Statistical Analysis of Equalization Enhanced Phase Noise in Coherent Fiber Optical Communications

THOMAS GAUDY
ROyal institute of technology (KTH)

abstract

School of Information and Communication Technology
School of Electrical Engineering

Master in Wireless Systems

statistical analysis of equalization enhanced phase noise in coherent fiber optical communications

by Thomas Gaudy

Coherent optical communications increase the receiver sensitivity and the total capacity. Together with high speed digital signal processing (DSP), compensation for linear transmission impairments (chromatic dispersion, polarization mode dispersion and phase noise) can be achieved. In the past, many algorithms have been developed for this purpose: we offer an overview of those. However, in dispersion unmanaged coherent optical links, even after DSP, the received constellation remains influenced by enhanced noise commonly known as equalization enhanced phase noise (EEPN). Its properties are explained in further details using statistical analysis. Statistical properties of EEPN are compared with that of additive white gaussian noise (AWGN) and pure phase noise (PN) by numerical simulations complemented with explanation of mean, variance and autocorrelation function. Finally, the possibility to use only digital post-processing for EEPN (similar to PN) is discussed and shown to be not feasible.
Acknowledgements

I want to deeply thank all the fiber optics team of Acreo Swedish ICT for their precious help: Oskars Ozolins and Gunnar Jacobsen for the advices and the constant supervisation, Aditya Kakkar for sharing his knowledge on EEPN, Xiaodan Pang for sharing his opinion and for the Matlab code, Aleksejs Udalcovs for sharing the VPI license, Richard Schatz for his expertise, Sergei Popov for listening at the meetings and all the others I forgot. I also want to thank my supervisor at KTH Slimane Ben Slimane for giving me the opportunity to do this thesis work and helping me with the formalities and all my teachers at KTH and Supélec for all the knowledge they gave me in my six and a half years of studies.
# Contents

Abstract iii

Acknowledgements v

1 Introduction 1
  1.1 Background ................................................. 1
  1.2 Two Different Paths ........................................ 2
  1.3 Problem .................................................. 2

2 Background Study on Coherent Fiber Optical Communications 5
  2.1 Coherent Optical Communications ............................ 5
  2.2 Impairments and Compensation Methods ....................... 7
    2.2.1 IQ Imbalance ........................................... 8
    2.2.2 Chromatic Dispersion Compensation ..................... 8
      Least Mean Square Adaptive Filter ......................... 9
      Fiber Dispersion Finitie-Impulse Response Filter ........... 10
      Blind Look-Up Filter - Frequency Domain Equalization .... 11
    2.2.3 Polarization Mode Dispersion Compensation ............. 12
      Decision-Directed Least Mean Square Algorithm ............ 12
      Constant Modulus Algorithm ................................. 13
      Multi Modulus Algorithm .................................. 14
    2.2.4 Carrier Phase Recovery ................................ 14
      Maximum A Posteriori ...................................... 15
      One-Tap Least Mean Square Filter ............................ 16
      Phase Lock-Loop Decision Directed .......................... 16
      Phase Lock-Loop Power Law-Average .......................... 17
      Blind Phase Search ....................................... 18
    2.2.5 Equalization Enhanced Phase Noise ...................... 19
      General Model of Coherent Optical System with Electronic
      Dispersion Compensation ...................................... 19
      Analysis of Equalization Enhanced Phase Noise .............. 20
      Statistical Properties of Equalization Enhanced Phase Noise 21
3 Simulations Setup in VPI and Matlab
  3.1 VPItransmissionMaker & VPIcomponentMaker .......................... 26
    3.1.1 Square M-QAM Transmitter ....................................... 26
    3.1.2 Transmission and Reception Filters .............................. 28
    3.1.3 Link .......................................................... 28
    3.1.4 Noise Loading Module ........................................... 29
    3.1.5 Local Oscillator ................................................ 29
    3.1.6 Polarization-Diversity Digital Coherent Receiver ............... 30
    3.1.7 Digital Signal Processing Module ................................ 31
  3.2 Digital Signal Processing in Matlab .................................. 32
    3.2.1 DC Removal .................................................... 32
    3.2.2 IQ Imbalance .................................................. 32
    3.2.3 Chromatic Dispersion Compensation .............................. 33
    3.2.4 Resampling .................................................... 33
    3.2.5 Equalizer ..................................................... 33
    3.2.6 Carrier Phase Recovery ....................................... 33
    3.2.7 Bit Error Rate Counter ...................................... 34

4 Equalization Enhanced Phase Noise Effect on Coherent Fiber Optical Communications
  4.1 Introduction to Numerical Simulations on Equalization Enhanced Pha
    se Noise .............................................................. 35
  4.2 Influence of Equalization Enhanced Phase Noise on the OSNR Penalty ... 36
    4.2.1 Simulation Method of OSNR Penalty ............................ 36
    4.2.2 BER vs OSNR Curves in Presence of Phase Noise ............... 37
    4.2.3 Analytical OSNR Penalty vs Numerically Simulated OSNR Pen
      alty ................................................................. 38
  4.3 Autocorrelation of the Signal Impaired by EEPN ........................ 40
    4.3.1 Simulation Method of Autocorrelation .......................... 40
    4.3.2 Reference Autocorrelation .................................... 43
      Length of the Autocorrelation Function ............................ 44
      Autocorrelation for QPSK ......................................... 45
      Autocorrelation for 16-QAM ....................................... 47
    4.3.3 Simulated Autocorrelation .................................... 49
      Autocorrelation for QPSK ......................................... 49
      Autocorrelation for 16-QAM ....................................... 55
  4.4 Autocorrelation of EEPN Separated from Data .......................... 60
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Coherent Optical Detector</td>
<td>5</td>
</tr>
<tr>
<td>2.2</td>
<td>Digital Signal Processing Chain</td>
<td>8</td>
</tr>
<tr>
<td>2.3</td>
<td>Importance of CD Compensation</td>
<td>9</td>
</tr>
<tr>
<td>2.4</td>
<td>Blind Loop-Up Filter</td>
<td>12</td>
</tr>
<tr>
<td>2.5</td>
<td>Importance of Carrier Phase Recovery</td>
<td>15</td>
</tr>
<tr>
<td>2.6</td>
<td>Phase Lock-Loop Decision Directed</td>
<td>17</td>
</tr>
<tr>
<td>2.7</td>
<td>Global System for Simulations</td>
<td>19</td>
</tr>
<tr>
<td>3.1</td>
<td>Simulation Setup on VPI</td>
<td>25</td>
</tr>
<tr>
<td>3.2</td>
<td>Block Representation of the Square M-QAM Transmitter</td>
<td>26</td>
</tr>
<tr>
<td>3.3</td>
<td>Square M-QAM Transmitter</td>
<td>26</td>
</tr>
<tr>
<td>3.4</td>
<td>Band-Pass Filter</td>
<td>28</td>
</tr>
<tr>
<td>3.5</td>
<td>Optical Fiber Block on VPI</td>
<td>28</td>
</tr>
<tr>
<td>3.6</td>
<td>Noise Loading Module Block</td>
<td>29</td>
</tr>
<tr>
<td>3.7</td>
<td>Noise Loading Module Detail</td>
<td>29</td>
</tr>
<tr>
<td>3.8</td>
<td>Local Oscillator Block</td>
<td>30</td>
</tr>
<tr>
<td>3.9</td>
<td>Local Oscillator Phase Modulation</td>
<td>30</td>
</tr>
<tr>
<td>3.10</td>
<td>Polarization-Diversity Digital Coherent Receiver Block</td>
<td>30</td>
</tr>
<tr>
<td>3.11</td>
<td>Polarization-Diversity Digital Coherent Receiver Block Detail</td>
<td>31</td>
</tr>
<tr>
<td>3.12</td>
<td>Digital Signal Processing Module</td>
<td>32</td>
</tr>
<tr>
<td>3.13</td>
<td>Digital Signal Processing Chain</td>
<td>32</td>
</tr>
<tr>
<td>4.1</td>
<td>OSNR Penalty Calculation</td>
<td>37</td>
</tr>
<tr>
<td>4.2</td>
<td>BER vs OSNR</td>
<td>39</td>
</tr>
<tr>
<td>4.3</td>
<td>OSNR Penalty vs Linewidth of the LO Laser</td>
<td>40</td>
</tr>
<tr>
<td>4.4</td>
<td>Autocorrelation Calculations Placement</td>
<td>42</td>
</tr>
<tr>
<td>4.5</td>
<td>Autocorrelation for different numbers of symbols and QPSK modulation</td>
<td>46</td>
</tr>
<tr>
<td>4.6</td>
<td>Autocorrelation for different numbers of symbols and 16QAM modulation</td>
<td>49</td>
</tr>
<tr>
<td>4.7</td>
<td>Global autocorrelation for different impairments and QPSK modulation</td>
<td>50</td>
</tr>
</tbody>
</table>
4.8 Zoom on the central peak of the autocorrelation for different impairments and QPSK modulation ........................................ 51
4.9 Zoom on the side peak of the autocorrelation for different impairments and QPSK modulation ........................................ 52
4.10 Constellations for different impairments and QPSK modulation . 53
4.11 Global autocorrelation for different impairments and 16QAM modulation ................................................................. 56
4.12 Zoom on the central peak of the autocorrelation for different impairments and 16QAM modulation ................................. 57
4.13 Zoom on the side peak of the autocorrelation for different impairments and 16QAM modulation ........................................ 58
4.14 Constellations for different impairments and 16QAM modulation . 59
4.15 Autocorrelation and Noise Constellation of only PN and EEPN . 60
4.16 Zoom on Autocorrelation of only EEPN ................................. 61
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Classification of the CPR algorithms.</td>
<td>15</td>
</tr>
<tr>
<td>4.1</td>
<td>Parameters used for the simulations to calculate the OSNR penalty.</td>
<td>38</td>
</tr>
<tr>
<td>4.2</td>
<td>Parameters used for the simulations of QPSK to calculate the autocorrelation for three different cases: AWGN, PN and EEPN.</td>
<td>41</td>
</tr>
<tr>
<td>4.3</td>
<td>Parameters used for the simulations of 16-QAM to calculate the autocorrelation for three different cases: AWGN, PN and EEPN.</td>
<td>41</td>
</tr>
<tr>
<td>4.4</td>
<td>Number of samples of the sequence at every stage of the DSP chain for 32 * 2^{N} samples transmitted.</td>
<td>45</td>
</tr>
<tr>
<td>4.5</td>
<td>Bits distribution in the transmitter for QPSK, I stands for In-Phase and Q for Quadrature-Phase.</td>
<td>45</td>
</tr>
<tr>
<td>4.6</td>
<td>Bits distribution in the transmitter for 16-QAM, I stands for In-Phase and Q for Quadrature-Phase. The indexes stand for 1st and 2nd bit on either I or Q (two bits are used on I and two bits on Q).</td>
<td>47</td>
</tr>
</tbody>
</table>
# List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADC</td>
<td>Analog to Digital Converter</td>
</tr>
<tr>
<td>AWGN</td>
<td>Additive White Gaussian Noise</td>
</tr>
<tr>
<td>BLU</td>
<td>Blind Look-Up</td>
</tr>
<tr>
<td>BPS</td>
<td>Blind Phase Search</td>
</tr>
<tr>
<td>CD</td>
<td>Chromatic Dispersion</td>
</tr>
<tr>
<td>CMA</td>
<td>Constant Modulus Algorithm</td>
</tr>
<tr>
<td>COD</td>
<td>Coherent Optical Detection</td>
</tr>
<tr>
<td>CPR</td>
<td>Carrier Phase Recovery</td>
</tr>
<tr>
<td>DAC</td>
<td>Digital to Analog Converter</td>
</tr>
<tr>
<td>DD-LMS</td>
<td>Decision Directed Least Mean Square</td>
</tr>
<tr>
<td>DSP</td>
<td>Digital Signal Processing</td>
</tr>
<tr>
<td>EDC</td>
<td>Electronic Dispersion Compensation</td>
</tr>
<tr>
<td>EDFA</td>
<td>Erbium-Doped Fiber Amplifier</td>
</tr>
<tr>
<td>EEPN</td>
<td>Equalization-Enhanced Phase Noise</td>
</tr>
<tr>
<td>FDE</td>
<td>Frequency Domain Equalization</td>
</tr>
<tr>
<td>FD-FIR</td>
<td>Fiber Dispersion Finite-Impulse Response</td>
</tr>
<tr>
<td>FEC</td>
<td>Forward Error Correction</td>
</tr>
<tr>
<td>FFT</td>
<td>Fast Fourier Transform</td>
</tr>
<tr>
<td>FWHM</td>
<td>Full Width at Half-Maximum</td>
</tr>
<tr>
<td>GVD</td>
<td>Group Velocity Dispersion</td>
</tr>
<tr>
<td>I</td>
<td>In-Phase</td>
</tr>
<tr>
<td>IFFT</td>
<td>Inverse Fast Fourier Transform</td>
</tr>
<tr>
<td>IFT</td>
<td>Inverse Fourier Transform</td>
</tr>
<tr>
<td>LMS</td>
<td>Least Mean Square</td>
</tr>
<tr>
<td>LO</td>
<td>Local Oscillator</td>
</tr>
<tr>
<td>MAP</td>
<td>Maximum A Posteriori</td>
</tr>
<tr>
<td>MMA</td>
<td>Multi Modulus Algorithm</td>
</tr>
<tr>
<td>M-PSK</td>
<td>M-ary Phase Shift Keying</td>
</tr>
<tr>
<td>M-QAM</td>
<td>M-ary Quadrature Amplitude Modulation</td>
</tr>
<tr>
<td>NLMS</td>
<td>Normalized Least Mean Square</td>
</tr>
<tr>
<td>OFDM</td>
<td>Orthogonal Frequency-Division Multiplexing</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------------------</td>
</tr>
<tr>
<td>OSNR</td>
<td>Optical Signal-to-Noise Ratio</td>
</tr>
<tr>
<td>PLL</td>
<td>Phase Lock-Loop</td>
</tr>
<tr>
<td>PMD</td>
<td>Polarization Mode Dispersion</td>
</tr>
<tr>
<td>PN</td>
<td>Phase Noise</td>
</tr>
<tr>
<td>PRBS</td>
<td>Pseudo-Random Binary Sequence</td>
</tr>
<tr>
<td>PSK</td>
<td>Phase Shift Keying</td>
</tr>
<tr>
<td>Q</td>
<td>Quadrature-Phase</td>
</tr>
<tr>
<td>QPSK</td>
<td>Quadrature Phase Shift Keying</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal-to-Noise Ratio</td>
</tr>
<tr>
<td>WGN</td>
<td>White Gaussian Noise</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 Background

The explosive growth of data traffic has transformed the landscape of modern optical telecommunication networks. In order to reach the total system capacity in the Tb/s range, multi-level modulation such as MPSK, MQAM, and optical orthogonal frequency division multiplexing (OOFDM) with MPSK or MQAM modulation on each multiplexed optical frequency are needed. Coherent multilevel modulation schemes are now becoming the next generation systems for high capacity (up to 10 Tb/s) communication on a single transmission wavelength, and 100 Tb/s on a single transmission fiber [1–5]. However, such modulation formats are quite sensitive where even minor imperfections in the coherence performance of transmitter and receiver (local oscillator) laser sources become a strong limiting factor.

