UNIVERSITÀ DEGLI STUDI DI PARMA

Dottorato di Ricerca in Tecnologie dell'Informazione

XXIV Ciclo

POLARIZATION-DEPENDENT NONLINEAR EFFECTS IN COHERENT DETECTION SYSTEMS

Coordinatore:

Chiar.mo Prof. Marco Locatelli

Tutor:

Chiar.mo Prof. Alberto Bononi

Dottorando: *Donato Sperti*

Gennaio 2012

Alla mia Famiglia

Contents

Introduction

In the last decades the demand for data capacity has increased exponentially. Optical Coherent Detection, firstly proposed at the end of the 1980s to improve receiver sensitivity, has proved as one of the most powerful techniques to increase the optical communication spectral efficiency and so the total per channel capacity. Indeed, thanks to the recent advances in digital signal processing (DSP) and high speed electronics, the DSP-based coherent detection in optical networks expedited the use of polarization division multiplexing (PDM) as a cost-effective way of doubling system capacity. Furthermore, coherent detection presents many others advantages with respect to direct detection such as the use of multilevel optical modulation formats like N-PSK and N-QAM and compensating linear propagation effects in the electrical domain as chromatic dispersion, polarization mode dispersion (PMD) and optical filtering. On the other hand, transmission reach of WDM systems is a major concern for the deployment of such a solution and is usually mainly limited by cross-nonlinear effects. In WDM transmission systems, the cross-nonlinearities make neighboring channels interact depending on their power and state of polarization (SOP). The last is of particular concern in PDM systems since they are more sensitive to a new kind of distortion that has been generally referred to as cross-polarization modulation (XPolM) [1] as a way to distinguish it from the well known cross-phase modulation (XPM).

At the beginning of our research activity in 2009, despite the growing interest and the number of publications on XPolM, many of its features were still unknown. For example, in Sept. 2009 Winter et al. provided a model that successfully measured the degree of polarization degradation in presence of XPolM [2], but it was still not clear when the bit error rate (BER) is dominated by XPolM and how XPolM relates to the other relevant nonlinear effects, such as XPM and self-phase modulation (SPM). With the investigations presented in this thesis we want to fill the gap, by providing a systematic simulation study of system performance where each nonlinear effect acts individually. Furthermore, thanks to the possibility in Optilux software [3] to take into account separately the nonlinear terms of the

propagation equation, we add some new piece of knowledge about XPolM. We quantify the XPolM-induced penalty as a function of transmission parameters such as the channel power, spacing and state of polarization (SOP). We also clarify the role of the Viterbi and Viterbi-based carrier phase estimator in mitigation of XPM and XPolM. We focused our investigation mainly on PDM-quadrature phase shift keying (QPSK) modulation format.

The thesis is organized as follows.

In the first chapter the principal impairments for long haul transmissions are briefly recalled. They are divided into linear and nonlinear effects, according to whether they are independent of the signal power or not. The first group is composed of fiber attenuation, chromatic dispersion and polarization mode dispersion. The second group is composed of nonlinear polarization-independent effects: such as SPM and XPM. Other linear effects such as polarization dependent loss and nonlinear effects as intra channel cross phase modulation, four wave mixing, nonlinear phase noise and non elastic scattering effects (stimulated Raman and Brillouin scattering) are not included in our discussion, while the XPolM is discussed at length in Ch. 3.

The second chapter discusses the joint use of PDM and the coherent detection, as a solution to increase the transmission capacity. We also discuss a new technique, namely mode division multiplexing (MDM), to further increase the transmission capacity thank to the joint use with PDM and coherent detection.

In Ch. 3, after the definition of the XPolM term in the propagation equation, we show the polarization rotation and the PDM-QPSK constellation distortion induced by XPolM as a function of the rotation axis orientation. We perform such analysis both mathematically and by simulation.

