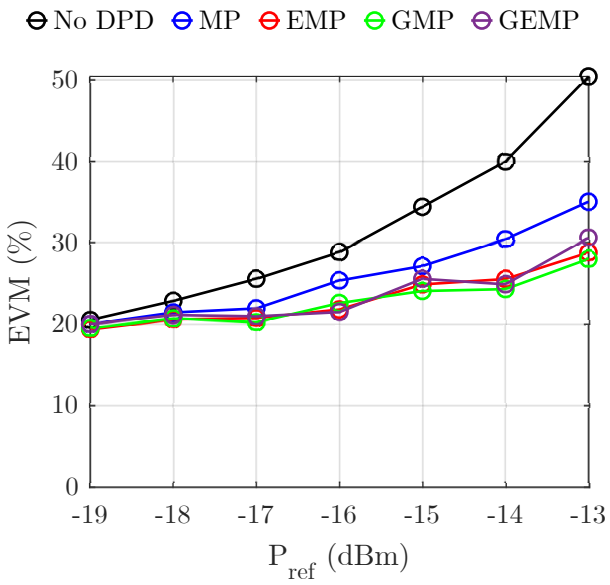


Figure 4.19 – Average EVM of multiple predistorters at different input powers into the SOA for five measurements of an 8 GHz OFDM frame. $I_{bias} = 240$ mA, 425 mV peak-to-peak AWG output

Figure 4.20 shows the baseband PSD with and without DPD at $P_{ref} = -13$ dBm. The no DPD case displays some distortions both in the passband where there is a slight

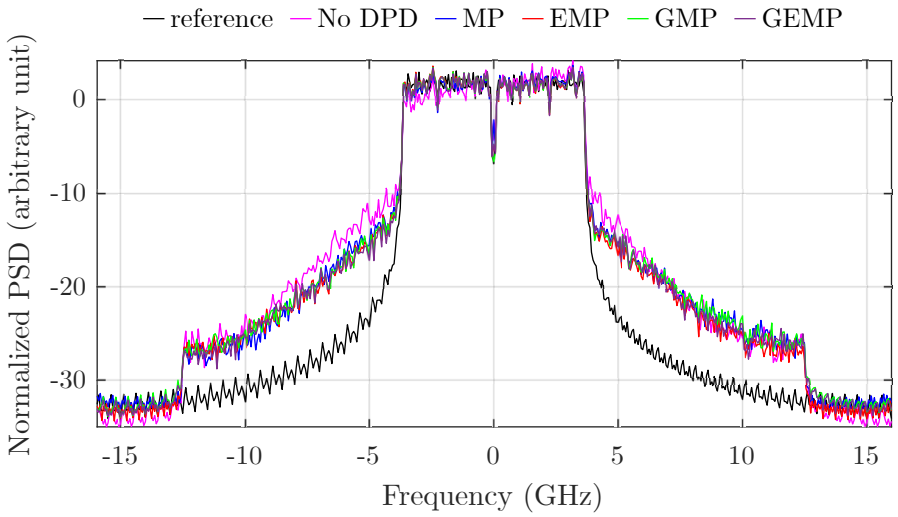


Figure 4.20 – PSD of multiple predistorters at $P_{ref} = -13$ dBm of an 8 GHz OFDM frame. $I_{bias} = 240$ mA, 425 mV peak-to-peak AWG output

tilt and in the adjacent channels where there is spectral regrowth. Inserting DPD rectifies the in-band effects of the SOA and decreases the out-band impact.

Figure 4.21 provides another way to visualize the benefits of DPD with the constellation of the no DPD and the different predistorter cases at $P_{ref} = -13$ dBm. The color coded constellation diagrams indicate that the highest densities of QAM symbols near the reference symbols belong to the EMP, GMP, and GEMP although the MP still performs better than the no DPD case.

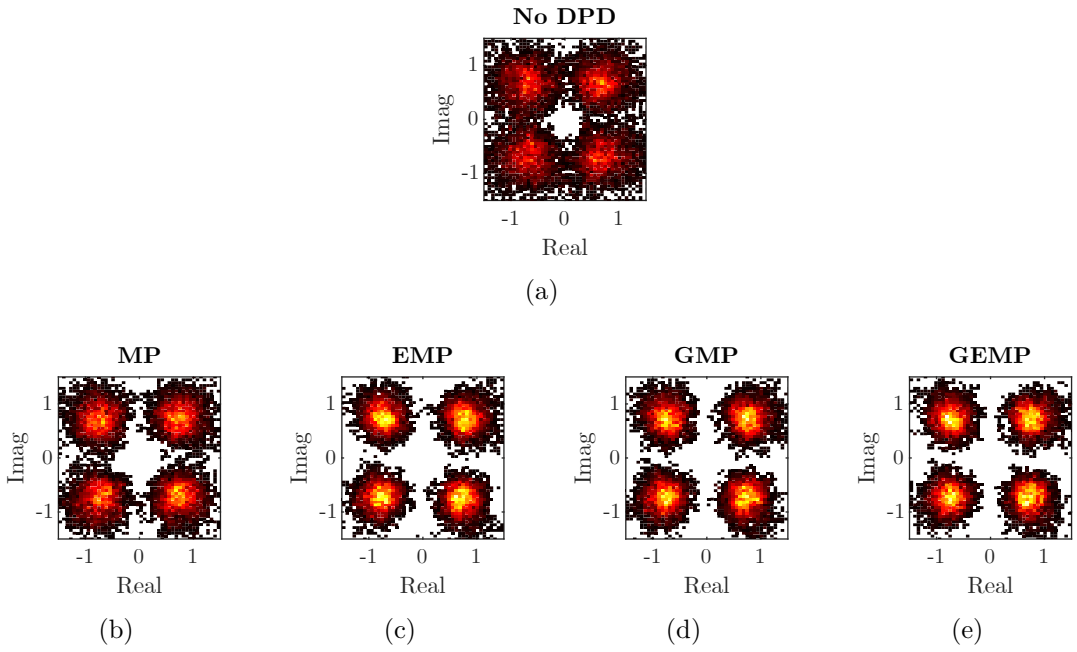


Figure 4.21 – Constellation diagrams of multiple predistorters at different input powers into the SOA for five measurements of an 8 GHz OFDM frame. $I_{bias} = 240$ mA, 425 mV peak-to-peak AWG output

The coefficients associated with each predistorter are the following: 12 for MP, 18 for EMP, 12 for GMP, and 17 for GEMP. In that regard, it seems that the GMP and the GEMP come out as top choices for DPD algorithms as they give the best performances for a lower complexity. When the bias current increases, the GMP stands out even more in terms of low complexity. A final important note is that the predistorters with the better results (EMP, GMP, GEMP) showcased more emphasis on memory depth than nonlinearity orders which would be in line with a more dispersive channel.

4.4 Conclusion

The mitigation at the emitter of nonlinear impairments caused by a booster SOA in a CO-OFDM access/metropolitan network application was investigated experimentally.

First, the experimental setup and tesbed was presented. The linearity of the emitter side of the transmission was established without SOA by looking at the amplitudes of the AWG, the RF drivers, and the IQ modulator. The impact of the correct operating point of the IQ modulator was evaluated as well as the frequency response of the AWG.