New generation coherent systems are unique in the use of high speed digital signal processing (DSP) in transmitters and/or receivers that can be utilized to mitigate the influence of system impairments. The impairments can occur in both optical domain, such as fiber chromatic and polarization dispersion, added spontaneous emission noise from in-line optical amplifiers, fiber nonlinearities; and electric domain, like imperfections in drive electronics and in external modulators (phase and amplitude), phase noise in the transmitter and local oscillator including equalization enhanced phase noise due to the use of DSP, etc.

In fact, coherent detection offers the advantage of access to optical electric field amplitude and phase in the electronic domain at the receiver. This allows fast recovery of linear channel transfer functions. Thus, DSP can be used to eliminate the influence of chromatic and polarization mode dispersion in the optical fiber, to adjust signal polarization imbalance at the receiver, to extract a reference carrier phase [3, 4], and to - at least in principle - eliminate the influence of fiber nonlinearities [5].
The aim of this thesis work is to investigate the phase noise influence for a number of multi-level modulation formats and its impact on coherent optical communications. The presence of phase noise along with the use of electronic dispersion compensation introduces Equalization Enhanced Phase Noise (EEPN) [6]. The objective is to investigate if EEPN can be directly compensated by suitable DSP techniques which are adopted from the equivalent radio frequency counterparts [7, 8]. We investigate how the phase noise statistics are modified by the use of DSP in the radio world and which further considerations are needed to transfer the radio principles into the optical transmission world, where the systems operate at much higher frequencies.

1.2 Two Different Paths

The drastic growth in the diverse IP based applications market will result in an annual traffic of 2 zettabyte by the end of 2019 [9]. Thus, operators are requesting upgrades to coherent transceivers, enabling the use of higher order modulation formats in order to cope with this demand. However, in systems with electronic dispersion compensation, after DSP, the received constellation remains influenced by EEPN [10, 11]. This limits the possibility to achieve either higher capacity and/or longer reach. Scientists around the world are trying to find solutions to mitigate EEPN and two different paths arise: hardware solution or software solution. The idea behind this thesis work is to investigate if the software solution (DSP) is practical. If not, then the hardware solution should be more deeply investigated.

1.3 Problem

The system we consider is using real laser that don’t transmit a unique frequency component but has a finite linewidth (due to phenomenons such as spontaneous emission amongst others). Each laser has a linewidth which is a representation of the spectral width of the laser output. We want to investigate the influence of the frequency noise of the LO laser over the overall transmission system. Thus, we are going to simulate a full transmission system using VPItransmissionMaker and Matlab in cosimulation.

For this thesis work, only white frequency noise at the LO laser is considered (which means that the frequency noise of the laser is white, such a laser has a
Lorentzian lineshape [12, 13]), the transmitter laser is considered ideal (transmitting only at its central frequency). This frequency noise will introduce two important impairments: Phase Noise (PN) and Equalization Enhanced Phase Noise (EEPN). Those impairment are introduced later and the methods to deal with them are explained.

Previous studies [14] show that EEPN can’t be treated by linear filtering. We want to confirm this statement, investigate if EEPN can be treated with DSP post-processing and understand the mitigation possibilities.
Chapter 2

Background Study on Coherent Fiber Optical Communications

In this Chapter, the different impairments that can be encountered when dealing with optical communications and the solutions and algorithms used to deal with them are introduced. If you need more details about each algorithm, please refer to the references given at the end of the report. This Chapter is a state of art and presents results from previous studies, we will not discuss more about each algorithm than what is given here.

2.1 Coherent Optical Communications

For this Master Thesis work, coherent optical detection (COD) is employed which enables us to fully recover the amplitude and the phase of the signal at the receiver, digital signal processing (DSP) is used to treat the signal post-reception [15]. The definition of coherent optical detection [4, 16, 17] is reminded. First, the diagram of a typical set for COD with DSP is given on Figure 2.1.

![Diagram of a coherent optical detector](image)

**Figure 2.1:** A coherent optical detector composed of: a local oscillator, a 90° hybrid, two photo-diodes, two A/D converters and a DSP block.
This configuration enables us to fully recover the information about the phase and the amplitude of the signal, which gives us two degrees of freedom. The incident powers are given by Equations 2.1a, 2.1b, 2.1c and 2.1d.

\[
P_{1a}(t) = |E_s(t)|^2 + |E_{LO}|^2 + 2\text{Re}[E_s(t)E_{LO}^*e^{j(\omega_S - \omega_{LO})t + j\phi_s(t) - j\phi_{LO}(t)}] \\
\text{(2.1a)}
\]

\[
P_{1b}(t) = |E_s(t)|^2 + |E_{LO}|^2 - 2\text{Re}[E_s(t)E_{LO}^*e^{j(\omega_S - \omega_{LO})t + j\phi_s(t) - j\phi_{LO}(t)}] \\
\text{(2.1b)}
\]

\[
P_{2a}(t) = |E_s(t)|^2 + |E_{LO}|^2 + 2\text{Im}[E_s(t)E_{LO}^*e^{j(\omega_S - \omega_{LO})t + j\phi_s(t) - j\phi_{LO}(t)}] \\
\text{(2.1c)}
\]

\[
P_{2b}(t) = |E_s(t)|^2 + |E_{LO}|^2 - 2\text{Im}[E_s(t)E_{LO}^*e^{j(\omega_S - \omega_{LO})t + j\phi_s(t) - j\phi_{LO}(t)}] \\
\text{(2.1d)}
\]

where \(E_s(t)\), \(\omega_s\) and \(\phi_s(t)\) are respectively the complex envelope, the angular frequency and the phase of the received optical signal; \(E_{LO}, \omega_{LO}\) and \(\phi_{LO}(t)\) are respectively the complex envelope, the angular frequency and the phase of the local oscillator (LO). \(\omega_s\) and \(\omega_{LO}\) are considered constant since they are drifting slowly.

The balanced photo-detector outputs are given by Equations 2.2a and 2.2b.

\[
\Delta P_1 = P_{1a}(t) - P_{1b}(t) = 4\text{Re}[E_s(t)E_{LO}^*e^{j(\omega_S - \omega_{LO})t + j\phi_s(t) - j\phi_{LO}(t)}] \\
\text{(2.2a)}
\]

\[
\Delta P_2 = P_{2a}(t) - P_{2b}(t) = 4\text{Im}[E_s(t)E_{LO}^*e^{j(\omega_S - \omega_{LO})t + j\phi_s(t) - j\phi_{LO}(t)}] \\
\text{(2.2b)}
\]

Hence, the electric field envelope of the signal is found by applying the mathematical relationship given in Equation 2.3.

\[
E_s(t) = \frac{e^{j(\omega_S - \omega_{LO})t + j\phi_s(t) - j\phi_{LO}(t)}}{4E_{LO}} (\Delta P_1(t) + j\Delta P_2(t)) \\
\text{(2.3)}
\]

But to achieve this, the frequency difference and phase difference from \(\Delta P_1\) and \(\Delta P_2\) must be retrieved which is further discussed in the part on the phase recovery.

In the system, the optical signal electric field envelope arriving at the receiver is the sum of the modulated transmitter laser and additive noise. Both shapes are chosen so that there is no inter-symbol interference and the received field is given
by Equation 2.4.

\[ E_s(t) = d(n)E_{s0}\exp(j\omega_s t + j\phi_s(n)) + p(n)E_{s0}\exp(j\omega_{LO} t + j\phi_{LO}(t)) \]  

(2.4)

where \( d(n) \) represents the digital values of the encoded bits (dependent of the type of constellation we use), \( E_{s0} \) is a constant and the second term is a Gaussian additive noise. The phase diverse coherent receiver observes the quantity \( r(n) \) given by Equation 2.5.

\[ r(n) = \frac{\Delta P_1(n) + j\Delta P_2(n)}{4E_{s0}E_{LO}^*} = d(n)\exp(j\omega t + j\phi(n)) + p(n) \]  

(2.5)

where \( \omega = \omega_s - \omega_{LO} \) and \( \phi(n) = \phi_s(n) - \phi_{LO}(n) \). \( p(n) \) is a complex Gaussian noise sequence where each complex part has variance \( \sigma_p^2 \).

And when the local oscillator laser has a Lorentzian line-shape given by Equation 2.7, the phase noise is a Wiener process [18] described by Equation 2.6.

\[ \phi(n) = \phi(n-1) + \omega(n) \]  

(2.6)

where \( \omega(n) \) is a white Gaussian noise [19] sequence of variance \( \sigma_\omega^2 \). \( r(n) \) is the signal we are interested in and that we are going to use to apply all the DSP algorithms.

\[ L(v) = \frac{\Delta v_L}{2\pi} \frac{1}{(v - v_0)^2 + \frac{\Delta v_L^2}{4}} \]  

(2.7)

where \( \Delta v_L \) is the Full Width at Half-Maximum (FWHM) of the Lorentzian line-shape.

## 2.2 Impairments and Compensation Methods

Even though all the parameters are retrieved, there are some impairments that need to be solved and there is a numerous amount of methods used to counter-effect the impairments in the fiber optic [20]. In this state of art, the following impairments and solutions [21] are presented:

- IQ Imbalance
- Chromatic Dispersion (CD)
- Polarization Mode Dispersion (PMD)
Chapter 2. Background Study on Coherent Fiber Optical Communications

- Phase Noise (PN)

The latter is the ultimate goal of this thesis work and will be subject to further studies. But, in order to understand the system globally, all the different impairments considered must be explained. The typical setup for DSP is given on Figure 2.2

![Block representation of the DSP algorithms.](image)

To illustrate the influence of each impairment on the system, the constellation for the QPSK modulation format [22] is given.

### 2.2.1 IQ Imbalance

The signal, as it arrives after the receiver, can be divided in two part: in-phase (I) and quadrature-phase (Q). It happens, in some part of the system, that these two components I and Q are separated and use different paths (in the 90-hybrid for example). These separations of the signal in two can lead to different losses in the I branch and the Q branch. These different losses will affect the average power in both branches with a different impact [23]. Thus, the received constellation could lose its square shape (and have a parallelogram shape). In order to preserve the efficiency of the DSP algorithms, the square shape of the constellation needs to be brought back. This is the role of the IQ imbalance module. It measures the average power on both I and Q of the received signal. Then, it divides each signal samples on both I and Q by the square root of the measured average power. In this way, the constellation is forced to have a square form and the DSP algorithms are ensured to work properly.

### 2.2.2 Chromatic Dispersion Compensation

Chromatic dispersion [24] is a phenomenon that is important in fiber optic communications. It appears as the wavelength-dependent dispersion of a light beam in time. For two different wavelengths, the travelling time on a given distance is different. Thus, a sharp light pulse travelling through the fiber will tend to get flattened or stretched. Without any compensation for this impairment, the
maximum capacity of the fiber can’t be reached. A non-exhaustive list of different methods employed to counter-effect the chromatic dispersion is put here.

On Figure 2.3 (to the left), the constellation for a link of 2000 km in presence of CD is given, with a CD coefficient of 16 ps/(nm.km) and an OSNR of 30 dB. It is obvious that none of the information provided by this constellation can be used if the CD is not compensated for. On the same figure (to the right), the resulting constellation after CD compensation using the Blind Look-Up Filter - Frequency Domain Equalization method [3] is shown. The CD compensated constellation is error-free in this case. This shows the importance of CD compensation.