In Ch. 4 we show when the bit error rate (BER) of a PDM-QPSK channel is dominated by XPolM, through a massive use of simulation in a wide range of system setups. We analyze different pulse shapes, transmission links and transmission fibers in both hybrid (PDM-QPSK – OOK) and homogeneous systems (PDM-QPSK). Furthermore we clarify the role of channel power, spacing, state of polarization (SOP) and Viterbi and Viterbi-based carrier phase estimator on the XPM- and XPolM-induced penalty. In the last part of the chapter we quantify the nonlinear penalty in a PDM-BPSK transmission system, showing the average performance and its fluctuation induced by the transmission sequences and SOPs.

In Ch. 5 we compare different optical methods to improve the resilience of coherent 112-Gb/s PDM-QPSK WDM transmissions against cross-channel nonlinearities. Such methods consist of increasing the line group velocity dispersion (GVD), or the line PMD, or inserting in-line XPM suppressors, which are passive devices that introduce different delays on

adjacent channels at specific points of the line.

In Ch. 6 we summarize the experimental results obtained during the research activity at Alcatel-Lucent Bell-Labs France on MDM. In such an activity we employ a mode converter based on a liquid-crystal on silicon (LCOS) spatial modulator and a prototype few mode fiber (FMF).

Last but not least, in the Appendix we discuss three different rules to correctly simulate the cross-nonlinearities, showing also some artifacts that can arise with a non-correct setting of some numerical parameters, such as the nonlinear step of Split-Step Fourier method, the sequence length and the sequence type.

Remarks on notation

The following notation is adopted throughout this thesis: Jones matrix and M*ü*eller matrix are denoted by capital letters, i.e., $T(z,\omega)$, $M(\omega)$. All vectors are denoted by an arrow and unit magnitude vectors by an hat, except 4-D Pauli vectors, which are underlined. Hence *(2x1)* Jones vectors are denoted by capital letters and appear as $\overrightarrow{E}(z,\omega)$, $\overrightarrow{A}(\omega)$ and $\widehat{J}$, while *(3x1)* real Stokes vectors are denoted by lower-case letters and appear as $\overrightarrow{e}(z,\omega)$, $\overrightarrow{a}(\omega)$ and $\widehat{j}$. A column vector with elements a,b,c is expressed as $[a;b;c]$. The symbols $\cdot$ and $\times$ denote vector scalar and cross product, ^{T} denotes matrix transpose and † stands for transpose-conjugate. The zero-th Pauli matrix σ_0 it the *(2x2)* identity matrix, while σ_i $i=1,2,3$ are the other Pauli matrices, defined in App. B.

The notation adopted in this thesis is summarized in the table 2:

	Examples of notation
Jones matrices	$U(z,\omega), T(z,\omega)$
Mueller matrix	$M(z,\omega)$
Jones vectors: unitary magnitude and not	$\widehat{J}(z,\omega), \overrightarrow{E}(z,\omega), \overrightarrow{A}(\omega)$
Stokes vectors: unitary magnitude and not	$\widehat{j}(z,\omega), \overrightarrow{e}(z,\omega), \overrightarrow{a}(\omega)$
Pauli vectors: 4-D and 3-D	$\underline{h}(z), \underline{u}(z,\omega) \in \mathbb{C}^4 \quad \check{h}(z) \in \mathbb{C}^3$
	$\Rightarrow \underline{h} = [h_0, \check{h}^T]^T$
Pauli matrices	$\sigma_i \quad$ with $i \in \{0,1,2,3\}$
Spin vectors: 4-D and 3-D	$\underline{\sigma}(4-D)$; $\overrightarrow{\sigma}(3-D)$

Table 2: Summary of notation.

Chapter 1

Fiber Optic Transmission channel

This first chapter discusses the propagation equations that describe the evolution of an electric field along the optical fiber both in absence and in presence of polarization effects, namely: 1) nonlinear Schrödinger equation (NLSE), 2) coupled nonlinear Schrödinger equation (CNLSE), and 3) Manakov-PMD equation.