As the baseband processing is implemented on MATLAB, a comparison of the MATLAB receiver and the Keysight VSA receiver showed a good agreement between the two receiver's EVM values. Out of all the linear effects of the transmission, the laser phase noise was shown to be impactful while still being well corrected through phase tracking. The laser phase noise was also suspected to have some influence in the learning phase of DPD.

Regardless, all DPD algorithms showed valuable results for different bias currents supplying the SOA. Some of these results were part of a conference contribution at the IPC 2021 [Sim+21d]. The GEMP was proven to be a worthy general low complexity predistorter whose scope can go beyond that of coherent optical applications. The GMP was the best choice of predistorter to keep low complexity on top of improving the quality of transmission.

Conclusion and perspectives

This thesis work was centered on the influence of digital predistortion (DPD) in semiconductor optical amplifier (SOA) based coherent optical orthogonal frequency division multiplexing (CO-OFDM) systems for access/metropolitan networks. Indeed, with the push for transparency in optical networks, the SOA places itself as a great prospect for optical amplification. The SOA is compact and integrable, low cost, and has a large optical bandwidth; all these qualities are valuable in the eyes of future access networks. On the other hand, orthogonal frequency division multiplexing (OFDM) modulation formats have the added value of a capacity for high spectral efficiency, simple digital signal processing (DSP) for its construction and the ability to allow for parallelization, a dynamic allocation of frequency bands, and a resilience to channel dispersions. The benefits of an SOA based CO-OFDM system are obvious; however, the intrinsic dynamic changes in instantaneous power of OFDM unfortunately tend to excite the nonlinearities in the SOA. This disadvantage in combining SOAs and OFDM modulation calls for linearization techniques. DPD is an attractive solution to this problem as it is known to be simple for implementation, flexible (a different predistorter can be selected if the format of the signal changes), relatively low cost, and effective.

The study of DPD in the context of this thesis was done through three types of analyses: numerical, stochastic, and experimental. For the numerical analysis, first a comparison was made between three known predistorters in the radio frequency field: memory polynomials (MP), envelope memory polynomials (EMP), generalized memory polynomials (GMP). In photonics, DPD has to be tackled in a different manner since a closed loop implementation is not exactly feasible since the amplifier is part of a bigger system with electro-optic conversion meaning the access to the SOA's input and output is limited. Photonics also deals with larger bandwidth and data rates which plays a significant role in the allowed complexity of a predistorter. The Hill-Climbing heuristic was validated as a reliable way to find the structural parameters of a predistorter and a comparative analysis revealed that the GMP gave slightly better performances than the MP and EMP without adding much in terms of parametric complexity. A deeper study of the GMP structure seemed relevant so an analysis of GMP in the context of increasing

data rate was performed. This analysis gave reason to believe that one could increase data rate by sending signals with larger bandwidths and still reap the benefits of DPD. More interestingly, the analysis confirmed that the same predistorter can be used regardless of the number of subcarriers per OFDM symbol. Similarly so, an investigation of an in-line SOA scenario revealed that the same predistorter identified without the fiber could be used once the fiber is added and still be effective up to a certain point. The final step in the numerical analysis was to contrast DPD and digital post-distortion (DPoD); here the comparison was done by using the same structure at the emitter vs. at the receiver. It seemed to be that opting to send a cleaner signal at the emitter was the better option though DPoD is still a viable solution when the SOA is closer to the receiver side.

The results obtained during the numerical analysis showed that DPD reduced the impact of the OFDM frame variability and could still function properly when the operating point of the SOA in which it was identified fluctuated. It seemed pertinent to test the validity of DPD through a sensitivity analysis where joint parameter variation is applied on the IQ modulator and the SOA. These parameters can mimic the variability of the environment, component fabrication or calibration in real applications. Varying the bias current of the SOA has an intuitive impact on the SOA behavior; the impact of varying the branches gain, peak-to-peak voltage and the phase difference of the IQ modulator is less obvious but quite apparent as it changes the input power into the SOA. To perform the stochastic analysis, polynomial chaos expansion (PCE) is found to be a reliable way to bypass the computational costs of Monte Carlo in the context of an optical transmission with SOAs showing a speedup ratio of 200. With PCE, DPD is discovered to be all the more useful for a booster SOA with and without fiber and in the presence of digital to analog conversion (DAC) quantization noise, and IQ imbalance since it reduces the impact of system uncertainties on top of improving the global quality of transmission.

DPD is discovered to be all the more useful for a booster SOA with and without fiber, and in the presence of DAC quantization noise and IQ imbalance since it reduces the impact of system uncertainties on top of improving the global quality of transmission. The experimental analysis is a way to verify the benefits of DPD under real system uncertainties beyond the limit of three parameters chosen during the stochastic analysis. It is also a way to make sure that the use of DPD is still valid when the simplifying assumptions made during the numerical analysis do not necessarily hold in a real environment (laser phase noise present, influence of the optical signal to noise ratio (OSNR)). The experimental analysis confirmed that DPD is able to lower the error vector magnitude (EVM)

computed at the receiver and can therefore increase transmission reach.

The analyses performed during this thesis demonstrated the value of using DPD for SOA based CO-OFDM systems. Further research can be done to solidify the findings of this thesis work. Other modulation formats that deal with non-constant envelopes can be tested. This would help confirm that the findings of this thesis are applicable in other contexts and the question of DPD implementation for example for a filter bank multicarrier (FBMC) modulation could be addressed. Other formats of OFDM like constant envelope OFDM (CE-OFDM) could also be explored. Reducing the issue of peak to average power ratio (PAPR) might allow DPD to be more effective since the dynamic and memory effects of the SOA should not be as present. This would facilitate nonlinearity characterization of the predistorters and activating the DPD block might be needed only at much higher input powers of the SOA. On top of that, the implementation of polarization division multiplexing (PDM) is a point to investigate as it brings higher data rates. Here, the way to identify the predistorter while compensating the polarization mode dispersion (PMD) will be issues to tackle.

The case of an intensity-modulation/direct-detection (IM/DD) scenario seems interesting as well where envelope based predistorters could be evaluated. During our thesis work, it was seen that envelope based predistorters were quite useful especially at high input powers. There is an open dialogue to have in what happens when a predistorter does not perform phase correction as is the case for coherent detection. There will also be a need to look at how will uncertain parameters in the system affect the efficiency of envelope based predistorters.

Further analysis can be done using other types of amplifiers like the reflective SOA (RSOA) which has functions that go beyond simple amplification or the gain clamped SOA (GC-SOA) which has a more stable gain characteristic. Exploring the scenario of cascaded SOAs seem pertinent as well where both the added nonlinearities and noise will cause some issues, and the process of identifying predistorters might have to be adjusted. In this case, the combination of intelligently weighted DPD and DPoD could prove useful or mixing DPD with envelope tracking (ET).