**Least Mean Square Adaptive Filter**

The least mean square algorithm (LMS) [3] is a particular case of the Kalman filter [25] which costs less in computation. It is suitable to use on a system that presents impairments that are varying in time. The CD is directly dependent on the channel since it’s the one creating it. Here, if the fiber is bended or if the temperature varies around the fiber, it can change the physical properties of the link and the CD will differ from before the perturbation. Therefore, an adaptive filter is used and will be able to compensate for the perturbation. LMS is suitable since it is fast and it presents good performances.

The basic of the LMS algorithm is that it relies on the error between what we want and what we obtained from the previous iteration of the algorithm. The
principle of LMS filter is given by the Equations 2.8a, 2.8b and 2.8c.

\[ y(n) = h^H(n)x(n) \]  \hspace{1cm} (2.8a)

\[ h(n + 1) = h(n) + \mu x(n)e^*(n) \]  \hspace{1cm} (2.8b)

\[ e(n) = d(n) - y(n) \]  \hspace{1cm} (2.8c)

where \( x(n) \) is the complex magnitude vector of the received signal, \( h(n) \) is the complex tap weights vector (the power \( H \) represents the Hermitian transform of the given vector \([26]\)), \( y(n) \) is the complex magnitude of the output (previously equalized received signal), \( d(n) \) is the desired symbol, which correspond to one of the different possible cases (depending on the modulation format used, it could be one of the following symbols for the QPSK constellation: \([1+i, 1-i, -1-i, -1+i]\)), \( e(n) \) represents the estimation error between the output signal and the desired symbol (the power \( * \) represents the conjugation of the given variable), and \( \mu \) is a coefficient called the step-size.

The goal of the algorithm is to reach a situation where \( e(n) \) is as small as possible (the ideal would be to have it to 0 all the time, which is of course not reachable). The tap-weights are updated depending on the value of the error.

The step-size parameter can not be set-up randomly and has to respect some contraints. If it is too big, the algorithm can diverge and become unstable. Thus, an interval of stability for the step-size is defined as follows: \( 0 < \mu < 1/\lambda_{\text{max}} \) where \( \lambda_{\text{max}} \) is the largest eigenvalue of the correlation matrix \( R = x(n)x^H(n) \). If the step-size is chosen very small, the algorithm will converge slowly but it will be stable, if the step-size is chosen large, the algorithm will converge faster but it will be less stable. Thus, a compromise between convergence speed and stability is to be found. A way to do this is to update the step-size variable depending on the largest eigenvalue, which is referred as variable-step-size LMS algorithm. But this method requires to check the largest eigenvalue at every step, so it reduces the computational efficiency of the classic LMS algorithm. Therefore, a small enough value of the step-size is picked to ensure the stability of the algorithm.

**Fiber Dispersion Finite-Impulse Response Filter**

Another solution is to implement a fiber dispersion finite-impulse response (FD-FIR) \([3]\) filter. Since the physics of the optical fiber are known, the equations that command the impairment can be written. A complementing filter can be implemented in order to compensate for the CD effect. The specification is quite
simple and the filter taps are given by Equation 2.9.

$$a_k = \sqrt{\frac{j c T^2}{D \lambda^2 z}} \exp \left( -j \frac{\pi c T^2}{D \lambda^2 z} k^2 \right) \text{ with } -\left\lfloor \frac{N}{2} \right\rfloor \leq k \leq \left\lfloor \frac{N}{2} \right\rfloor$$ (2.9)

where \( D \) is the fiber chromatic dispersion coefficient, \( \lambda \) is the central wavelength of the transmitted optical signal, \( z \) is the fiber length and \( T \) is the sampling period.

In order to avoid aliasing, the cutting frequency of the frequency spreading of the pass-band filter needs to be lower than the Nyquist frequency. The filter needs to be windowed to limit the number of taps. This constraint is given by Equation 2.10.

$$-\frac{|D|\lambda^2 z}{2cT} \leq t \leq \frac{|D|\lambda^2 z}{2cT}$$ (2.10)

The time window of the filter is then given by Equation 2.11.

$$T_W^A = \frac{|D|\lambda^2 z}{cT}$$ (2.11)

And so, the maximum number of taps for the filter is given by Equation 2.12.

$$N^A = 2 \left\lfloor \frac{T_W^A}{2T} \right\rfloor + 1 = 2 \left\lfloor \frac{|D|\lambda^2 z}{2cT^2} \right\rfloor + 1$$ (2.12)

**Blind Look-Up Filter - Frequency Domain Equalization**

The idea under the blind look-up filter (BLU) [3] is similar to the idea of the FD-FIR filter. Since the physics of the optical fiber are known, a complementing filter can be implemented. But this time, the filtering operates in frequency domain (instead of time domain for the FD-FIR filter). The Fourier transform [19] is applied to the received signal, the resulting FT is multiplied by the inverse transfer function of the fiber and finally the inverse Fourier transform is applied. The multiplication by the inverse transfer function of the fiber is called frequency domain equalization (FDE). To get the algorithm to run fast, the fast Fourier transform (FFT) and inverse fast Fourier transform (IFFT) are used. And to properly reconstruct the signal, it is divided in blocks of convenient size for the FFT algorithm; the FFT, the multiplication by the inverse transfer function and the IFFT are applied, and the overlap-save or the overlap-add reconstructs the signal. The system is presented on Figure 2.4.

Here the function \( G(z, \omega) = \exp(-jD \frac{\lambda^2}{2zc} \frac{\omega^2}{z} z) \) is the inverse transfer function of the fiber span.

The number of filter taps and the overlap ratio have to be chosen. The most convenient way is to fix a 50% overlap and then pick a number of taps that
will ensure the best compensation of the CD. The Nyquist frequency has to be respected by the constraint given by Equation 2.13.

$$-\omega_N \leq \omega_{BLU} \leq \omega_N$$  \hspace{1cm} (2.13)

where $\omega_N$ is the Nyquist angular frequency of the system. Then, the tap weights of the BLU filter are given by Equation 2.14.

$$b_k = \exp \left[ -j \frac{D\lambda^2 z}{\pi c} \left( \frac{k\omega_N}{N_{FFT}} \right)^2 \right] \text{ with } -\frac{N_{FFT}}{2} \leq k \leq \frac{N_{FFT}}{2}$$  \hspace{1cm} (2.14)

where $N_{FFT}$ is the size of the FFT.

### 2.2.3 Polarization Mode Dispersion Compensation

Polarization mode dispersion (PMD) [27] is another form of impairment in fiber optic communications. This is due to random imperfections and asymmetries of the fiber. The effect is that the two polarizations of light could travel at different speed in the fiber and it could introduce a spreading of optical pulses. A list of the main methods that have been used to counter-effect this dispersion is given here.

#### Decision-Directed Least Mean Square Algorithm

The influence of PMD and the polarization fluctuation can be compensated by using the adaptive filter decision-directed least mean square algorithm (DD-LMS) [3]. This algorithm is similar to the previously shown LMS algorithm, but it consists of applying 4 times the LMS algorithm with 4 equations: 2 equations for
inter-polarization compensation and 2 equations for intra-polarization compensation. The algorithm is designed with the Equations 2.15a, 2.15b and 2.15c.

\[
\begin{pmatrix}
    x_{\text{out}}(n) \\
    y_{\text{out}}(n)
\end{pmatrix}
= 
\begin{pmatrix}
    h_{xx}^H(n) & h_{xy}^H(n) \\
    h_{yx}^H(n) & h_{yy}^H(n)
\end{pmatrix}
\begin{pmatrix}
    x_{\text{in}}(n) \\
    y_{\text{in}}(n)
\end{pmatrix}
\]  

(2.15a)

\[
\begin{align*}
    h_{xx}(n+1) &= h_{xx}(n) + \mu p \epsilon_x(n) x_{\text{in}}^*(n) \\
    h_{yx}(n+1) &= h_{yx}(n) + \mu p \epsilon_y(n) y_{\text{in}}^*(n) \\
    h_{xy}(n+1) &= h_{xy}(n) + \mu p \epsilon_x(n) y_{\text{in}}^*(n) \\
    h_{yy}(n+1) &= h_{yy}(n) + \mu p \epsilon_y(n) y_{\text{in}}^*(n)
\end{align*}
\]  

(2.15b)

\[
\begin{align*}
    \epsilon_x(n) &= d_x(n) - x_{\text{out}}(n) \\
    \epsilon_y(n) &= d_y(n) - y_{\text{out}}(n)
\end{align*}
\]  

(2.15c)

where \(x_{\text{in}}(n)\) and \(y_{\text{in}}(n)\) are the complex magnitude vectors of the input signals on X-polarization and Y-polarization respectively and \(x_{\text{out}}(n)\) and \(y_{\text{out}}(n)\) are the complex magnitudes of the equalized output signals. The taps weights are represented by the complex vectors \(h_{xx}(n), h_{xy}(n), h_{yx}(n)\) and \(h_{yy}(n)\). The desired symbols are represented by \(d_x(n)\) and \(d_y(n)\); and the associated estimation errors are given by \(\epsilon_x(n)\) and \(\epsilon_y(n)\). And the parameter \(\mu p\) represents the step-size of the LMS algorithm. The discussion to choose the step-size parameter is the same as for the regular LMS algorithm.

**Constant Modulus Algorithm**

The constant modulus algorithm (CMA) [28] was proposed by Godard and is part of the dynamic-channel equalization theory. It can only be applied to transmission systems which don’t use amplitude modulation (QPSK or PSK for example). In this algorithm, the symbols are considered to have a constant modulus (e.g. 1 which could correspond to a QPSK modulation format). The errors to be minimized are \(\epsilon_x^2(n) = (1 - |x_{\text{out}}(n)|^2)^2\) and \(\epsilon_y^2(n) = (1 - |y_{\text{out}}(n)|^2)^2\), where \(x_{\text{out}}(n)\) and \(y_{\text{out}}(n)\) are the outputs of the algorithm on X-polarization and Y-polarization respectively. This specification gives an algorithm close to the LMS algorithm (it’s also an adaptive filter) but slightly different, the algorithm is designed by
Equations 2.16a and 2.16b.

\[
\begin{pmatrix}
x_{\text{out}}(n) \\
y_{\text{out}}(n)
\end{pmatrix} =
\begin{pmatrix}
h_{xx}^* & h_{xy}^* \\
h_{yx}^* & h_{yy}^*
\end{pmatrix}
\begin{pmatrix}
x_{\text{in}}(n) \\
y_{\text{in}}(n)
\end{pmatrix}
\]

\[ (2.16a) \]

\[
\begin{align*}
h_{xx}(n+1) &= h_{xx}(n) + \mu \epsilon_x(n) x^*_{\text{in}}(n)x_{\text{out}}(n) \\
h_{yx}(n+1) &= h_{yx}(n) + \mu \epsilon_y(n) y^*_{\text{in}}(n)x_{\text{out}}(n) \\
h_{xy}(n+1) &= h_{xy}(n) + \mu \epsilon_x(n) x^*_{\text{in}}(n)y_{\text{out}}(n) \\
h_{yy}(n+1) &= h_{yy}(n) + \mu \epsilon_y(n) y^*_{\text{in}}(n)y_{\text{out}}(n)
\end{align*}
\]

\[ (2.16b) \]

where \( x_{\text{in}}(n) \) and \( y_{\text{in}}(n) \) are the complex magnitude vectors of the input signals on X-polarization and Y-polarization respectively. The taps weights are represented by the complex vectors \( h_{xx}(n), h_{xy}(n), h_{yx}(n) \) and \( h_{yy}(n) \). The parameter \( \mu \) represents the step-size of the algorithm.

**Multi Modulus Algorithm**

The principle of the multi-modulus algorithm (MMA) [28] is the same as the CMA, but it is extended so that higher modulation formats such as 16-QAM can be considered. The difference is in the calculation of the errors, which are given by Equation 2.17. This time, the transmission system is allowed to use amplitude modulation by introducing different possible amplitudes of the symbols in the constellation.

\[
\begin{align*}
\epsilon_x^2(n) &= (R^2(i) - |x_{\text{out}}(n)|^2)^2 \\
\epsilon_y^2(n) &= (R^2(i) - |y_{\text{out}}(n)|^2)^2
\end{align*}
\]

\[ (2.17) \]

where \( R(i) \) lists the possible radius in the constellation and \( i \) is the index of the closest radius to the one of the received signal.

**2.2.4 Carrier Phase Recovery**

It is necessary to have a Carrier Phase Recovery [29] algorithm in the DSP because it happens as well that the phase of the information bits is being shifted randomly, leading in a random rotation of the constellation. This phenomenon is called phase-noise and is inducted by the frequency noise of the lasers. The shape of the real laser emission is broadened around the central frequency which doesn’t lead in a perfect demodulation. The phase noise can be modeled in different ways, the most studied one is the Lorentzian line-shape of the laser emission. The following presented algorithms have been studied for the Lorentzian line-shape.
Figure 2.5: To the left: the constellation without carrier phase recovery; to the right: the constellation with carrier phase recovery.

Figure 2.5 (to the left) shows the effect of white frequency noise on the QPSK constellation for a laser linewidth of 5MHz and an OSNR of 30dB in B2B propagation. The constellation symbols have been rotated in a random way so that rings appear. On the same figure (to the right), the constellation after using the carrier-phase recovery algorithm is shown (the Blind Phase Search algorithm [30] was used).

Two types of CPR algorithms are distinguished: feedforward and feedback. Table 2.1 gives the classification.