Starting from these equations, the principal impairments for long haul transmissions are briefly recalled. An exhaustive description of such a fiber impairments is out of the scope of this thesis and more details on them can be found in the references given throughout the text.

The fiber impairments are divided into linear and nonlinear effects, according to whether they are independent of the signal power or not. The first group is composed of fiber attenuation, chromatic dispersion and polarization mode dispersion (PMD). The second group is composed of nonlinear polarization-independent effects such as SPM and XPM. Other linear effects such as polarization dependent loss and nonlinear effects as intra channel cross phase modulation, four wave mixing, nonlinear phase noise and non elastic scattering effects (stimulated Raman and Brillouin scattering) are not included in our analysis, while the XPolM will be discussed in depth in Ch. 3.

1.1 NLSE

The scalar nonlinear Schrödinger equation (NLSE) is a good model to study the linear and nonlinear transmission impairments in the fiber in absence of polarization effects, i.e., when the signals are lunched in a single polarization. The NLSE for an electric field $A(z,t)\,[\sqrt{W}]$ is the following partial differential equation (PDE) [4]:

$$\frac{\partial A(z,t)}{\partial z}+\frac{\alpha}{2}A(z,t)+\beta_1\frac{\partial A(z,t)}{\partial t}-\frac{i}{2}\beta_2\frac{\partial^2 A(z,t)}{\partial t^2}-\frac{\beta_3}{6}\frac{\partial^3 A(z,t)}{\partial t^3}=-i\gamma|A(z,t)|^2A(z,t) \tag{1.1}$$

where z is the distance [m], t the time [s], i the imaginary unit, α the fiber attenuation and γ the nonlinear coefficient. $\beta_m=\left(\frac{d^m\beta}{d\omega^m}\right)_{\omega=\omega_0}$ $(m=1,2,..)$, being $\beta(\omega)$ the wave propagation constant and $\omega_0=2\pi f_0=\frac{2\pi c}{\lambda_0}$, with f_0, λ_0 the central frequency/wavelength of $A(z,t)$, respectively and c the speed of light.

1.1.1 Linear effects

In this section we briefly discuss the fiber linear effects, namely attenuation, group velocity and chromatic dispersion, i.e., the terms on the left-side of eq. (1.1). More details on these linear effects are given in [4].

1.1.1.1 Attenuation

The **attenuation** α [1/km] is a measure of the power loss with distance, due to absorption of the SiO_2 and Rayleigh scattering loss, coming from imperfections of the silica such as microscopic fluctuations in the material density.

Assuming all parameters zero except α, the NLSE (1.1) becomes:

$$\frac{\partial A}{\partial z}=-\frac{\alpha}{2}A$$

whose solution is $A(z,t)=A(0,t)e^{-\frac{\alpha}{2}z}$ $[\sqrt{W}]$.

If P_0 [W] is the power launched at the input of a fiber, the power of the optical signal $P(z)$ [W] decreases exponentially along the propagation, following the equation:

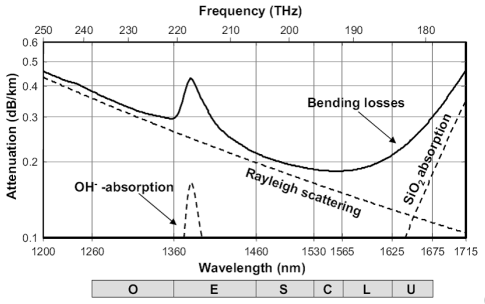
$$P(z)=P_0\,\mathrm{e}^{-\alpha z} \tag{1.2}$$

where z [km] is the distance. The attenuation coefficient is usually expressed in [dB/km], using the relation:

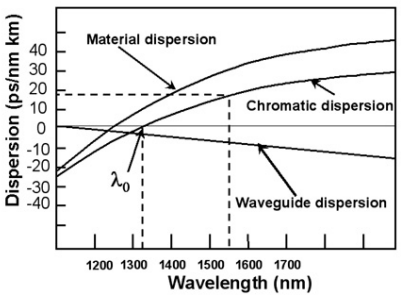
$$\alpha\,[dB/km]=10\log\left(\mathrm{e}^{\alpha[1/\mathrm{km}]}\right)\simeq 4.343\alpha\,[1/km].$$

Fiber attenuation is spectrally varying, showing a minimum around 1550 nm, as shown in Fig. 1.1(a). Most of today's transmissions use the C band, where the attenuation value is typically around 0.2 dB/km. To the attenuation one can associate the ***attenuation length***

$L_A = 1/\alpha$ as a measure of the distance over which the loss effect is significant. For a typical system having $\alpha = 0.2$ [dB/km] it is $L_A = 21.715$ [km].



(a)



(b)

Figure 1.1: (a) Fiber attenuation of a single-mode silica fiber, indicating lowest-loss regions associated with C- and L-transmission bands. Dashed curves show the contributions resulting from Rayleigh scattering, OH^- and SiO_2 absorptions. (b) Two contribution to Chromatc Dispersion: material and waveguide in a standard single-mode fiber (SMF) as a function of wavelength. λ_0 is the zero-dispersion wavelength.

1.1.1.2 Group velocity and Walk-off

$\beta_1 = \left(\frac{d\beta}{d\omega}\right)_{\omega=\omega_0} = 1/v_g(\omega_0)$ [s/m] accounts for the **group velocity** v_g of the signal along the fiber, and hence is a delay per unit length. In presence of only β_1 the NLSE writes as:

$$\frac{\partial A}{\partial z} = -\beta_1 \frac{\partial A}{\partial t}$$

whose solution is:$A(z,t) = A\left(0, t - \frac{z}{v_g}\right)$.

Two signals centered at wavelengths λ_1 and λ_2 generally have different group velocities v_{g1} and v_{g2}, respectively, and hence travel at different speed. The delay per unit length between the two signals is the ***walk-off parameter*** d_{12} equal to:

$$d_{12} = \frac{1}{v_{g1}} - \frac{1}{v_{g2}} \tag{1.3}$$

$d_{12} > 0$ means that the channel having group velocity v_{g1} travels slower than the other. The walk-off weights the impact of the cross-channel nonlinear effects, such as XPM, XPolM and FWM effects, as we will see in Ch. 4. To the walk-off one can associate the ***walk-off length*** $L_W = T_0/d_{12}$, being T_0 a reference time, generally the symbol time.

1.1.1.3 Chromatic Dispersion and Group-Velocity Dispersion

When an electromagnetic wave interacts with the fiber, the medium response depends on the optical frequency ω. This property, referred to as ***chromatic dispersion*** (CD), manifests through the frequency dependence of the refractive index $n(\omega)$. There are generally two sources of dispersion: material dispersion and waveguide dispersion. Material dispersion, which is the main contribution, comes from a frequency-dependent response of a material to waves. Waveguide dispersion occurs when the speed of a wave in the optical fiber depends on its frequency for geometric reasons, for example when the waves are confined in some region, as in the core of fiber. Fig. 1.1(b) depicts the two contributions to CD in a standard single-mode fiber (SMF) as a function of wavelength.

Mathematically, the effects of fiber dispersion are accounted for by expanding the mode-propagation constant $\beta(\omega)$ in a Taylor series about the frequency ω_0 at which the pulse spectrum is centered:

$$\beta(\omega) = n(\omega)\frac{\omega}{c} = \beta_0 + \beta_1(\omega-\omega_0) + \frac{1}{2}\beta_2(\omega-\omega_0)^2 + \frac{1}{6}\beta_3(\omega-\omega_0)^3 + ...,$$

where $\beta_m = \left(\frac{d^m\beta}{d\omega^m}\right)_{\omega=\omega_0}$ $(m = 0,1,2,....)$.