List of publications

Journal paper

[Sim+20] Jacqueline E. Sime, Pascal Morel, Mohamad Younes, Igor Simone Stievano, Mihai Telescu, Noël Tanguy, and Stéphane Azou, « The Effects of Digital Predistortion in a CO-OFDM System—A Stochastic Approach », *in*: *IEEE Photonics Technology Letters* 32.*13* (2020), Publisher: IEEE, pp. 763–766

Conference papers

[Sim+21c] Jacqueline E. Sime, Pascal Morel, Mihai Telescu, Noël Tanguy, and Stéphane Azou, « Digital Predistortion for CO-OFDM Systems Using Generalized Memory Polynomials », *in*: *2020 IEEE Eighth International Conference on Communications and Electronics (ICCE)*, IEEE, Jan. 2021, pp. 7–11

[Sim+21b] Jacqueline E. Sime, Pascal Morel, Igor Simone Stievano, Mihai Telescu, Noël Tanguy, and Stéphane Azou, « SOAs and Digital Linearization in Optical Networks–A Stochastic Investigation », *in*: *Conference on Lasers and Electro-Optics*, Optical Society of America, May 2021, STh1F–6

[Sim+21a] Jacqueline E. Sime, Pascal Morel, Igor Simone Stievano, Mihai Telescu, Noël Tanguy, and Stéphane Azou, « Analyse stochastique de la linéarisation en bande de base des SOAs dans les réseaux optiques », *in*: *Journée Nationale d'Optique Guidée*, Dijon, France, July 2021

[Sim+21d] Jacqueline E. Sime, Pascal Morel, Mihai Telescu, Noël Tanguy, and Stéphane Azou, « Experimental Investigation of the Linearization of an SOA based CO-OFDM System », *in*: *IEEE Photonics Conference*, Vancouver, Canada, Oct. 2021, ThE1.5

[Fru+21] Alexandru Frunza, Jacqueline E. Sime, Vincent Choqueuse, Pascal Morel, and Stéphane Azou, « Joint Estimation and Compensation of Transmitter IQ imbalance and Laser Phase Noise in Coherent Optical Systems », *in*: *IEEE Photonics Conference*, Vancouver, Canada, Oct. 2021, ThE1.4

BIBLIOGRAPHY

[Agr12] Govind P Agrawal, *Fiber-optic communication systems*, vol. 222, John Wiley & Sons, 2012.

[Agr+16] Erik Agrell, Magnus Karlsson, AR Chraplyvy, David J Richardson, Peter M Krummrich, Peter Winzer, Kim Roberts, Johannes Karl Fischer, Seb J Savory, Benjamin J Eggleton, et al., « Roadmap of optical communications », *in*: *Journal of Optics* 18.*6* (2016), Publisher: IOP Publishing, p. 063002.

[Ami+15] Siamak Amiralizadeh, An T Nguyen, and Leslie A Rusch, « Modeling and compensation of transmitter nonlinearity in coherent optical OFDM », *in*: *Optics express* 23.*20* (2015), Publisher: Optical Society of America, pp. 26192–26207.

[Arm09] Jean Armstrong, « OFDM for optical communications », *in*: *Journal of lightwave technology* 27.*3* (2009), Publisher: IEEE, pp. 189–204.

[Azo+14] Stéphane Azou, Serban Bejan, Pascal Morel, and Ammar Sharaiha, « A comparative study of nonlinear companding schemes for CO-OFDM transmissions », *in*: *2014 13th International Conference on Optical Communications and Networks (ICOCN)*, IEEE, 2014, pp. 1–4.

[Bar+90] John R Barry and Edward A Lee, « Performance of coherent optical receivers », *in*: *Proceedings of the IEEE* 78.*8* (1990), Publisher: IEEE, pp. 1369–1394.

[Bar+18a] Sylvain Barthomeuf, Fabienne Saliou, Luiz Anet Neto, Philippe Chanclou, and Didier Erasme, « TDM-PON PAM downstream transmission for 25 Gbit/s and beyond », *in*: *Photonics*, vol. 5, Issue: 4, Multidisciplinary Digital Publishing Institute, 2018, p. 45.

[Bar+18b] Sylvain Barthomeuf, Fabienne Saliou, L Anet Neto, Philippe Chanclou, and Didier Erasme, « High optical budget 25Gbit/s PON with PAM4 and optically amplified O band downstream transmission », *in*: *2018 European Conference on Optical Communication (ECOC)*, IEEE, 2018, pp. 1–3.

[Bej+15] Ș Bejan, Stéphane Azou, Pascal Morel, Cherif Diouf, Mihai Telescu, Noel Tanguy, and Ammar Sharaiha, « A joint linearization/companding approach for improving a CO-OFDM transmitter », *in*: *IEEE photonics technology letters* 27.*20* (2015), Publisher: IEEE, pp. 2162–2165.

[Ben+79] Sergio Benedetto, Ezio Biglieri, and Riccardo Daffara, « Modeling and performance evaluation of nonlinear satellite links-a Volterra series approach », *in*: *IEEE Transactions on Aerospace and Electronic Systems 4* (1979), Publisher: IEEE, pp. 494–507.

[Ber+15] Pablo Wilke Berenguer, Markus Nölle, Lutz Molle, Talha Raman, Antonio Napoli, Colja Schubert, and Johannes Karl Fischer, « Nonlinear digital predistortion of transmitter components », *in*: *Journal of lightwave technology* 34.*8* (2015), Publisher: IEEE, pp. 1739–1745.

[Bla+11] Géraud Blatman and Bruno Sudret, « Adaptive sparse polynomial chaos expansion based on least angle regression », *in*: *Journal of computational Physics* 230.*6* (2011), Publisher: Elsevier, pp. 2345–2367.

[Blu+19] Christian Bluemm, Maximilian Schaedler, Maxim Kuschnerov, Fabio Pittalà, and Changsong Xie, « Single carrier vs. OFDM for coherent 600Gb/s data centre interconnects with nonlinear equalization », *in*: *2019 Optical Fiber Communications Conference and Exhibition (OFC)*, IEEE, 2019, pp. 1–3.

[Bor+14] Robert Borkowski, Darko Zibar, and Idelfonso Tafur Monroy, « Anatomy of a digital coherent receiver », *in*: *IEICE Transactions on Communications* 97.*8* (2014), Publisher: The Institute of Electronics, Information and Communication Engineers, pp. 1528–1536.

[Bor+04] P Borwein and R Ferguson, « A complete description of Golay pairs for lengths up to 100 », *in*: *Mathematics of computation* 73.*246* (2004), pp. 967–985.

[Bud+15] Russell A Budd, Laurent Schares, Benjamin G Lee, Fuad E Doany, Christian Baks, Daniel M Kuchta, Clint L Schow, and Frank Libsch, « Semiconductor optical amplifier (SOA) packaging for scalable and gain-integrated silicon photonic switching platforms », *in*: *2015 IEEE 65th Electronic Components and Technology Conference (ECTC)*, IEEE, 2015, pp. 1280–1286.

[Car+15] Filipe Carvalho and Adolfo Cartaxo, « Broad baseband nonlinear distortion mitigation using digital pre-and post-distortion in OFDM-based WDM LR-PON », *in*: *Optics express* 23.*6* (2015), Publisher: Optical Society of America, pp. 7062–7074.

[Che+16] Qun Chen, Feng Hu, Yang Song, and Chaoshi Cai, « Effects of even-order terms on behavior models in predistortion linearization », *in*: *2016 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, IEEE, 2016, pp. 860–865.

[Che+13] Zhiyu Chen, Lianshan Yan, Wei Pan, Bin Luo, Xihua Zou, Yinghui Guo, Hengyun Jiang, and Tao Zhou, « SFDR enhancement in analog photonic links by simultaneous compensation for dispersion and nonlinearity », *in*: *Optics express* 21.*18* (2013), Publisher: Optical Society of America, pp. 20999–21009.