Table 2.1: Classification of the CPR algorithms.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Feedback</th>
<th>Feedforward</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum A Posteriori</td>
<td>X</td>
<td>-</td>
</tr>
<tr>
<td>One-tap Least Mean Square</td>
<td>X</td>
<td>-</td>
</tr>
<tr>
<td>Phase Lock-Loop Decision Directed</td>
<td>X</td>
<td>-</td>
</tr>
<tr>
<td>Phase Lock-Loop Power Law Average</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Blind Phase Search</td>
<td>-</td>
<td>X</td>
</tr>
</tbody>
</table>

**Maximum A Posteriori**

The maximum a posteriori (MAP) [4] is the best possible estimate that can be made given the observed values \( r(n) \). It relies on the statistical behavior of the additive noise and phase noise. It tries to find the sequence of values of \( \hat{\phi}(n) \) that
maximizes the log-probability given by Equation 2.18.

\[
\log(P) = \sum_n \left( -\frac{|r(n) - \hat{d}(n)\exp(j\hat{\phi}(n))|^2}{2\sigma_P^2} - \frac{(\hat{\phi}(n) - \hat{\phi}(n-1))^2}{2\sigma_\omega^2} \right)
\] (2.18)

This estimator is way too complex for real-time implementation but can be used for offline simulations and comparison to results from other techniques. It is set to be the reference for the carrier phase recovery algorithms.

**One-Tap Least Mean Square Filter**

A one-tap normalized LMS (NLMS) [3] algorithm can be used to estimate the phase. The output of the filter is given by Equation 2.19.

\[
y_{PN}(n) = w_{NLMS}(n)x_{PN}(n)
\] (2.19)

The error is given by Equation 2.20.

\[
\epsilon_{NLMS}(n) = d_{PN}(n) - y_{PN}(n)
\] (2.20)

The update rule for the filter tap weight is given by Equation 2.21.

\[
w_{NLMS}(n + 1) = w_{NLMS}(n) + \frac{\mu_{PN}}{|x_{PN}(n)|^2}x_{PN}^*(n)\epsilon_{NLMS}(n)
\] (2.21)

where \(x_{PN}(n)\) and \(y_{PN}(n)\) are respectively the input and the output of the NLMS filter, \(w_{NLMS}(n)\) is the filter tap weight, \(\epsilon_{NLMS}(n)\) is the error to the desired symbol \(d_{PN}(n)\) and \(\mu_{PN}\) is the step-size.

**Phase Lock-Loop Decision Directed**

The phase lock-loop decision directed [4] approach is illustrated by Figure 2.6. It can be considered not as efficient as expected since a delay is introduced in the transmission. It will not be used for small transmission lines where the delay is important, but it can be used for long-haul transmission.

The main idea is to buffer the input signal on a length of \(L\) symbols and to multiply the \(L^{th}\) symbol with the corresponding conjugate of the estimate as specified in Equation 2.22.

\[
\hat{d}^*(n-L)r(n-L) = \hat{d}^*(n-L)d(n-L)\exp(j\hat{\phi}(n-L)) + \hat{d}^*(n-L)p(n-L)
\] (2.22)
With the approximation of a low bit error rate, most of the time we have
\( \hat{d}(n - L) = d(n - L) \) which transforms Equation 2.22 to Equation 2.23.

\[
\hat{d}^*(n - L)r(n - L) = \exp(j\phi(n - L)) + p_1(n - L) \tag{2.23}
\]

By smoothing, an old phase estimate \( \hat{\phi}(n - L) \) can be retrieved and applied to
the newly observed signal. The new estimate is specified in Equation 2.24.

\[
\hat{d}(n) = \Gamma \left( r(n)\exp(-j\hat{\phi}(n - L)) \right) \tag{2.24}
\]

where \( \Gamma(x) \) is the decision function (dependent of the modulation format used).
This method is relevant since the phase noise doesn’t significantly drift during \( L \)
symbols.

**Phase Lock-Loop Power Law-Average**

The phase lock-loop power law-average [4] is an alternative way to retrieve
the phase of the signal by raising the M-ary PSK signal to the power M. This implies
to use PSK modulation but 16-QAM can be used and be divided in four underlying
4-PSK constellations. The mathematical model is given by Equations 2.25a and
2.25b.

\[
s(n) = r^M(n) = d^M(n)\exp(jM\phi(n)) + d^{M-1}(n)\exp(j(M - 1)\phi(n))Mp(n) + O[p^2(n)] \tag{2.25a}
\]

\[
s(n) = \exp(jM\phi(n)) + q(n) \tag{2.25b}
\]
From unwrapping $s(n)$, the phase of the signal is obtained. The obtained phase is not necessarily between $-\pi$ and $\pi$, so one needs to be careful. The phase has to be incremented or decremented by $2\pi$ every time the signal crosses the negative real axis. The angle given by Equation 2.26.

$$\theta(n) = M\phi(n) + q_1(n)$$  \hspace{1cm} (2.26)

Where $q_1(n)$ is a real Gaussian noise sequence of variance $\sigma_q^2$. Then to retrieve the phase $\phi(n)$ from $\theta(n)$, a smoothing, averaging filter, or Wiener filter is applied. Here are different options: a zero-lag filter given by Equation 2.27a or a finite-lag filter given by equation 2.27b.

$$\hat{\phi}(z) = \frac{1 - \alpha}{M(1 - \alpha z^{-1})} \theta(z)$$  \hspace{1cm} (2.27a)

$$\hat{\phi}(z) = \frac{1}{M} \left(1 - \alpha\right)^{D} + (1 - \alpha)^2 \sum_{k=1}^{D} \alpha^{D-k} z^{-k}$$

$$\frac{1 - \alpha z^{-1}}{1 - \alpha z^{-1}} z^{D} \theta(z)$$  \hspace{1cm} (2.27b)

**Blind Phase Search**

The blind phase search algorithm (BPS) [30] is the one to be used for the simulations. It consists of dividing the complex plane in many different angular parts and finding which is the closest to the received signal. One idea, which is hidden behind this, is that the phase-noise varies slower than the signal baud rate, so that previously received symbols can be used to estimate the phase noise of the received signal by using a sliding window. With this sliding window, a number $N$ of received symbols are averaged and the BPS algorithm is applied over it. This averaging helps to mitigate the additive white Gaussian noise (AWGN) that pollutes the received signal and thus to have a more accurate estimation of the phase noise.

This algorithm applies for square M-QAM constellation. First, the input symbol-rate samples are rotated by test phase angles $\phi_b$ which gives Equation 2.28.

$$z(n, b) = r(n) \exp(j\phi_b) \text{ with } \left[ \phi_b = b \frac{\pi}{B/2} - \frac{\pi}{4}, b = 1, 2, ..., B \right]$$  \hspace{1cm} (2.28)

Where $B$ is the number of test phase angles used (depends on the wanted precision of the algorithm, and note that increasing $B$ will increase significantly the computational cost of the algorithm). The previous equation is applied to all the symbols in the sliding window. Then the closest points of the constellation to the rotated symbols is found and the angle shift $\phi_b$ for every symbol is stored.
Given this, the angle shifts are averaged over the samples in the window and an accurate estimate of the phase noise for the first symbol in the window is found. The window is finally shifted of one symbol and the previous steps are applied again.

### 2.2.5 Equalization Enhanced Phase Noise

EEPN is another impairment that appears in the system. It is the focus point of this thesis.

#### General Model of Coherent Optical System with Electronic Dispersion Compensation

We want to understand why EEPN appears in coherent optical systems with Electronic Dispersion Compensation (EDC). The mathematical analysis of this system is performed as explained in [14]. The base band equivalent representation of the system is used as shown on Figure 2.7.

![Figure 2.7: System Model of Coherent Optical System with Electronic Dispersion Compensation. The Fourier transform pairs of the signals and the response of components are indicated.](image)

In coherent digital communication, information is encoded in phase and/or amplitude respectfully to the modulation format. A power normalized representation of the system is considered, so that the gain is unity and independent on the modulation. The incoming bits are mapped into complex symbols \( c_n \) at a rate \( R_b/m \), where \( m \) is the number of bits per symbol and \( R_b \) is the bit rate (corresponding to a symbol period \( T_s = m/R_b \)). The symbol sequence is convoluted with a pulse shaping filter represented by the Fourier transform pair \( h_{ps}(t) \) to generate a band limited signal. This signal is modulated by a transmission laser represented by \( e^{j\phi_{Tx}(t)}X_{Tx}(f) \). The modulated signal can be represented as the Fourier transform pair \( r(t)|R(f) \). The signal is then transmitted through a dispersive fiber with response \( h_f(t)|e^{jkf^2} \). The receiver part coherently detects the signal with a LO with base band representation \( e^{j\phi_{LO}(0)}X_{LO}(f) \). The detected signal is
oversampled, followed by a dispersion equalization, downsampled and filtered with a low pass matched filter. Dispersion equalization is represented by the inverse of the channel transfer function \( h_f^{-1} e^{-jkf^2} \) and the signal after it is represented by the Fourier transform pair \( r'(t)|R'(f) \). Finally, the signal goes through carrier phase recovery and demodulation. The carrier phase recovery is considered ideal in the mathematical study.

Since the convolution operation is associative with itself, the operations can be interchanged such as oversampling and downsampling to merge them into a single downsampling operation. Which leads us to represent the system as on Figure 2.7.

Analysis of Equalization Enhanced Phase Noise

Figure 2.7 is the base to calculate the received signal after EDC given by Equation 2.29.

\[
R'(f) = \left[ R(f).e^{jkf^2} \otimes X_{LO}(f) \right].e^{-jkf^2} \quad (2.29)
\]

where \( k = 2\pi^2\beta_2 l = \pi Dlcf_0^{-2} \) is the accumulated dispersion factor, where \( \beta_2 \) is the group velocity dispersion parameter, \( l \) is the length of the fiber, \( D \) is the dispersion coefficient, \( c \) is the speed of light and \( f_0 \) is the central optical frequency. The convolution term is calculated, which leads to Equation 2.30.

\[
R'(f) = e^{-jkf^2}. \int_{-\infty}^{+\infty} R(f - f_1).e^{jk(f-f_1)^2} X_{LO}(f_1) df_1
\]

\[
R'(f) = \int_{-\infty}^{+\infty} R(f - f_1).e^{jk(f^2 - 2ff_1)} X_{LO}(f_1) df_1
\]

\[
R'(f) = \int_{-\infty}^{+\infty} R(f - f_1).e^{jk(f_1 - 2ff_1)} X_{LO}(f_1) df_1
\]

This shows us that the accumulated channel dispersion has been fully compensated by the dispersion equalizer. However, the sidebands of the dispersed signal and LO output are intermixing, resulting in an enhancement of the noise [6]. To understand the impact on the time domain signal, the inverse Fourier transform (IFT) is taken as shown on Equation 2.31. Since the integration is a linear operation, we can swap the order of the terms in the equation to obtain the second formula.

\[
r'(t) = \int_{-\infty}^{+\infty} R(f - f_1).e^{jk(f_1 - 2ff_1)} X_{LO}(f_1) df_1 .e^{j2\pi ft} df
\]

\[
r'(t) = \int_{-\infty}^{+\infty} X_{LO}(f_1).e^{jkf_1^2} \left[ \int_{-\infty}^{+\infty} R(f - f_1).e^{-j2f_1f} .e^{j2\pi ft} df \right] df_1 \quad (2.31)
\]
A multiplication in the frequency domain corresponds to a convolution in the
time domain for the terms in the second integration. We obtain Equation 2.32.

\[ r'(t) = \int_{-\infty}^{+\infty} X_{LO}(f_1).e^{jkf_1^2} \left[ IFT(R(f - f_1)) \otimes IFT(e^{-jkf_1^2}) \right] df_1 \quad (2.32) \]

We use the properties of the Fourier transform to derive

\[ IFT(R(f - f_1)) = r(t).e^{j2\pi f_1 t} \]
\[ IFT(e^{-jkf_1^2}) = \delta(t - \frac{k f_1}{\pi}) \]

And we know that a convolution by a Dirac leads in a delay, then

\[ IFT(R(f - f_1)) \otimes IFT(e^{-jkf_1^2}) = r(t - \frac{k f_1}{\pi}).e^{j2\pi f_1 t}.e^{-jkf_1^2} \]

We plug this expression in Equation 2.32 to obtain Equation 2.33.

\[ r'(t) = \int_{-\infty}^{+\infty} X_{LO}(f_1).e^{-jkf_1^2} \left[ r(t - \frac{k f_1}{\pi}).e^{j2\pi f_1 t} \right] df_1 \quad (2.33) \]

Equation 2.33 shows that each side band component of the LO introduces a
delayed version of the transmitted signal \( r(t) \). The delay is proportional to the
frequency component relatively to the central frequency. EEPN is caused by the
interference of these multiple delayed versions of the signal with itself. Thus, EEPN
corresponds to intra and inter symbol interference [31].

From Equation 2.30, one can see that the noise spectrum is a broadened version
of the signal spectrum. This broadening is governed by the linewidth of the LO
which is negligible compared to the baudrate, thus the noise spectrum should be
almost identical to the signal spectrum. But linear filtering cannot remove noise
with same spectral characteristic as the signal itself. This is the reason why EEPN
cannot be mitigated with any linear filter [32].

**Statistical Properties of Equalization Enhanced Phase Noise**

In the previous section, it is shown that EEPN was caused by intra and inter
symbol interference and that it could not be treated by linear filtering. We want
to verify this by a deeper statistical study of EEPN. The autocorrelation function
of the signal at different points in the DSP chain is computed by numerical
simulations. Therefore, some statistical properties of EEPN are given, based on
[14].