[Cis20] Cisco, U, « Cisco annual internet report (2018-2023) white paper », *in*: (2020), URL: `https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.pdf`.

[Con07] Michael J Connelly, *Semiconductor optical amplifiers*, Springer Science & Business Media, 2007.

[Dia+17] Vinicius OC Dias, Ezequiel da V Pereira, Helder RO Rocha, Marcelo EV Segatto, and Jair AL Silva, « Performance evaluation of CO-OFDM systems based on electrical constant-envelope signals », *in*: *Optical Fiber Technology* 37 (2017), Publisher: Elsevier, pp. 30–34.

[Din+04] Lei Ding and Guo Tong Zhou, « Effects of even-order nonlinear terms on power amplifier modeling and predistortion linearization », *in*: *IEEE Transactions on Vehicular Technology* 53.*1* (2004), Publisher: IEEE, pp. 156–162.

[Dio+17] Cherif Diouf, Mohamad Younes, Alexandru Noaja, Stéphane Azou, Mihai Telescu, Pascal Morel, and Noël Tanguy, « Robustness analysis of a parallel two-box digital polynomial predistorter for an SOA-based CO-OFDM system », *in*: *Optics Communications* 402 (2017), Publisher: Elsevier, pp. 442–452.

[Efr+04] Bradley Efron, Trevor Hastie, Iain Johnstone, and Robert Tibshirani, « Least angle regression », *in*: *The Annals of statistics* 32.*2* (2004), Publisher: Institute of Mathematical Statistics, pp. 407–499.

[Egh+14] Amir Eghbali, Håkan Johansson, Oscar Gustafsson, and Seb J Savory, « Optimal least-squares FIR digital filters for compensation of chromatic dispersion in digital coherent optical receivers », *in*: *Journal of Lightwave Technology* 32.*8* (2014), Publisher: IEEE, pp. 1449–1456.

[Elg+14] Hany Elgala and Thomas DC Little, « SEE-OFDM: Spectral and energy efficient OFDM for optical IM/DD systems », *in*: *2014 IEEE 25th annual international symposium on personal, indoor, and mobile radio communication (PIMRC)*, IEEE, 2014, pp. 851–855.

[Est+18] José Manuel Estarán, Haik Mardoyan, Filipe Jorge, Oskars Ozolins, Aleksejs Udalcovs, Agnieszka Konczykowska, Muriel Riet, Bernadette Duval, Virginie Nodjiadjim, Jean-Yves Dupuy, et al., « 140/180/204-Gbaud OOK transceiver for inter-and intra-data center connectivity », *in*: *Journal of Lightwave Technology* 37.*1* (2018), Publisher: IEEE, pp. 178–187.

[Far+17] Md Saifuddin Faruk and Seb J Savory, « Digital signal processing for coherent transceivers employing multilevel formats », *in*: *Journal of Lightwave Technology* 35.*5* (2017), Publisher: IEEE, pp. 1125–1141.

[Fat+08] Irshaad Fatadin, Seb J Savory, and David Ives, « Compensation of quadrature imbalance in an optical QPSK coherent receiver », *in*: *IEEE Photonics Technology Letters* 20.*20* (2008), Publisher: IEEE, pp. 1733–1735.

[Fre+13] Francesco Fresi, Marco Secondini, Gianluca Berrettini, Gianluca Meloni, and Luca Poti, « Impact of optical and electrical narrowband spectral shaping in faster than Nyquist Tb superchannel », *in*: *IEEE Photonics Technology Letters* 25.*23* (2013), Publisher: IEEE, pp. 2301–2303.

[Fru+21] Alexandru Frunza, Jacqueline E. Sime, Vincent Choqueuse, Pascal Morel, and Stéphane Azou, « Joint Estimation and Compensation of Transmitter IQ imbalance and Laser Phase Noise in Coherent Optical Systems », *in*: *IEEE Photonics Conference*, Vancouver, Canada, Oct. 2021, ThE1.4.

[Gam+16] Heshani Gamage, Nandana Rajatheva, and Matti Latva-aho, « High PAPR Sequence Scrambling for Reducing OFDM Peak-to-Average Power Ratio », *in*: *European Wireless 2016; 22th European Wireless Conference*, VDE, 2016, pp. 1–5.

[Gha+09] Fadhel M Ghannouchi and Oualid Hammi, « Behavioral modeling and predistortion », *in*: *IEEE Microwave magazine* 10.*7* (2009), Publisher: IEEE, pp. 52–64.

[Gha+15] Fadhel M Ghannouchi, Oualid Hammi, and Mohamed Helaoui, *Behavioral modeling and predistortion of wideband wireless transmitters*, John Wiley & Sons, 2015.

[Gia+10] E Giacoumidis, JL Wei, XL Yang, A Tsokanos, and JM Tang, « Adaptive-modulation-enabled WDM impairment reduction in multichannel optical OFDM transmission systems for next-generation PONs », *in*: *IEEE Photonics Journal* 2.*2* (2010), Publisher: IEEE, pp. 130–140.

[Gil+20] Pere L Gilabert, R Neil Braithwaite, and Gabriel Montoro, « Beyond the Moore–Penrose inverse: Strategies for the estimation of digital predistortion linearization parameters », *in*: *IEEE Microwave Magazine* 21.*12* (2020), Publisher: IEEE, pp. 34–46.

[Gol61] Marcel Golay, « Complementary series », *in*: *IRE transactions on information theory* 7.*2* (1961), Publisher: IEEE, pp. 82–87.

[Ham+15] Mohamad Hamze, Ammar Sharaiha, Pascal Morel, Mikael Guegan, and Ali Hamié, « Validation expérimentale d'une chaine de co-simulation d'émission et de réception de formats IM-OFDM à base de RSOA », *in*: *Optique Bretagne, Journées Nationales d'Optique Guidée (JNOG'35)*, 2015.

[Hil+10] D Hillerkuss, M Winter, M Teschke, A Marculescu, J Li, G Sigurdsson, K Worms, S Ben Ezra, N Narkiss, W Freude, et al., « Simple all-optical FFT scheme enabling Tbit/s real-time signal processing », *in*: *Optics express* 18.*9* (2010), Publisher: Optical Society of America, pp. 9324–9340.

[Hor+20] Tomas Horvath, Jan Radil, Petr Munster, and Ning-Hai Bao, « Optical Amplifiers for Access and Passive Optical Networks: A Tutorial », *in*: *Applied Sciences* 10.*17* (2020), Publisher: Multidisciplinary Digital Publishing Institute, p. 5912.

[Ibr+17] AN Ibrahim and MFL Abdullah, « The potential of FBMC over OFDM for the future 5G mobile communication technology », *in*: *AIP Conference Proceedings*, vol. 1883, Issue: 1, AIP Publishing LLC, 2017, p. 020001.

[Idl+17] Wilfried Idler and Fred Buchali, « Higher-order modulation formats–concepts and enabling devices », *in*: *Fibre Optic Communication*, Springer, 2017, pp. 291–357.

[IEE17] IEEE 802.11ah, *802.11ah-2016 - IEEE Standard for Information technology Telecommunications and information exchange between systems - Local and metropolitan area networks–Specific requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY)*. OCLC: 1112400862, IEEE, 2017, ISBN: 978-1-5044-3911-4.

[ITU18] ITU-T G Suppl. 64, *PON transmission technologies above 10 Gbit/s per wavelength*, ITU, Feb. 2018, URL: http://handle.itu.int/11.1002/1000/13589.

[ITU09] ITU-T G.698.1, *Multichannel DWDM applications with single-channel optical interfaces*, ITU, Nov. 2009, URL: https://www.itu.int/rec/T-REC-G.698.1-200911-I/en.

[ITU20a] ITU-T G.8310, *Architecture of the metro transport network*, ITU, Dec. 2020, URL: http://handle.itu.int/11.1002/1000/14516.

[ITU20b] ITU-T G.9807.1 Amendment 2, *10-Gigabit-capable symmetric passive optical network (XGS-PON) Amendment 2*, ITU, Oct. 2020, URL: https://www.itu.int/rec/T-REC-G.9807.1-202010-I!Amd2/en.

[Jia16] Tao Jiang, « Behavioral modeling and FPGA implementation of digital predistortion for RF and microwave power amplifiers », *in*: *Politecnico di Torino, Tech. Rep* 116 (2016).

[Jin+20] Wei Jin, Abdulai Sankoh, Yixian Dong, Zhu-Qiang Zhong, Roger Philip Giddings, Maurice O'Sullivan, Jeffrey Lee, Tim Durrant, and Jianming Tang, « Hybrid SSB OFDM-digital filter multiple access PONs », *in*: *Journal of Lightwave Technology* 38.*8* (2020), Publisher: IEEE, pp. 2095–2105.

[Jun+16] Sang-Min Jung, Kyoung-Hak Mun, Sun-Young Jung, and Sang-Kook Han, « IQ imbalance tolerable parallel-channel DMT transmission for coherent optical OFDMA access network », *in*: *Optical Fiber Technology* 32 (2016), Publisher: Elsevier, pp. 43–49.

[Kah06] Joseph M Kahn, « Modulation and detection techniques for optical communication systems », *in*: *Coherent Optical Technologies and Applications*, Optical Society of America, 2006, CThC1.

[Kai+18] Arun Kaintura, Tom Dhaene, and Domenico Spina, « Review of polynomial chaos-based methods for uncertainty quantification in modern integrated circuits », *in*: *Electronics* 7.*3* (2018), Multidisciplinary Digital Publishing Institute, p. 30.

[Key16] Keysight Technologies, *M8195A 65 GSa/s arbitrary waveform generator and M8197 multi-channel syncrhonization module - data sheet*, 2016, URL: `http://literature.cdn.keysight.com/litweb/pdf/5992-0014EN.pdf`.

[Kha12] Hamidreza Khaleghi, « Influence des amplificateurs optiques à semi-conducteurs (SOA) sur la transmission cohérente de signaux optiques à format de modulation multi-porteuses (CO-OFDM) », PhD Thesis, Université de Bretagne occidentale-Brest, 2012.

[Kha+12] Hamidreza Khaleghi, Pascal Morel, Ammar Sharaiha, and Thierry Rampone, « Experimental validation of numerical simulations and performance analysis of a coherent optical-OFDM transmission system employing a semiconductor optical amplifier », *in*: *Journal of lightwave technology* 31.*1* (2012), Publisher: IEEE, pp. 161–170.

[Kik08] Kazuro Kikuchi, « Polarization-demultiplexing algorithm in the digital coherent receiver », *in*: *2008 Digest of the IEEE/LEOS Summer Topical Meetings*, IEEE, 2008, pp. 101–102.

[Koe+14] Swen Koenig, Rene Bonk, Harald Schmuck, Wolfgang Poehlmann, Th Pfeiffer, Christian Koos, Wolfgang Freude, and Juerg Leuthold, « Amplification of advanced modulation formats with a semiconductor optical amplifier cascade », *in*: *Optics express* 22.*15* (2014), Publisher: Optical Society of America, pp. 17854–17871.

[Kuc+12] Sergei Kucherenko, Stefano Tarantola, and Paola Annoni, « Estimation of global sensitivity indices for models with dependent variables », *in*: *Computer physics communications* 183.*4* (2012), Publisher: Elsevier, pp. 937–946.

[Kum11] Shiva Kumar, *Impact of nonlinearities on fiber optic communications*, vol. 7, Springer Science & Business Media, 2011.

[Laf+07] J Laferrière, G Lietaert, R Taws, and S Wolszczok, *Reference guide to fiber optic testing*, JDS Uniphase Corporation, 2007.

[Li+16] Chao Li and Qi Yang, « Optical OFDM/OQAM for the future fiber-optics communications », *in*: *Procedia engineering* 140 (2016), Publisher: Elsevier, pp. 99–106.

[Lia+06] Shien-Kuei Liaw, Cheng-Kai Huang, Yuan–Lung Hsiao, Li-Chuan Mau, Hong-Xi Cao, Ji-Bin Horng, and Keang-Po Ho, « Low noise-figure miniature erbium-doped fibre amplifier using uncooled pump laser », *in*: *Optical and quantum electronics* 38.*7* (2006), Publisher: Springer, pp. 625–631.

[Liu+16] Anliang Liu, Xin Wang, Qi Shao, Teng Song, Hongxi Yin, and Nan Zhao, « A low cost structure of radio-over-fiber system compatible with WDM-PON », *in*: *2016 25th Wireless and optical communication conference (Wocc)*, IEEE, 2016, pp. 1–3.

[Low+07] Arthur James Lowery, Liang Bangyuan Du, and Jean Armstrong, « Performance of optical OFDM in ultralong-haul WDM lightwave systems », *in*: *Journal of Lightwave Technology* 25.*1* (2007), Publisher: IEEE, pp. 131–138.

[Man+17a] Paolo Manfredi, Dries Vande Ginste, Igor S Stievano, Daniël De Zutter, and Flavio G Canavero, « Stochastic transmission line analysis via polynomial chaos methods: an overview », *in*: *IEEE Electromagnetic Compatibility Magazine* 6.*3* (2017), Publisher: IEEE, pp. 77–84.

[Man+17b] CT Manimegalai, Sabitha Gauni, Nikita Raghavan, and T Rama Rao, « Investigations on suitable modulation techniques for visible light communications », *in*: *2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, IEEE, 2017, pp. 1818–1822.

[Mar+09] Nicola Marchetti, Muhammad Imadur Rahman, Sanjay Kumar, and Ramjee Prasad, « OFDM: Principles and challenges », *in*: *New directions in wireless communications research*, Springer, 2009, pp. 29–62.

[Mar+18] PRN Marciano, DVN Coelho, JAL Silva, MJ Pontes, and MEV Segatto, « A New All-Optical OFDM Architecture for NG-PON2 », *in*: *Latin America Optics and Photonics Conference*, Optical Society of America, 2018, W3D–4.

[Mar+14] Stefano Marelli and Bruno Sudret, « UQLab: A framework for uncertainty quantification in Matlab », *in*: *Vulnerability, uncertainty, and risk: quantification, mitigation, and management*, 2014, pp. 2554–2563.

[McK+00] Michael D McKay, Richard J Beckman, and William J Conover, « A comparison of three methods for selecting values of input variables in the analysis of output from a computer code », *in*: *Technometrics* 42.*1* (2000), Publisher: Taylor & Francis, pp. 55–61.

[Min+03] Hlaing Minn, Vijay K Bhargava, and Khaled Ben Letaief, « A robust timing and frequency synchronization for OFDM systems », *in*: *IEEE Transactions on Wireless communications* 2.*4* (2003), Publisher: IEEE, pp. 822–839.