The mean and the variance are to be derivated. Based on Equation 2.33, the
mean of the signal is given by Equation 2.34 (the expectation and the intergration
operations are linear).

\[ E[r'(t)] = \int_{-\infty}^{+\infty} E \left[ X_{LO}(f_1).r(t - \frac{k f_1}{\pi}) \right].e^{-jkf_1^2}.e^{j2\pi f_1 t} df_1 \quad (2.34) \]
where $E[.]$ is the expectation operation. The two stochastic signals $r(t)$ and $X_{LO}(f)$ are from different origins and are independent, the product operation is taken out of the expectation as follows $E [X_{LO}(f_1).r(t - \frac{kf_1}{\pi})] = E[X_{LO}(f_1)].E[r(t - \frac{kf_1}{\pi})]$. It is known that $E[r(t)] = \sum_n E[c_n].E[e^{j\phi_{tx}(t)}].h_p(t - nT_s)$ which leads us to Equation 2.35.

$$E[r'(t)] = \sum_n E[c_n]. \int_{-\infty}^{+\infty} E[X_{LO}(f_1)].E[e^{j\phi_{tx}(t-\frac{kf_1}{\pi})}].h_p(t-nT_s-\frac{kf_1}{\pi}).e^{-jkf_1^2}e^{j2\pi f_1 t} df_1$$

(2.35)

For the constellations in which the transmitted symbols $c_n$ are independent and identically distributed with zero mean, the mean of the received signal is given by $E[r'(t)] = 0$. It will be the case for our study.

The general formula of the variance of a signal is given by $\sigma_{r'(t)}^2 = E[|r'(t)|^2] - [E[r'(t)]]^2$ which corresponds to $\sigma_{r'(t)}^2 = E[|r'(t)|^2]$ since the mean of our signal is 0. We know that $E[|r'(t)|^2] = E[r'(t).r^*(t)]$. We obtain Equation 2.36.

$$E[|r'(t)|^2] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} E\left[r(t - \frac{kf_1}{\pi}),r^*(t - \frac{kf_2}{\pi})\right].E[X_{LO}(f_1).X_{LO}^*(f_2)]e^{j2\pi f_1 t}e^{jkf_1^2} df_1 df_2$$

(2.36)

The symbols and the LO laser output are uncorrelated and, for the LO laser, the spontaneous emissions at two distinct frequencies are independent with zero cross-correlation. We obtain Equation 2.37.

$$E[|r'(t)|^2] = \int_{-\infty}^{+\infty} E\left|r(t - \frac{kf_1}{\pi})\right|^2].E[|X_{LO}(f_1)|^2] df_1$$

(2.37)

where $E[|X_{LO}(f_1)|^2]$ is the power spectral density of the LO laser. The LO laser will be considered to have a Lorentzian lineshape given by $E[|X_{LO}(f_1)|^2] = \frac{r_{sp}}{f_1^2 + (LW/2)^2}$ corresponding to Equation 2.7 and where $r_{sp}$ is the spontaneous emission noise power and $LW$ the linewidth of the laser. For digital modulation format, the correlation function of the transmitted signal is given by Equation 2.38.

$$E[|r(t)|^2] = \sum_i \sum_j E[c_i,c_j^*]h_p(t - iT_s)h^*_p(t - jT_s)$$

(2.38)

The transmitted symbols are independent and identically distributed. Which leads to $E[|r(t)|^2] = \sum_n E[|c_n|^2]|h_p(t - nT_s)|^2$ and to the final Equation 2.39 for the
variance of the received signal.

\[ \sigma^2_{r(t)} = \sum_n E[|c_n|^2] \int_{-\infty}^{+\infty} \left| h_{ps} \left( t - nT_s - \frac{kf}{\pi} \right) \right|^2 \cdot \frac{r_{sp}}{f_1^2 + (LW/2)^2} df \] (2.39)

To extend the statistical study, the autocorrelation function is derived. The signal has a zero mean, which corresponds to \( \gamma_r(\tau) = E[r'(t)r'^*(t + \tau)] \). We plug in Equation 2.33 knowing that expectations and integrals are linear operations and can swap and that the transmitter signal and the LO laser emissions are independent and identically distributed. We obtain Equation 2.40.

\[ \gamma_r(\tau) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} E \left[ r(t - \frac{kf_1}{\pi}) r^*(t + \tau - \frac{kf_2}{\pi}) \right] . E[X_{LO}(f_1) X_{LO}^*(f_2)] \cdot e^{j2\pi(f_1-f_2)\tau} . e^{-j2\pi f_2 \tau} . e^{jk(f_2-f_1)^2} df_1 df_2 \] (2.40)

Once again, the LO laser spontaneous emissions are independent at two different frequencies and have zero cross correlation. We obtain Equation 2.41.

\[ \gamma_{r'}(\tau) = \int_{-\infty}^{+\infty} E \left[ r(t - \frac{kf_1}{\pi}) r^*(t + \tau - \frac{kf_1}{\pi}) \right] . E[|X_{LO}(f_1)|^2] . e^{-j2\pi f_1 \tau} df_1 \] (2.41)

The first expectation term corresponds to the autocorrelation function of the transmitted signal: \( \gamma_r(\tau) = E [r(t - \frac{kf_1}{\pi}) r^*(t + \tau - \frac{kf_1}{\pi})] \). The second expectation term corresponds to the Lorentzian lineshape: \( E[|X_{LO}(f_1)|^2] = \frac{r_{sp}}{f_1^2 + (LW/2)^2} \). We plug in Equation 2.41 and obtain Equation 2.42.

\[ \gamma_{r'}(\tau) = \int_{-\infty}^{+\infty} \gamma_r(\tau) \frac{r_{sp}}{f_1^2 + (LW/2)^2} . e^{-j2\pi f_1 \tau} df_1 \] (2.42)

The autocorrelation function is slightly changed from that of the transmitted signal. This is verified in the simulations further in this thesis.
Chapter 3

Simulations Setup in VPI and Matlab

The simulation setup is provided by Acreo and uses the cosimulation tool between VPItranmissionMaker & VPIcomponentMaker and Matlab. The simulations run via VPItranmissionMaker & VPIcomponentMaker and fetch the data to be transmitted to Matlab for the DSP part. The results are then saved as Matlab matrix files. The global system can be represented as on the Figure 3.1.

**Figure 3.1:** Representation of the simulation setup in the VPI software.

On Figure 3.1, in order of appearance from left to right, one can observe:

- M-QAM transmitter
- Transmission and reception filters
- Dispersive fiber
- Noise loading module
- Frequency modulated LO laser
- Single polarization receiver
• DSP processing

3.1 VPItransmissionMaker & VPIcomponentMaker

The global system is implemented in VPI and then the data from the photodiodes is given to Matlab for the DSP. Each parts of the system is further explained.

3.1.1 Square M-QAM Transmitter

The square M-QAM transmitter from VPI shown as the block on Figure 3.2 is used to transmit QPSK and 16-QAM which are the two constellations to be used for this thesis project.

![Figure 3.2: The block representation of the square M-QAM transmitter from VPI.](image)

Figure 3.2: The block representation of the square M-QAM transmitter from VPI.

![Figure 3.3: A classical square M-QAM transmitter from VPI.](image)

Figure 3.3: A classical square M-QAM transmitter from VPI.

An inside view of the module is shown on Figure 3.3. The bit sequence used to encode the symbols goes through an encoder (bit mapping) which uses its output
to modulate the divided light of a transmitting laser of wavelength $\lambda$ (frequency $f_{Tx}$) in I and Q (original and $\pi/2$ shifted version) before recombining them in a power combiner. Further, the blocks set the carrier frequency and bandwidth to the logical channel and update the labels.

The coder and modulator driver block uses a sequence of bits to encode the symbols depending on the bit mapping. A pseudo-random binary sequence (PRBS) \cite{33} is used. A binary sequence of length $N$ is a sequence of bits $b_0, ..., b_{N-1}$ where $b_i \in \{0, 1\}$ for $i = 0, ..., N - 1$. The appellation binary means that each object in the sequence can only take two different values here 0 and 1. The parameter $m = \sum b_j$ gives the number of bits equal to 1 in the binary sequence.

It is defined that a binary sequence is a PRBS if its autocorrelation function is of the form of Equation 3.1.

$$\gamma[k] = \begin{cases} m, & \text{if } k \equiv 0 \ [\text{mod } N] \\ mc, & \text{otherwise} \end{cases}$$  \hspace{1cm} (3.1)

where $c = \frac{m-1}{N-1}$ is called the duty cycle of the PRBS. Half the bits are equal to 1 and the other half to 0, which corresponds to the case $c = 1/2$. The autocorrelation function of the PRBS is then given by Equation 3.2.

$$\gamma[k] = \begin{cases} N/2, & \text{if } k \equiv 0 \ [\text{mod } N] \\ N/4, & \text{otherwise} \end{cases}$$  \hspace{1cm} (3.2)

The order of PRBS used is 15, called PRBS15. The bit sequences of PRBS15 are concatenated to obtain a longer bit sequence and encode the symbols. The initial bit sequence is divided into odd (I) and even (Q) part: $I_j = b_{2j-1}$ and $Q_j = b_{2j}$, and IQ symbols are formed by Equation 3.3 \cite{34}. This formula holds for the square M-QAM transmitter.

$$IQ_{M-QAM}(k) = \sum_{l=1}^{N/2} 2^{N/2-l}(2.I(k_{l-1}).N/2+l-1) + i \cdot \sum_{l=1}^{N/2} 2^{N/2-l}(2.Q(k_{l-1}).N/2+l-1)$$  \hspace{1cm} (3.3)

Still in the coder and modulator driver, the symbols are given the pulse shape. They are scaled with the MZM characteristic voltage $V_\pi$. The four voltages at the output of the block are given on Equation 3.4 \cite{34}.

$$I_{rfup} = I_{rflo} = \frac{V_\pi}{\pi}.asin(Re(IQ))$$

$$Q_{rfup} = Q_{rflo} = -\frac{V_\pi}{\pi}.asin(Im(IQ))$$  \hspace{1cm} (3.4)
where $\text{Im}(.)$ and $\text{Re}(.)$ stand for imaginary and real part, respectively, and $\text{asin}(.)$ is the inverse sine in radians. These voltages are used to drive the modulation of the laser light in the transmitter to finally obtain the modulated optical output.

### 3.1.2 Transmission and Reception Filters

![Figure 3.4: Representation of the band-pass filters used at the transmission and at the reception.](image)

Two band-pass filters (represented by the block on Figure 3.4) are used, one after the transmitter and one before the receiver, to ensure that the signal has a limited bandwidth. The bandwidth is set to be 12.5 GHz. It is one bandwidth used to measure in optical systems.

### 3.1.3 Link

![Figure 3.5: The universal optical fiber block on VPI.](image)

The link used is an optical fiber with a dispersion coefficient $D$. The length of the fiber will vary depending on the simulation parameters. The block representing the optical fiber in VPI is given by Figure 3.5. PMD is not considered for this thesis work and the system works with a single polarization. The dispersion operator is specified by the group velocity dispersion (GVD) and the dispersion slope and follows Equation 3.5.

$$
\hat{D} = j \frac{\beta_2}{2} \frac{\partial^2}{\partial t^2} + \frac{\beta_3}{6} \frac{\partial^3}{\partial t^3}
$$

(3.5)

where $\beta_2 = -\frac{c}{2\pi f_{ref}} D_\lambda$ and $\beta_3 = -\frac{c}{(2\pi)^2 f_{ref}^3} \left( \frac{c}{f_{ref}} S_\lambda + 2D_\lambda \right)$. Here, $D_\lambda$ is the fiber dispersion, $S_\lambda$ is the dispersion slope, $c$ is the speed of light and $f_{ref}$ is the reference frequency.
3.1.4 Noise Loading Module

![Figure 3.6: Custom block to set the power of the noise to fit a given OSNR.](image)

The noise loading module is a custom block that is represented as on Figure 3.6. The way this module works is more detailed on Figure 3.7. It measures the power of the incoming signal with a power meter. Then depending on the wanted value of the OSNR, it will modulate the amplitude of the WGN that will be added to the incoming signal. This block emulate the noisy behaviour of the EDFA amplificators normally used for real transmission chains (which are placed approximately every 80 km) and which generate noise on the chain. This noise can be modeled as an AWGN and it is what the module is doing.

Using this module, any wanted value of the OSNR can be fixed. The OSNR is not changed by changing the optical input power but by adapting the amplitude of the AWGN in a smart way.

3.1.5 Local Oscillator

The local oscillator is depicted as on Figure 3.8. It is composed of an ideal laser transmitting at the wavelength $\lambda$ which correspond to a frequency of $f_{LO}$ and of
a module to load the pulse spectral shape of the laser. The ideal laser places
the pulse shape at the wanted frequency and the shape is given by the second
module. The second module uses one WGN, filtered until a given frequency,
which gives the white frequency noise bandwidth. The frequency given by the
ideal laser is modulated to emulate the local oscillator frequency noise. This
process is represented on Figure 3.9.

3.1.6 Polarization-Diversity Digital Coherent Receiver

Figure 3.10: Receive the optical signals from the link and the
local oscillator and converts it to digital after detection.
The coherent receiver is the same as the one showed on Figure 2.1. The block is pictured on Figure 3.10. The detail of the block is given on Figure 3.11. The two optical signals (one from the link and one from the LO) are demodulated using a 90-hybrid then fed into 4 photodiodes to retrieve the amplitude and phase information of the signal with an electrical signal. This electrical signal is then converted to digital with an ADC and is to be fed into the DSP module.