[Mor06] Pascal Morel, « Modélisation des amplificateurs optiques à semi-conducteurs: du composant au système », PhD Thesis, Université de Bretagne occidentale-Brest, 2006.

[Mor+08] Pascal Morel and Ammar Sharaiha, « Wideband time-domain transfer matrix model equivalent circuit for short pulse propagation in semiconductor optical amplifiers », *in*: *IEEE Journal of Quantum Electronics* 45.*2* (2008), Publisher: IEEE, pp. 103–116.

[Mor+06] Dennis R Morgan, Zhengxiang Ma, Jaehyeong Kim, Michael G Zierdt, and John Pastalan, « A generalized memory polynomial model for digital predistortion of RF power amplifiers », *in*: *IEEE Transactions on signal processing* 54.*10* (2006), Publisher: IEEE, pp. 3852–3860.

[El-+20] Fady I El-Nahal and Nabih M Abualoff, « Coherent optical communication systems based on orthogonal frequency division multiplexing », *in*: *Optoelectronics Letters* 16.*4* (2020), Publisher: Springer, pp. 303–305.

[Nie92] Harald Niederreiter, *Random number generation and quasi-Monte Carlo methods*, SIAM, 1992.

[Nun+14] Reginaldo B Nunes, R de O Helder, Marcelo EV Segatto, and Jair AL Silva, « Experimental validation of a constant-envelope OFDM system for optical direct-detection », *in*: *Optical Fiber Technology* 20.*3* (2014), Publisher: Elsevier, pp. 303–307.

[Off+17] Funmilayo B Offiong, Sinan Sinanović, and Wasiu O Popoola, « On PAPR reduction in pilot-assisted optical OFDM communication systems », *in*: *IEEE Access* 5 (2017), Publisher: IEEE, pp. 8916–8929.

[Ort+20] Julio Cesar Ortiz-Cornejo, Pascal Morel, Stéphane Azou, and Jorge Arturo Pardiñas-Mir, « A numerical assessment of an effective envelope-tracking semiconductor optical amplifier design for coherent-optical OFDM transmission », *in*: *Optics Communications* 454 (2020), Publisher: Elsevier, p. 124474.

[Ozd+07] Mehmet Kemal Ozdemir and Huseyin Arslan, « Channel estimation for wireless OFDM systems », *in*: *IEEE Communications Surveys & Tutorials* 9.*2* (2007), Publisher: IEEE, pp. 18–48.

[Pec+09] Vladimir Pechenkin and Ivan J Fair, « Correlation between peak-to-average power ratio and four-wave mixing in optical OFDM systems », *in*: *Journal of Optical Communications and Networking* 1.*7* (2009), Publisher: Optical Society of America, pp. 636–644.

[Per+15] Jose Krause Perin, Milad Sharif, and Joseph M Kahn, « Modulation schemes for single-laser 100 Gb/s links: multicarrier », *in*: *Journal of Lightwave Technology* 33.*24* (2015), Publisher: IEEE, pp. 5122–5132.

[Peu12] Christophe Peucheret, « Generation and detection of optical modulation formats », *in*: *Dept. of Photonics Engg, Technical University of Denmark* (2012).

[Rah+13] Yasir Rahmatallah and Seshadri Mohan, « Peak-to-average power ratio reduction in OFDM systems: A survey and taxonomy », *in*: *IEEE communications surveys & tutorials* 15.*4* (2013), Publisher: IEEE, pp. 1567–1592.

[Ras+18] Ahmed Nabih Zaki Rashed and Mohammed Salah F Tabbour, « The trade off between different modulation schemes for maximum long reach high data transmission capacity optical orthogonal frequency division multiplexing (OOFDM) », *in*: *Wireless Personal Communications* 101.*1* (2018), Publisher: Springer, pp. 325–337.

[Rob+09] Kim Roberts, Maurice O'Sullivan, Kuang-Tsan Wu, Han Sun, Ahmed Awadalla, David J Krause, and Charles Laperle, « Performance of dual-polarization QPSK for optical transport systems », *in*: *Journal of lightwave technology* 27.*16* (2009), Publisher: IEEE, pp. 3546–3559.

[Ros+17] Narek Rostomyan, Jefy A Jayamon, and Peter M Asbeck, « 15 GHz Doherty power amplifier with RF predistortion linearizer in CMOS SOI », *in*: *IEEE Transactions on Microwave Theory and Techniques* 66.*3* (2017), Publisher: IEEE, pp. 1339–1348.

[Rou14] Sudhir K Routray, « The changing trends of optical communication », *in*: *IEEE Potentials* 33.*1* (2014), Publisher: IEEE, pp. 28–33.

[Rub+16] Reuven Y Rubinstein and Dirk P Kroese, *Simulation and the Monte Carlo method*, vol. 10, John Wiley & Sons, 2016.

[Sai+11] Youssef Said and Houria Rezig, « Numerical modeling and experimental investigation of the nonlinear polarization rotation phenomenon in semiconductor optical amplifiers », *in*: *Advances in OptoElectronics* 2011 (2011), Publisher: Hindawi.

[Sal+08] Andrea Saltelli, Marco Ratto, Terry Andres, Francesca Campolongo, Jessica Cariboni, Debora Gatelli, Michaela Saisana, and Stefano Tarantola, *Global sensitivity analysis: the primer*, John Wiley & Sons, 2008.

[San+15] Max Fréjus Sanya, Leopold Djogbe, A Vianou, and Christelle Aupetit-Berthelemot, « DC-biased optical OFDM for IM/DD passive optical network systems », *in*: *Journal of Optical Communications and Networking* 7.*4* (2015), Publisher: Optical Society of America, pp. 205–214.

[Sch76] Martin Schetzen, « Theory of pth-order inverses of nonlinear systems », *in*: *IEEE Transactions on Circuits and Systems* 23.*5* (1976), Publisher: IEEE, pp. 285–291.

[Sch+11] Rene Schmogrow, Bernd Nebendahl, Marcus Winter, Arne Josten, David Hillerkuss, Swen Koenig, Joachim Meyer, Michael Dreschmann, Michael Huebner, Christian Koos, et al., « Error vector magnitude as a performance measure for advanced modulation formats », *in*: *IEEE Photonics Technology Letters* 24.*1* (2011), Publisher: IEEE, pp. 61–63.

[Sch+14] H Schmuck, R Bonk, W Poehlmann, C Haslach, W Kuebart, D Karnick, J Meyer, D Fritzsche, E Weis, J Becker, et al., « Demonstration of an SOA-assisted open metro-access infrastructure for heterogeneous services », *in*: *Optics express* 22.*1* (2014), Publisher: Optical Society of America, pp. 737–748.

[Sch+13a] Jochen Schröder, Liang Bangyuan Du, Joel Carpenter, Benjamin J Eggleton, and Arthur J Lowery, « All-optical OFDM with cyclic prefix insertion using flexible wavelength selective switch optical processing », *in*: *Journal of lightwave technology* 32.*4* (2013), Publisher: IEEE, pp. 752–759.

[Sch+13b] Benjamin Schubert, Ahmet Gökceoglu, Lauri Anttila, and Mikko Valkama, « Augmented Volterra predistortion for the joint mitigation of power amplifier and I/Q modulator impairments in wideband flexible radio », *in*: *2013 IEEE Global Conference on Signal and Information Processing*, IEEE, 2013, pp. 1162–1165.