Figure 3.11: The received signal and the LO signal are mixed in 90-hybrid then detected by 4 photodiodes before being converted to digital with an ADC.

### 3.1.7 Digital Signal Processing Module

The DSP module is using the cosimulation tool with Matlab. It is pictured on Figure 3.12. It transmits the received signal to Matlab and the different parameters used on VPI and useful for the DSP, such as:

- The CD parameter
- The ADC factor
- The dispersion slope
- The length of the link
- The wavelength
- The symbol rate
- The number of bits per symbol


**3.2 Digital Signal Processing in Matlab**

In the Matlab environment, the DSP part of the system is implemented. It is an important part since it influences directly the performances. Without any DSP, the system will not work. The different algorithms used are explained here. To understand this part, it is necessary to have a full understanding of the algorithms presented in Chapter 2. The global DSP chain is given on Figure 3.13.

### 3.2.1 DC Removal

This part of the code ensures us that any DC component is removed from the incoming signal.

### 3.2.2 IQ Imbalance

The IQ imbalance module function was specified in Chapter 2.
3.2.3 Chromatic Dispersion Compensation

The CD equalizer chosen is the Blind Look-Up Filter, or Frequency Domain Equalization, which ensures that the CD will be totally compensated for. It is a key part of the DSP.

3.2.4 Resampling

The resampler has not been presented in the state of art, but it is an important part of the DSP as well. The signal is transmitted at higher rate than our actual baud rate by increasing the number of samples per symbol. In the simulation framework, 32 samples per symbol are transmitted. It helps us to synchronize our receiver on the good sampling time and improves significantly the quality of the transmission. The resampler will not give back only one sample per symbol but two samples per symbol (two samples per symbol will be needed for the equalizer). The resampler interpolates its input signal to provide 12 times more samples (increasing the sampling frequency by 12) and then down-samples to the wanted number of samples per symbols. The interpolator is the typical spline interpolator provided by Matlab.

To summarize, the resampler receives a signal containing 32 samples per symbol, then increases it to 384 samples per symbol and finally decreases it to 2 samples per symbol.

3.2.5 Equalizer

The equalizer is based on the Constant Modulus Algorithm (for QPSK modulation) and Multi Modulus Algorithm (for 16-QAM modulation and higher orders). It uses 3 iterations over the whole signal to ensure a better convergence of the filter taps. It is a fractionally spaced equalizer since it receives 2 input samples to produce one output sample and update the tap weights. There are 13 taps with initial weights set to 1. The value of the parameter $\mu = 6.10^{-4}$ ensures the convergence of the algorithm. This convergence is also ensured by the preconvergence length of 10000 symbols. Those symbols will not be considered after the equalizer.

3.2.6 Carrier Phase Recovery

The carrier phase recovery algorithm is the blind phase search. The algorithm is dividing the complex plane in 32 angle sectors, and uses 19 symbols to smooth the decision. The algorithm will average over 19 symbols to have a smoothed
approximation of the shifting angle and find the closest of the 32 angles which divide the complex plane. Then it will rotate back the last symbol respectively to this angle to retrieve the original position of the constellation. The algorithm has an angle resolution of $\frac{\pi}{16}$.

Though, the algorithm has a limitation, which appears in case of very significant phase noise. It can happen that some points rotated of an angle superior to $2\pi$. In this case, the algorithm misses the $2\pi$ rotation, which will lead in a cycle slip. The problem is that one needs to identify these cycle slips to be able to properly count the BER in the following module.

### 3.2.7 Bit Error Rate Counter

The BER counter knows the data we transmitted (as explained in the Section 3.1.1), it is a symbol sequence generated with a concatenation of PRBS15 bit sequences. Using this, one can easily predict what should be the received symbol sequence and compare it to the actual received symbol sequence. To properly count the errors, the cycles slips need to be considered. But we will not go further in this direction since it is not the purpose of this thesis work.
Chapter 4

Equalization Enhanced Phase Noise Effect on Coherent Fiber Optical Communications

4.1 Introduction to Numerical Simulations on Equalization Enhanced Phase Noise

The numerical simulations on EEPN are presented in this chapter. The simulation setup was specified in Chapter 3. The main parts of the setup are reminded here. The system is working in single polarization mode (same data sent over X and Y polarization). The transmitter sends either QPSK or 16-QAM symbols encoding a binary sequence based on concatenations of PRBS15 sequences. The baud rate is 28 GBaud. The transmitter laser operates at the frequency $f_{Tx} = 193.1THz$ (corresponds to a wavelength $\lambda = 1.5525\mu m$) and is ideal. The dispersive fiber does not implement any PMD, the link is purely dispersive with a dispersion coefficient $D = 16.10^{-6}s/m^2$. The LO laser operates with a central frequency $f_{LO} = 193.1THz$ and is modulated in frequency by a driving white noise. Finally, after the receiver, DSP is applied to the signal as presented on Figure 3.13.

The statistical study of EEPN is separated in three parts:

- OSNR penalty: the figure of merit is used to show how EEPN degrades the performances of the system and motivates the statistical study.

- Autocorrelation of the signal impaired by EEPN: the analytical formulas derived in Chapter 2 are confirmed by numerical simulations in order to show that EEPN has an autocorrelation function similar to the data.
• Autocorrelation of EEPN separated from the data: the transmitted symbols (which are known) are subtracted to the received signal after DSP to observe the noise only.

4.2 Influence of Equalization Enhanced Phase Noise on the OSNR Penalty

Here, the results are presented in two steps. First, a simulation of a transmission chain with two sets of parameters as presented on table 4.1. These simulations will give the BER vs OSNR graphs. Then, these graphs are used to calculate the OSNR penalty and obtain the OSNR penalty vs Impairment (linewidth of the LO laser) graph.

4.2.1 Simulation Method of OSNR Penalty

The OSNR penalty is used as a basis to understand how the linewidth of the LO laser degrades the system. The OSNR penalty can be calculated by using the graph of the BER against the OSNR. The difference between the reference OSNR (BER at FEC limit for a zero linewidth) and the OSNR at FEC limit for a different value of the impairment gives the OSNR penalty for the given impairment. The differences are gathered and plotted against the linewidth of the LO laser to obtain the OSNR penalty curve.

The FEC limit is set to be $10^{-3}$, below this limit it is considered that all the errors can be corrected by correcting codes.

The OSNR penalty curve gives the penalty in terms of OSNR that the impairment brings. An illustration is given on Figure 4.1 of how the OSNR penalty is measured, based on the BER vs OSNR curves for a varying linewidth. These measurements give the OSNR penalty curve to be compared to the theoretical value of the OSNR penalty [14] which is given by Equation 4.1.

$$Penalty = -10 \log_{10} \left( 1 - OSNR_{ref.inband} \cdot EVM^2_{EEPN} \right)$$  \hspace{1cm} (4.1)

where $EVM^2_{EEPN}$ is the square of error vector magnitude and is given by Equation 4.2 and $OSNR_{ref.inband}$ is the linear reference OSNR inside the signal bandwidth and is given by equation 4.3.

$$EVM^2_{EEPN} = \pi^2 \cdot |\beta_2| \cdot L.Baudrate.LW$$  \hspace{1cm} (4.2)
where $\beta_2 = \frac{\pi D}{2f_{Tx}}$ is the group velocity dispersion parameter which is given by $D$ the dispersion coefficient and $f_{Tx}$ the central frequency, $L$ is the length of the link, Baudrate is the baudrate and $LW$ is the linewidth of the LO.

$$OSNR_{ref,inband} = \frac{NBW_{ref}}{NBW_{inband}} \cdot 10^{OSNR_{ref}/10}$$ \hspace{1cm} (4.3)

where $NBW_{ref} = 12.5 \text{ GHz}$ is the reference noise bandwidth, $NBW_{inband} = 1.8 \times \text{Baudrate \ GHz}$ is the in-band noise bandwidth (the 1.8 coefficient comes from the Bessel filter used to filter the signal at the reception) and $OSNR_{ref}$ is the reference OSNR for a zero-impairment at the FEC limit. The linear OSNR inband corresponds to the linear OSNR which is normalized inside the bandwidth of the system.

### 4.2.2 BER vs OSNR Curves in Presence of Phase Noise

The simulations use the set of parameters specified by table 4.1. The modulation format for the simulations is 16-QAM, which for similar impairments will give a higher OSNR penalty than a lower modulation format (such as QPSK) would give. At similar average power, the constellation for 16-QAM has a shorter distance
between the points of the constellation, than the constellation for QPSK. This implies that the same impairment will lead in more errors when detecting the bits at the receiver, so that the BER is lowered for higher modulation format.

For the OSNR simulations, $2^{18}$ symbols for each points are used to ensure good statistical properties.

The simulations collect the BER for different values of the impairment considered (the linewidth of the LO laser) and for different values of the OSNR. Figure 4.2 for the link of 2000 km (top graph) and for the link of 4000 km (bottom graph) shows all the gathered data.

For the 4000 km link, the impairment brings a bigger penalty than for the 2000 km link. The longer the link, the more important EEPN is. The linewidth impairment (and EEPN indirectly) brings a huge penalty to the system. For a linewidth of 0 MHz to 1 MHz, the OSNR has to be increased of at least 1 dB to keep up with the BER which is a lot for optical communications. Considering this, one can understand why it is important to study EEPN and to try to mitigate it as much as possible to have a performant system.

### 4.2.3 Analytical OSNR Penalty vs Numerically Simulated OSNR Penalty

Figure 4.3 gathers the results for the analytical study and the previous simulations. For the analytical part, the Equation 4.1 is used, given the parameters specified above.

As expected, EEPN brings an important OSNR penalty to the system. In numerical and analytical simulations, the penalty goes up to 1.5dB if we use high linewidth. It implies that for very long links, lasers with narrow-linewidth [35–37] have to be used to minimize EEPN and maximize the channel capacity. But lasers with low linewidth can have other limitations that lasers with higher linewidth won’t have. Particularly, the price and the output power of the laser. There exist also other methods to mitigate EEPN [38, 39], but we will not discuss
Figure 4.2: This graph shows the curves of the BER for different values of the linewidth of the LO laser and different values of the OSNR. The link is 2000km on the first graph and 4000km on the second graph. The baudrate is 28 GBaud.
Figure 4.3: Gathers the results of the simulations on OSNR penalty for 2000km and 4000km link. We compare the Analytical curve to the Simulation curve.

more about these. Those conclusions motivate us to go in a deeper study of the EEPN impairment to understand if other mitigation techniques can be used.

4.3 Autocorrelation of the Signal Impaired by EEPN

This section is concentrating on the autocorrelation function of the signal at the reception, the expected autocorrelation function will be explained first, then the simulated autocorrelation functions will be given for the parameters given in Table 4.2 for QPSK and Table 4.3 for 16-QAM.

4.3.1 Simulation Method of Autocorrelation

One important step in understanding EEPN and finding a treatment to it, is to study its autocorrelation [40]. The problem is that EEPN occurs in presence of a dispersive link, high linewidth of the LO laser and electronic dispersion compensation. Thus, different impairments will occur at the same time: AWGN, CD,
Table 4.2: Parameters used for the simulations of QPSK to calculate the autocorrelation for three different cases: AWGN, PN and EEPN.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>AWGN</th>
<th>PN</th>
<th>EEPN1</th>
<th>EEPN2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linewidth LO (MHz)</td>
<td>0</td>
<td>2 or 5</td>
<td>2 or 5</td>
<td>2 or 5</td>
</tr>
<tr>
<td>Length of the link (km)</td>
<td>0</td>
<td>0</td>
<td>5000</td>
<td>10000</td>
</tr>
<tr>
<td>OSNR (dB)</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 4.3: Parameters used for the simulations of 16-QAM to calculate the autocorrelation for three different cases: AWGN, PN and EEPN.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>AWGN</th>
<th>PN</th>
<th>EEPN1</th>
<th>EEPN2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linewidth LO (MHz)</td>
<td>0</td>
<td>2 or 4</td>
<td>2 or 4</td>
<td>2 or 4</td>
</tr>
<tr>
<td>Length of the link (km)</td>
<td>0</td>
<td>0</td>
<td>5000</td>
<td>10000</td>
</tr>
<tr>
<td>OSNR (dB)</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
</tbody>
</table>

Phase noise and EEPN. To characterize EEPN, it needs to be dominant in the system. The choice has been made to have a long link (high CD which is almost perfectly compensated for, high EEPN which is desired), a low linewidth (low residual phase noise, high enough EEPN) and a high OSNR (low AWGN).

For the simulations, a fair comparison between PN and EEPN is crucial. To ensure this, PN and EEPN are compared for the same linewidth of the LO laser. In this way, the influence of the introduction of the link between the transmitter and the receiver will be the only changing parameter. The OSNR is fixed high, the linewidth of the LO laser is chosen low enough and the link is chosen accordingly to the considered case. The parameters used in the simulations are given by Table 4.2 for QPSK and Table 4.3 for 16-QAM.

To properly verify that the results from the autocorrelation respect the specifications of the system, we will calculate the autocorrelation at 5 different points in the DSP chain, which are depicted on Figure 4.4.

These autocorrelations will be called as follows:

- 1 - Autocorrelation before CD
- 2 - Autocorrelation before resampling
- 3 - Autocorrelation before equalization
- 4 - Autocorrelation before CPR
- 5 - Autocorrelation after DSP
The signal is a complex signal, which means that at each point the autocorrelation function will be calculated as in Equation 4.4.

\[ \gamma_{xx}[k] = E[E_X[n]E_X^*[n + k]] \]  

(4.4)

where \( E[.] \) is the expectation, and \( E_X[n] = I[n] + jQ[n] \) is the \( n^{th} \) sample of the signal \( E_X \), and I and Q stand for In-Phase and Quadrature-Phase components of the signal. Which leads to the final Equation 4.5.

\[ \gamma_{xx}[k] = E[I[n]I[n + k]] + E[Q[n]Q[n + k]] - jE[I[n]Q[n + k]] + jE[Q[n]I[n + k]] \]  

(4.5)

In this equation, there are the two first terms which are autocorrelation terms (real-real and imaginary-imaginary) and the two last terms which are crosscorrelation terms (real-imaginary and imaginary-real). The difference between these terms will play a significant role in the understanding of the shape of the autocorrelation function.

Another consideration is that the autocorrelation can either be power normalized if one wants to relatively compare the amplitude of the peaks from one impairment to the other. But later, we will remove the data part of the signal to obtain the autocorrelation function of only the noise part. In this second case, the noise power is important. Then, the autocorrelation is normalized by the length of the sequence to obtain the average power of the noise as the maximum value of the autocorrelation function.

In this work, the following characteristics of the autocorrelation functions are verified:

- The position of the side peaks, which has to fit the system’s specifications in terms of length of the PRBS pattern used.
• The amplitude of the side peaks, to understand how degrading the impairment is.

• The shape of the peaks, which will give the wave’s shape and will tell if the impairment adds intersymbol (or intersample) correlation.

• The number of peaks and the number of samples/symbols, which will ensure that the system works as expected.

4.3.2 Reference Autocorrelation

The shape of the autocorrelation function is understood in this section. To do this, we need to consider how the signal is generated at the transmitter with VPI. We know that the transmitter uses a PRBS of order 15 (PRBS15) to generate the incoming bits.

The PRBS15 has a length of $2^{15} - 1 = 32767$. In our study case, we want to generate very long sequences of symbols to have a signal with strong statistical properties, thus we will concatenate the PRBS15 at the transmitter to generate the symbols. For example, if we want to generate $2^{16}$ symbols for the QPSK modulation, we will need $2^{17}$ bits. So, we will repeat PRBS15 five times to have enough bits. Since: $5 \times (2^{15} - 1) > 2^{17}$. If we wanted the same number of symbols for 16-QAM modulation, we would have repeated PRBS15 nine times, and so on. We understand already that the last PRBS15 sequence to be concatenated will not be fully used, but it does not matter since the pseudorandom behaviour is conserved anyway.

After the PRBS15 has been repeated, we use the bit-mapping to generate the symbols that will be modulated and transmitted over the link. We explained previously, that the bit-mapping made the symbols generated with PRBS have an autocorrelation form which is 0 everywhere but when the delay is 0. We want to understand what will the autocorrelation function of the generated symbols will look like.

We know that in VPI, the way the bit mapping is done is that one bit over two is taken for I and the other one for Q, and on and on until all the bits have been distributed. Then the bit mapping is used to generate the symbols of the constellation. But one thing that can be seen is that the period of repetition of PRBS is 32767 which is not even, we want to find the period of repetition of the symbols sequence generated. We need to separate the case where the modulation format is QPSK and when it is 16-QAM.
Length of the Autocorrelation Function

This section concerns both 16-QAM and QPSK modulations. Here, the number of samples at every stage of the DSP routine is predicted. To do this, one needs to consider all the filters and delays that are added in the DSP part. That work is essential since it will be shown later that the number of samples of the sequence is an important parameter to determine the shape of the autocorrelation function (triangle window function in the following sections).

For the simulations, the number of transmitted symbols is a power of 2, of the form: $2^N$.

The system works at a higher rate than the symbol rate represented by the ADC factor of 32. It means that for one transmitted symbol, 32 samples are used to represent it. The number of samples transmitted through the link is then $32 \times 2^N$. That number will be the number of samples received at the very beginning of the DSP chain. The number of symbols before any DSP is then $32 \times 2^N$. The CD compensation and the IQ compensation modules don’t change the number of samples.

The first module in the DSP chain to change the number of samples is the resampler. The resampler changes the number of samples per symbol from 32 to 2 (the CMA/MMA equalizer uses 2 samples per symbol). The number of samples after the equalizer should be: $\frac{32 \times 2^N}{16} = 2 \times 2^N$. But the resampler is a bit more complicated. It is composed of two different modules: one spline interpolator and one decimator. The interpolator increases the sampling rate by 12 before it is decreased to the desired value by the decimator. And for technical constraints, the interpolator makes us lose 1000 samples at a rate 12 times higher. After both the interpolator and the decimator, 168 samples are lost. The number of samples after the resampler is: $2 \times 2^N - 168$.

The second module in the DSP chain to change the number of samples is the CMA/MMA equalizer. The equalizer has a preconvergence length of 10000 samples. The algorithm also uses 2 samples per symbol to convert them to 1 sample per symbol. Added to this, we have 3 iterations of the algorithm which looses 5 samples. After the equalizer, the number of samples in the sequence is: $\frac{2^N + 168}{2} - 10000 - 5 = 2^N - 10089$.

The last module in the DSP chain to change the number of samples is the CPR. The BPS algorithm uses an averaging window of length $2 \times 9 + 1 = 19$, 19 samples (that are symbols in the BPS case) are lost. After the BPS algorithm, the number of samples in the sequence is: $2^N - 10108$. 
### Table 4.4: Number of samples of the sequence at every stage of the DSP chain for $32 \times 2^N$ samples transmitted.

<table>
<thead>
<tr>
<th>Stage</th>
<th>Number of samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before any DSP</td>
<td>$32 \times 2^N$</td>
</tr>
<tr>
<td>After CD and IQ</td>
<td>$32 \times 2^N$</td>
</tr>
<tr>
<td>After resampler</td>
<td>$2 \times 2^N - 168$</td>
</tr>
<tr>
<td>After equalizer</td>
<td>$2^N - 10089$</td>
</tr>
<tr>
<td>After CPR</td>
<td>$2^N - 10108$</td>
</tr>
</tbody>
</table>

Table 4.5: Bits distribution in the transmitter for QPSK, I stands for In-Phase and Q for Quadrature-Phase.

<table>
<thead>
<tr>
<th>1st Sequence</th>
<th>2nd Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_1$ I $b_2$ Q $b_3$ I ... $Q b_{N-1}$ $b_N$</td>
<td>$Q b_1$ I $b_2$ Q ... $I b_{N-1}$ $b_N$</td>
</tr>
<tr>
<td>3rd Sequence</td>
<td></td>
</tr>
<tr>
<td>$b_1$ I $b_2$ Q $b_3$ I ... $Q b_{N-1}$ $b_N$</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.4 summarizes the number of samples at every stage of the DPS chain. The number of samples is decreasing along the DSP chain. Thus, the autocorrelation function will be slightly changed.

### Autocorrelation for QPSK

In the case of QPSK, Table 4.5 illustrates how the bits from PRBS are mapping on I and Q. For one sequence of PRBS15 the bits are assigned to I and Q. For the next sequence of PRBS15, the bits that were assigned to I are assigned to Q and the bits that were assigned to Q are assigned to I. And for the next sequence, it comes back to the situation of the first sequence.

This way, the symbols have a period that is the same as the repetition of PRBS15: 32767. Which is calculated as follows: $\frac{L_{PRBS15} \times n_s}{n_b}$, where $L_{PRBS15} = 32767$ is the length of PRBS15, $n_s = 2$ is the number of repetition of PRBS15 before the symbols take the same values again and $n_b = 2$ is the number of bits per symbol.

But, the autocorrelation is calculated with Equation 4.5. It says that the autocorrelation considers both the autocorrelation and the crosscorrelation terms of I and Q. In our case, for an integer $i$ and $k$ taken from Equation 4.5:

- for $k = i \times L_{PRBS15}$, the autocorrelation terms are non-null and the crosscorrelation terms are equal to zero.
Chapter 4. Equalization Enhanced Phase Noise Effect on Coherent Fiber Optical Communications

![Figure 4.5](image)

**Figure 4.5:** The autocorrelation function in ideal conditions after all the DSP chain for two different numbers of symbols: $2^{16}$ and $2^{17}$.

- for $k = \frac{L_{PRBS15}}{2} + i \times L_{PRBS15}$, the autocorrelation terms and the term $E[Q[n]I[n + k]]$ are equal to zero and the last crosscorrelation term is non-null.

- for $k = \frac{L_{PRBS15}+1}{2} + i \times L_{PRBS15}$, the autocorrelation terms and the term $E[I[n]Q[n + k]]$ are equal to zero and the last crosscorrelation term is non-null.

- for any other delay, all the terms are equal to zero.

For the non-null terms, what is the influence of each of them on the amplitude of the autocorrelation. If the symbol sequences were of infinite length, each non-null term would add $1/2$ to the amplitude of the autocorrelation. Because half of the bits of the whole sequence are ones (either I or Q). Then, the amplitude of the peaks and the autocorrelation function if the sequence was of infinite length:

- for $k = i \times L_{PRBS15}$, the amplitude of the autocorrelation function is 1.
### Table 4.6: Bits distribution in the transmitter for 16-QAM, I stands for In-Phase and Q for Quadrature-Phase. The indexes stand for 1st and 2nd bit on either I or Q (two bits are used on I and two bits on Q).

<table>
<thead>
<tr>
<th>1st Sequence</th>
<th>2nd Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I_1$ $b_1$ $Q_1$ $b_2$ $I_2$ $b_3$ $Q_2$ $b_4$ ... $Q_1$ $b_N$</td>
<td>$Q_2$ $b_1$ $I_1$ $b_2$ $Q_1$ $b_3$ $I_2$ $b_4$ ... $I_1$ $b_N$</td>
</tr>
<tr>
<td>$I_2$ $b_1$ $Q_2$ $b_2$ $I_1$ $b_3$ $Q_1$ $b_4$ ... $Q_1$ $b_N$</td>
<td>$Q_2$ $b_1$ $I_2$ $b_2$ $Q_1$ $b_3$ $I_1$ $b_4$ ... $I_1$ $b_N$</td>
</tr>
<tr>
<td>$I_1$ $b_1$ $Q_1$ $b_2$ $I_2$ $b_3$ $Q_2$ $b_4$ ... $Q_1$ $b_N$</td>
<td>$Q_2$ $b_1$ $I_1$ $b_2$ $Q_1$ $b_3$ $I_2$ $b_4$ ... $I_1$ $b_N$</td>
</tr>
</tbody>
</table>

- for $k = \frac{L_{\text{PRBS15}}-1}{2} + i L_{\text{PRBS15}}$, the amplitude of the autocorrelation function is $1/2$.
- for $k = \frac{L_{\text{PRBS15}}+1}{2} + i L_{\text{PRBS15}}$, the amplitude of the autocorrelation function is $1/2$.
- for any other delay, the amplitude of the autocorrelation function is 0.

But the sequence is of finite length. It corresponds to multiply the infinite length sequence by a rectangular window of the length of the finite length sequence. It will modify the amplitude of the autocorrelation function. That amplitude is multiplied by a triangular function of maximum amplitude 1 for $k = 0$ and minimum amplitude 0 for $k > L_{\text{sequence}}$ or $k < -L_{\text{sequence}}$ where $L_{\text{sequence}}$ is the length of the finite length sequence.

Figure 4.5 shows the simulations for $2^{16}$ and $2^{17}$ symbols with ideal conditions of transmission. The parameters used for these simulations are: 30dB OSNR, 0MHz linewidth of the LO laser and a link of 0km. This confirms all the results presented in this section and on the Table 4.4.

**Autocorrelation for 16-QAM**

In the case of 16-QAM, Table 4.6 illustrates how the bits from PRBS are mapping on I and Q. For one sequence of PRBS15 the bits are assigned to I and Q. The idea is similar to the one for QPSK but this time twice as many sequences of PRBS15 are needed to get a periodicity. It means that the same bits map to I and Q after 4 repetitions of PRBS15. Thus, in the same way as QPSK, one can show that:
for $k = i \ast L_{PRBS15}$, the autocorrelation terms are non-null and the crosscorrelation terms are equal to zero.

- for $k = \frac{L_{PRBS15}+1}{4} + i \ast L_{PRBS15}$ or $k = \frac{L_{PRBS15}-3}{4} + i \ast L_{PRBS15}$ or $k = 3 \ast \frac{L_{PRBS15}+1}{4} + i \ast L_{PRBS15}$ or $k = 3 \ast \frac{L_{PRBS15}-3}{4} + i \ast L_{PRBS15}$, the autocorrelation terms are equal to zero, a part of the crosscorrelation terms are equal to zero and the rest of the crosscorrelation terms is non-null.

- for $k = \frac{L_{PRBS15}+1}{2} + i \ast L_{PRBS15}$, the autocorrelation terms are equal to zero, half of the crosscorrelation terms is equal to zero and the rest of the crosscorrelation terms is non-null.

- for any other delay, all the terms are equal to zero.

The amplitude of the autocorrelation peaks is harder to predict than for QPSK. But to simplify, it is considered that the bit-sequence has an infinite length, and that it contains as many ones than zeros. It leads to give the following theoretical amplitude for the autocorrelation function:

- for $k = i \ast L_{PRBS15}$, the amplitude of the autocorrelation function is 1.