[Shi+06] William Shieh and C Athaudage, « Coherent optical orthogonal frequency division multiplexing », *in*: *Electronics letters* 42.*10* (2006), Publisher: IET, pp. 587–589.

[Shi+09] William Shieh and Ivan B Djordjevic, *OFDM for optical communications*, Academic press, 2009.

[Shi+08] William Shieh, Xingwen Yi, Yiran Ma, and Qi Yang, « Coherent optical OFDM: has its time come? », *in*: *Journal of Optical Networking* 7.*3* (2008), Publisher: Optical Society of America, pp. 234–255.

[Sim+21a] Jacqueline E. Sime, Pascal Morel, Igor Simone Stievano, Mihai Telescu, Noël Tanguy, and Stéphane Azou, « Analyse stochastique de la linéarisation en bande de base des SOAs dans les réseaux optiques », *in*: *Journée Nationale d'Optique Guidée*, Dijon, France, July 2021.

[Sim+21b] Jacqueline E. Sime, Pascal Morel, Igor Simone Stievano, Mihai Telescu, Noël Tanguy, and Stéphane Azou, « SOAs and Digital Linearization in Optical Networks–A Stochastic Investigation », *in*: *Conference on Lasers and Electro-Optics*, Optical Society of America, May 2021, STh1F–6.

[Sim+21c] Jacqueline E. Sime, Pascal Morel, Mihai Telescu, Noël Tanguy, and Stéphane Azou, « Digital Predistortion for CO-OFDM Systems Using Generalized Memory Polynomials », *in*: *2020 IEEE Eighth International Conference on Communications and Electronics (ICCE)*, IEEE, Jan. 2021, pp. 7–11.

[Sim+21d] Jacqueline E. Sime, Pascal Morel, Mihai Telescu, Noël Tanguy, and Stéphane Azou, « Experimental Investigation of the Linearization of an SOA based CO-OFDM System », *in*: *IEEE Photonics Conference*, Vancouver, Canada, Oct. 2021, ThE1.5.

[Sim+20] Jacqueline E. Sime, Pascal Morel, Mohamad Younes, Igor Simone Stievano, Mihai Telescu, Noël Tanguy, and Stéphane Azou, « The Effects of Digital Predistortion in a CO-OFDM System—A Stochastic Approach », *in*: *IEEE Photonics Technology Letters* 32.*13* (2020), Publisher: IEEE, pp. 763–766.

[Sin+19] Sarika Singh, Sandeep K Arya, and Shelly Singla, « A Study & Review of Various Optical Linearization Techniques for Next Generation RoF Networks », *in*: *2019 5th International Conference on Signal Processing, Computing and Control (ISPCC)*, IEEE, 2019, pp. 125–133.

[Sin+21] Sukwinder Singh and Jagannath Malik, « Review of efficiency enhancement techniques and linearization techniques for power amplifier », *in*: *International Journal of Circuit Theory and Applications* 49.*3* (2021), Publisher: Wiley Online Library, pp. 762–777.

[Sob93] Ilya M Sobol, « Sensitivity analysis for non-linear mathematical models », *in*: *Mathematical modelling and computational experiment* 1 (1993), pp. 407–414.

[Ste+07] Heidi Steendam and Marc Moeneclaey, « Different guard interval techniques for OFDM: performance comparison », *in*: *Multi-Carrier Spread Spectrum 2007*, Springer, 2007, pp. 11–24.

[Sud08] Bruno Sudret, « Global sensitivity analysis using polynomial chaos expansions », *in*: *Reliability engineering & system safety* 93.*7* (2008), Publisher: Elsevier, pp. 964–979.

[Sud14] Bruno Sudret, « Polynomial chaos expansions and stochastic finite element methods », *in*: *Risk and reliability in geotechnical engineering* (2014), Publisher: CRC Press Boca Raton, FL, pp. 265–300.

[Tal+19] Giuseppe Talli, Cleitus Antony, Mark Power, and Paul Townsend, « SOA-based metro-access coherent transmission systems », *in*: *CLEO: Science and Innovations*, Optical Society of America, 2019, SM4G–1.

[Teh+10] Ali Soltani Tehrani, Haiying Cao, Sepideh Afsardoost, Thomas Eriksson, Magnus Isaksson, and Christian Fager, « A comparative analysis of the complexity/accuracy tradeoff in power amplifier behavioral models », *in*: *IEEE Transactions on Microwave Theory and Techniques* 58.*6* (2010), Publisher: IEEE, pp. 1510–1520.

[Ter+02] John Terry and Juha Heiskala, *OFDM wireless LANs: A theoretical and practical guide*, Sams publishing, 2002.

[Tho+08] Steve C Thompson, Ahsen U Ahmed, John G Proakis, James R Zeidler, and Michael J Geile, « Constant envelope OFDM », *in*: *IEEE transactions on communications* 56.*8* (2008), Publisher: IEEE, pp. 1300–1312.

[Van96] Richard DJ Van Nee, « OFDM codes for peak-to-average power reduction and error correction », *in*: *Proceedings of GLOBECOM'96. 1996 IEEE Global Telecommunications Conference*, vol. 1, IEEE, 1996, pp. 740–744.

[Vas+17] Olga Vassilieva, Inwoong Kim, Tomofumi Oyama, Shoichiro Oda, Hisao Nakashima, Takeshi Hoshida, and Tadashi Ikeuchi, « Reach extension with 32-and 64 GBaud single carrier vs. multi-carrier signals », *in*: *2017 Optical Fiber Communications Conference and Exhibition (OFC)*, IEEE, 2017, pp. 1–3.

[VIA21] VIAVI Solutions, « Challenges in Next-Gen PON Deployment white paper », *in*: (2021), URL: `https://www.viavisolutions.com/en-us/literature/challenges-next-gen-pon-deployment-white-papers-books-en.pdf`.

[Vit83] Andrew Viterbi, « Nonlinear estimation of PSK-modulated carrier phase with application to burst digital transmission », *in*: *IEEE Transactions on Information theory* 29.*4* (1983), Publisher: IEEE, pp. 543–551.

[Wan+16] Siqi Wang, Mazen Abi Hussein, Olivier Venard, and Geneviève Baudoin, « Optimal sizing of generalized memory polynomial model structure based on hill-climbing heuristic », *in*: *2016 46th European Microwave Conference (EuMC)*, IEEE, 2016, pp. 190–193.

[Wan+18] Siqi Wang, Mazen Abi Hussein, Olivier Venard, and Genevieve Baudoin, « A novel algorithm for determining the structure of digital predistortion models », *in*: *IEEE Transactions on Vehicular Technology* 67.*8* (2018), Publisher: IEEE, pp. 7326–7340.

[Waq+21] Abi Waqas, Paolo Manfredi, and Daniele Melati, « Performance Variability Analysis of Photonic Circuits with Many Correlated Parameters », *in*: *Journal of Lightwave Technology* (2021), Publisher: IEEE.

[Wel67] Peter Welch, « The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms », *in*: *IEEE Transactions on audio and electroacoustics* 15.*2* (1967), Publisher: IEEE, pp. 70–73.