- for $k = \frac{L_{PRBS15}+1}{4} + i \ast L_{PRBS15}$ or $k = \frac{L_{PRBS15}-3}{4} + i \ast L_{PRBS15}$ or $k = 3 \ast \frac{L_{PRBS15}+1}{4} + i \ast L_{PRBS15}$ or $k = 3 \ast \frac{L_{PRBS15}-3}{4} + i \ast L_{PRBS15}$, the amplitude of the autocorrelation function is 1/4.

- for $k = \frac{L_{PRBS15}+1}{2} + i \ast L_{PRBS15}$, the amplitude of the autocorrelation function is 1/2.

- for any other delay, all the terms are equal to zero.

The reader needs to keep in mind that these values are expected values in the ideal situation, and that they are only approximations. The real implementation will not provide exactly the same results.

As for QPSK, the sequence is of finite length. It corresponds to multiply the infinite length sequence by a rectangular window of the length of the finite length sequence. It modifies the amplitude of the autocorrelation function. The amplitude is multiplied by a triangular function of maximum amplitude 1 for $k = 0$ and minimum amplitude 0 for $k > L_{sequence}$ or $k < -L_{sequence}$ where $L_{sequence}$ is the length of the finite length sequence.

Figure 4.6 shows the simulations for $2^{16}$ and $2^{17}$ symbols with ideal conditions of transmission. The parameters used for the simulations are: 30dB OSNR, 0MHz linewidth of the LO laser and a link of 0km. It enables to confirm all the results presented in this section and on the Table 4.4.
Chapter 4. Equalization Enhanced Phase Noise Effect on Coherent Fiber Optical Communications

4.3.3 Simulated Autocorrelation

In this section, the simulation results of the computation of the autocorrelation function are shown. The different sets of parameters are specified and the results for three types of impairment (AWGN, PN and EEPN) are compared. It will be considered that the case with AWGN is the reference since the value of OSNR has been taken high enough to ensure an error-free transmission.

Autocorrelation for QPSK

In the case of QPSK, the set of parameters specified on Table 4.2 is used. And the 5 points where the autocorrelation function is calculated are the ones specified by Figure 4.4.

Two series of 5 graphs from the simulations of the autocorrelation function at the 5 different points of the DSP are given on Figure 4.7.

**Figure 4.6:** The autocorrelation function in ideal conditions after all the DSP chain for two different numbers of symbols: $2^{16}$ and $2^{17}$. 
Chapter 4. Equalization Enhanced Phase Noise Effect on Coherent Fiber Optical Communications

Figure 4.7: The autocorrelation function at the 5 previously described points and for a linewidth of the LO of 2MHz (top) and 5MHz (bottom).
Figure 4.8: Zoom on the central peak of the autocorrelation function at the 5 previously described points and for a linewidth of the LO of 2MHz (top) and 5MHz (bottom).
Figure 4.9: Zoom on the side peak of the autocorrelation function at the 5 previously described points and for a linewidth of the LO of 2MHz (top) and 5MHz (bottom).
Figure 4.10: Constellations for different impairments after all DSP for a linewidth of the LO of 2MHz (top) and 5MHz (bottom).
Figure 4.7 gives us a global view of the autocorrelation and we can see that the final shape (5th graph) after all the DSP looks like what we expected. The side peaks are spaced as expected:

- every 524272 samples before any DSP
- every 524272 samples after CD compensation
- every 32767 samples after resampling
- every 16384 samples/symbols after equalization
- every 16384 samples/symbols after all the DSP

The amplitude of the side peaks in the case of AWGN is close to the ideal case. The zones of interest are zoomed in: the central peak on Figure 4.8 and the second side peak on Figure 4.9. To understand the graphs, the constellations are provided on Figure 4.10.

Figure 4.8 shows that the broadening of the central peaks vary inside the DSP chain. The autocorrelation function is calculated with the samples of the signal, not the symbols, so that when simulated, the peaks should be of width:

- every 2*32+1 samples before any DSP
- every 2*32+1 samples after CD compensation
- every 2*2+1 samples after resampling
- every 2*1+1 samples/symbols after equalization
- every 2*1+1 samples/symbols after all the DSP

The graphs confirm this assumption.

The zoom on the central peaks also shows that, wherever the autocorrelation is calculated and whatever the case (AWGN, PN or EEPN), the central peaks are all the same. Since the power normalized autocorrelations is used for the computations, the function is expected to take 1 for value when the lag is 0. Phase noise is a slow varying process compared to the symbol rate, that implies that the phase noise term in the calculation of the autocorrelation function (represented by a complex exponential), when the lag is small, have almost similar values on both the delayed and not delayed signals used to calculate the autocorrelation. This explains why the central peaks are similar on the graphs.

Figure 4.9 shows that before the CPR algorithm, the amplitude of the side peaks is only dependent on the linewidth of the LO laser. So, the higher the
Chapter 4. Equalization Enhanced Phase Noise Effect on Coherent Fiber Optical Communications

linewidth of the LO laser, the lower the amplitude of the side peaks. The drop in amplitude of the peak before the CPR algorithm is the same for PN and EEPN.

Now, the autocorrelation function after the CPR shows that that for PN, the peak always goes up almost as high as the peak for AWGN is. In the case of EEPN, the appearance of errors in the DSP (visible on the constellation) lowers the amplitude of the side peak significantly. When the system is error free, the simulation results join the analytical study of the autocorrelation function derived in Chapter 2: the signal impaired by EEPN has an autocorrelation function similar to the transmitted signal.

**Autocorrelation for 16-QAM**

In the case of 16-QAM, the set of parameters specified on Table 4.3 is used. And the 5 points where the autocorrelation function is calculated are the ones specified by Figure 4.4.

Two series of 5 graphs from the simulated autocorrelation function at the 5 different points of the DSP are given on Figure 4.11.

Figure 4.11 gives us a global view of the autocorrelation and we can see the final shape (5th graph) after all the DSP looks like what we expected. The side peaks are spaced as expected:

- every 262136 samples before any DSP
- every 262136 samples after CD compensation
- every 16384 samples after resampling
- every 8192 samples/symbols after equalization
- every 8192 samples/symbols after all the DSP

The amplitude of the side peaks respects the expectations in the case of AWGN since it is very close to the ideal transmission case (considering the triangular windowing function). The parts of interest are zoomed in: the central peak on Figure 4.12 and the 4th side peak on Figure 4.13. To understand the graphs, one needs to consider the constellations provided on Figure 4.14.

The 16-QAM case shows similar results than the QPSK case: the central peaks look the same at all the calculation points and the 4th peak shows similar properties than the 2nd peak for QPSK. But one can wonder why are the peaks for 5000km and 10000km of similar amplitude in the case of a linewidth of 4MHz. The answer to this is still the same; the amplitude of the side peaks is induced by
Figure 4.11: The autocorrelation function at the 5 previously described points and for a linewidth of the LO of 2MHz (top) and 4MHz (bottom).
Figure 4.12: Zoom on the central peak of the autocorrelation function at the 5 previously described points and for a linewidth of the LO of 2MHz (top) and 4MHz (bottom).
Figure 4.13: Zoom on the side peak of the autocorrelation function at the 5 previously described points and for a linewidth of the LO of 2MHz (top) and 4MHz (bottom).
Figure 4.14: Constellations for different impairments after all DSP for a linewidth of the LO of 2MHz (top) and 4MHz (bottom).
the errors of the CPR, and the constellations for 5000km and 10000km look the same, which explains why the amplitudes are the same.

Similarly to the case of QPSK, EEPN does not increase the correlation in the received signal and the analytical results are coherent with the simulation results.

4.4 Autocorrelation of EEPN Separated from Data

![Figure 4.15: Noise Constellation (top) and autocorrelation (bottom) of PN (left) and EEPN (middle and right) for the parameters specified on Table 4.2 for the case of a linewidth of the LO laser of 2MHz.](image)

We wonder what would the autocorrelation look like if we were to consider only the autocorrelation of the noise. Since the transmitted sequence is known, the received symbols can be predicted. The transmitted data is subtracted to the received signal after all the DSP chain so that the autocorrelation of the noise itself can be calculated after all DSP. It is a tricky operation because of the phase noise. The problem is that if the linewidth of the LO laser is too high, cycle slips can appear, which needs to be considered to predict the received data at the very end. Since we don’t want to deal with cycle slips, the data subtraction for the QPSK modulation will be made at the linewidth of 2MHz. The QPSK modulation is more robust to cycle slips than the 16-QAM modulation. And with the chosen linewidth, we avoid them.
The autocorrelation function of the noise for 3 cases are compared: one case with PN and two cases with EEPN. The parameters used for the simulation are given on Table 4.2 for the case of a linewidth of the LO laser of $2\,MHz$. The results are given on Figure 4.15.

To calculate these autocorrelation functions, the autocorrelation normalized by the length of the sequence (to be able to read the power of the noises on the central peaks) is used. The Figure shows that the amplitude of the central peak of the autocorrelation function increases with the length of the link. In other words, the power of the noise increases with the length of the link. If the noise power increases, the power needed at the transmission to keep a BER value has to be increased so that the OSNR needs to be increased. That links us directly to the influence of EEPN on the OSNR penalty, the bigger the product “distance-linewidth”, the bigger is the penalty on the system.

![Figure 4.16: Zoom on the autocorrelation function of EEPN only in the cases EEPN1 and EEPN2 of Figure 4.15.](image)

The autocorrelation function for EEPN shows side peaks positioned exactly like the side peaks of the data (see Figure 4.16). The amplitude of the side peaks is different than for the data, but it shows that the autocorrelation of EEPN depends on the autocorrelation of the signal. Thus, the spectrum of EEPN is similar to the spectrum of the signal. That confirms that EEPN is introduced by multiple delayed versions of the signal with itself. The autocorrelation function of the signal is not the autocorrelation function of a white noise which confirms the hypothesis that EEPN in not a white noise. And since EEPN is dependent on the data, it can not be compensated for by DSP post processing.
Chapter 5

Conclusion

EEPN is an enhanced noise induced by the insertion of a dispersive link between the transmitter and the receiver in the presence of imperfection in the LO laser (linewidth) and electronic dispersion compensation. It is caused due to interference of multiple delayed versions of the dispersed signal, generated by intermixing of the received dispersed signal and the noise side bands of the LO in the photodetectors. The goal of this study was to investigate if this noise can be compensated by only digital post-processing.

The figure of merit OSNR penalty is used to motivate the study. It shows that EEPN is degrading significantly the performances of the system. For a reasonable linewidth of the LO laser and length of the link (up to $1MHz$ with a link of $2000km$ to $4000km$), it has been seen that EEPN introduces up to $1.5dB$ of OSNR penalty. This value is high and motivates the research for mitigation techniques regarding EEPN.

The study was centered on statistical analysis of EEPN both in analytical and numerical simulations. The analytical study shows that the autocorrelation function of the signal impaired by EEPN is similar to the autocorrelation of the transmitted signal. The assumption made is that EEPN has the same spectral characteristics as the transmitted signal and that it can’t be compensated by classical digital post-processing.

Further on, the numerical simulations of the autocorrelation function confirmed the previous assumption. They have been made for two modulation formats: QPSK and 16-QAM. The system was using the following parameters: a dispersive link of dispersion coefficient $D = 16.10^{-6}s/m^2$, a central frequency $f_0 = 193.1THz$, $30dB$ OSNR, single polarization, no PMD and a baud rate of $28GBaud$. Three types of impairments were compared: AWGN (considered the reference, B2B propagation and no linewidth), PN (B2B propagation and linewidth of $2 - 4 - 5MHz$) and EEPN (link of $5000 - 10000km$ and linewidth of $2 - 4 - 5MHz$). The results
Chapter 5. Conclusion

showed that the central peaks of the autocorrelation functions are the same independently of the impairment due to the fact that the linewidth of the LO laser is negligible against the baud rate. The side peaks of the autocorrelation function are also similar when the system is error free (the introduction of errors in the DSP lowers the peaks). It confirms that the signal impaired by EEPN has a similar autocorrelation function than the transmitted signal.

Finally, the numerical simulations of the autocorrelation function of the noise itself have been made. The setup was the same as previously for QPSK modulation format and the set of parameters: $2MHz$ linewidth and $5000km$ link. This simulation could only be made for a small value of the “linewidth-length of the link” because otherwise cycle slips appear and make the simulation too complicated. It showed that the higher the previous product the higher the power of the noise. And, the autocorrelation of the noise itself shows side peaks at the same position than the signal impaired by EEPN. This confirmed even more that EEPN is dependent on the transmitted data and that its spectral characteristics are similar to the signal.

To summarize:

• EEPN is caused due to interference of multiple delayed versions of the dispersed signal, generated by intermixing of the received dispersed signal and the noise side bands of the LO in the photodetectors.

• The signal impaired by EEPN has an autocorrelation function similar to the transmitted signal since the linewidth of the LO laser is a slow varying process compared to the baud rate of the system.

• EEPN is not a white noise since it has been shown that it has an autocorrelation function dependent of the autocorrelation function of the signal. And the spectrum of EEPN itself is similar to the spectrum of the signal.

The analytical and numerically simulated studies of EEPN lead us to the conclusion that it would not be practical to treat EEPN by classical digital post-processing since it has the same spectral characteristics as the signal itself. As a future work, one could consider to investigate hardware techniques to mitigate EEPN [41].
Bibliography


[34] “Coder and modulator driver.” In: VPI help ().