[Xie+12] Xiaojun Xie, Yitang Dai, Kun Xu, Jian Niu, Ruixin Wang, Li Yan, Yuefeng Ji, and Jintong Lin, « Digital joint compensation of IMD3 and XMD in broadband channelized RF photonic link », *in*: *Optics express* 20.*23* (2012), Publisher: Optical Society of America, pp. 25636–25643.

[Yan+19] Siyuan Yan, Xianzhen Yang, and Fu Li, « Effects of even-order nonlinear terms on dual-band power amplifier modeling », *in*: *Microwave and Optical Technology Letters* 61.*1* (2019), Publisher: Wiley Online Library, pp. 163–166.

[Yan+11] Qi Yang, Abdullah Al Amin, and William Shieh, « Optical OFDM basics », *in*: *Impact of nonlinearities on fiber optic communications*, Springer, 2011, pp. 43–85.

[Yi+07] Xingwen Yi, William Shieh, and Yan Tang, « Phase estimation for coherent optical OFDM », *in*: *IEEE Photonics Technology Letters* 19.*12* (2007), Publisher: IEEE, pp. 919–921.

[You+19] Biyu You, Liu Yang, Fengguang Luo, Songnian Fu, Shuailong Yang, Bin Li, and Deming Liu, « Joint carrier frequency offset and phase noise estimation based on pseudo-pilot in CO-FBMC/OQAM system », *in*: *IEEE Photonics Journal* 11.*1* (2019), Publisher: IEEE, pp. 1–11.

[You19] Mohamad Younes, « Techniques de pré-distorsion numérique pour l'optimisation des réseaux d'accès optiques », PhD Thesis, Université de Bretagne occidentale-Brest, 2019.

[You+20] Mohamad Younes, Mihai Telescu, Stéphane Azou, Pascal Morel, and Noël Tanguy, « Very low-complexity filter lookup-table design with non-uniform spacing for SOA linearisation », *in*: *IET Optoelectronics* 14.*5* (2020), Publisher: IET, pp. 306–311.

[Zha+20] Jiao Zhang, Jianjun Yu, Jun Shan Wey, Xinying Li, Li Zhao, Kaihui Wang, Miao Kong, Wen Zhou, Jiangnan Xiao, Xiangjun Xin, et al., « SOA preamplified 100 Gb/s/λ PAM-4 TDM-PON downstream transmission using 10 Gbps O-band transmitters », *in*: *Journal of Lightwave Technology* 38.*2* (2020), Publisher: IEEE, pp. 185–193.

[Zha+14] Xiupu Zhang, Ran Zhu, Dongyan Shen, and Taijun Liu, « Linearization technologies for broadband radio-over-fiber transmission systems », *in*: *Photonics*, vol. 1, Issue: 4, Multidisciplinary Digital Publishing Institute, 2014, pp. 455–472.

[Zha+19] Qingsong Zhao, Shiqi Hao, Yong Wang, Jingnan Wang, Xiongfeng Wan, and Chenlu Xu, « Optimization of pilot pattern in ACO-OFDM systems channel estimation based on compressed sensing », *in*: *17th International Conference on Optical Communications and Networks (ICOCN2018)*, vol. 11048, International Society for Optics and Photonics, 2019, 110481B.

[Zhu+09] Guanghao Zhu, Wei Liu, and Harold R Fetterman, « A broadband linearized coherent analog fiber-optic link employing dual parallel Mach–Zehnder modulators », *in*: *IEEE Photonics Technology Letters* 21.*21* (2009), Publisher: IEEE, pp. 1627–1629.

[Zhu+14a] Ran Zhu and Xiupu Zhang, « Broadband predistortion circuit design for electro-absorption modulator in radio over fiber system », *in*: *Optical Fiber Communication Conference*, Optical Society of America, 2014, W2A–10.

[Zhu+14b] Ran Zhu and Xiupu Zhang, « Linearization of radio-over-fiber systems by using two lasers with different wavelengths », *in*: *2014 IEEE MTT-S International Microwave Symposium (IMS2014)*, IEEE, 2014, pp. 1–3.

[Zhu+13] Ran Zhu, Xiupu Zhang, Bouchaib Hraimel, Dongya Shen, and Taijun Liu, « Broadband predistortion circuit using zero bias diodes for radio over fiber systems », *in*: *IEEE Photonics Technology Letters* 25.*21* (2013), Publisher: IEEE, pp. 2101–2104.

[Zib+08] Darko Zibar, Xianbin Yu, Christophe Peucheret, Palle Jeppesen, and Idelfonso T Monroy, « Digital coherent receiver for phase-modulated radio-over-fiber optical links », *in*: *IEEE Photonics Technology Letters* 21.*3* (2008), Publisher: IEEE, pp. 155–157.

DOCTORAT BRETAGNE LOIRE / MATHSTIC

ÉCOLE NATIONALE D'INGÉNIEURS DE BREST

Titre : Réduction en bande de base des distorsions non-linéaires dans les systèmes OFDM optiques cohérents basés sur les SOAs : analyses stochastiques et expérimentales

Mot clés : Communication optique, formats de modulation multi-porteuses, amplification optique, systèmes non linéaires, techniques de pré-distorsion, macro-modélisation

Résumé : Avec la montée des demandes en débit, les systèmes de communications optiques cherchent à répondre à un certain cahier des charges en termes de débit, de portée de transmission, et de flexibilité. Grâce aux avancées technologiques récentes, la question des possibilités de traitement numérique se pose pour les réseaux d'accès et métropolitains. Ce travail de thèse se penche, d'une part, sur le cas des transmissions utilisant les amplificateurs optiques à semiconducteurs, et d'autre part le cas des modulations multi-porteuses. Ces amplificateurs permettent d'augmenter la bande optique utilisable en limitant le coût de l'architecture. Ce type de modulation, ici la modulation OFDM, donne à l'opérateur une certaine flexibilité tout en augmentant l'efficacité spectrale. La combinaison de ces amplificateurs et de la modulation OFDM engendre des non-linéarités qui affecteront négativement le système de transmission. A travers des analyses numériques, expérimentales, et stochastiques, ce travail de thèse aide à établir la faisabilité et l'apport de la prédistorsion et de la post-distorsion numérique en termes d'amélioration de la qualité globale de transmission, d'augmentation de portée de transmission, et de robustesse à la variabilité du système.

Title: Baseband mitigation of nonlinear impairments in SOA based coherent optical OFDM systems: stochastic and experimental analyses

Keywords: Optical communication, multi-carrier modulation formats, optical amplification, nonlinear systems, predistortion techniques, macromodeling

Abstract: With the demand for higher data rates, optical communication systems are seeking to meet certain specifications in terms of data rate, transmission range, and flexibility. With the recent technological advances, the question of digital signal processing possibilities arises for access and metropolitan networks. This thesis focuses, on the one hand, on transmissions using semiconductor optical amplifiers, and on the other hand, on multi-carrier modulations. These amplifiers make it possible to increase the usable optical bandwidth while limiting the cost of the architecture. This type of modulation, here OFDM modulation, gives the operator a certain flexibility while increasing the spectral efficiency. The combination of these amplifiers and the OFDM modulation generates non-linearities that will negatively affect the transmission system. Through numerical, experimental, and stochastic analyses, this thesis helps to establish the feasibility and advantages of digital predistortion and post-distortion in terms of improving overall transmission quality, increasing transmission range, and robustness to system variability.