Assuming all parameters zero except $\beta_{1,2,3}$ the NLSE becomes a linear PDE and writes as:

$$\frac{\partial A}{\partial z} = -\beta_1\frac{\partial A}{\partial t} + j\frac{\beta_2}{2}\frac{\partial^2 A}{\partial t^2} + \frac{\beta_3}{6}\frac{\partial^3 A}{\partial t^3}$$

which writes in a simple form in the frequency domain $\widetilde{A}(z,\omega) = \mathscr{F}\{A(z,t)\}$:

$$\frac{\partial \widetilde{A}}{\partial z} = -j\left(\beta_1\omega + \frac{\beta_2}{2}\omega^2 + \frac{\beta_3}{6}\omega^3\right)\widetilde{A} \tag{1.4}$$

whose solution is:

$$\widetilde{A}(z,\omega) = \widetilde{A}(0,\omega)e^{-j\left(\beta_1\omega+\frac{\beta_2}{2}\omega^2+\frac{\beta_3}{6}\omega^3\right)z}. \tag{1.5}$$

Physically speaking the parameter β_2 represents the dispersion of the group velocity and is responsible for pulse broadening. This phenomenon is known as the **group-velocity dispersion** (GVD), and β_2 is the GVD parameter.

Note from (1.5) that, being the system linear, the behavior of each frequency along the fiber depends only by itself. Since in (1.4) the loss is absent, for the energy conservation principle the energy carried by frequency ω must remain unaltered, i.e., $\left|\widetilde{A}(z,\omega)\right|^2 = \left|\widetilde{A}(0,\omega)\right|^2$, so that the $\beta_{1,2,3}$ induce a pure phase rotation in the frequency domain.

Most of the times β_2 and β_3 are expressed as functions of the wavelength through the fiber ***dispersion parameter***, D (Fig. 1.1(b)), and through the ***fiber dispersion slope***, $D^{'}$. The following relations hold:

$$\begin{aligned} D &= \left(\frac{d\beta_1}{d\lambda}\right)_{\lambda=\lambda_0} & \beta_2 &= -\frac{\lambda_0^2}{2\pi c}D \\ D' &= \left(\frac{d^2\beta_1}{d\lambda^2}\right)_{\lambda=\lambda_0} & \beta_3 &= \left(\frac{\lambda_0}{2\pi c}\right)^2\left(2\lambda_0 D+\lambda_0^2 D^{'}\right), \end{aligned}$$

where $\lambda_0 = c/f_0$.

The chromatic dispersion D is usually expressed in [ps/(nm·km)] while the dispersion slope $D^{'}$ in [ps^2/(nm·km)]. $D>0$ ($\beta_2<0$) corresponds to anomalous dispersion, $D<0$ ($\beta_2>0$) corresponds to normal dispersion.

To the β_2 and β_3 parameters one can associate the ***dispersion length***, $L_D = \beta_2/T_0^2$, and the ***dispersion slope length***, $L_D^{'} = \beta_3/T_0^3$, being T_0 the symbol time. The previous lengths have signs, hence sometimes the absolute value is used.

By chromatic dispersion the various spectral components of the signal do not travel at the same speed, causing a pulse broadening in time domain. The pulses broading leads to bit-to-bit overlaps, called Inter-Symbol Interference (ISI). For example in a OOK signal, due to fiber chromatic dispersion, each "1" broadens and "0" and "1" are increasingly difficult to distinguish.

1.1.2 Nonlinear effects

The response of any dieletric medium to light becomes nonlinear for intense electromagnetic fields, and the optical fibers are no exception. The change in the refractive index of the material, in response to an applied electromagnetic field, is called Kerr effect and it depends on the optical power of the field $|A|^2$ [4, 5]:

$$\tilde{n}(\omega,|A|^2) = n(\omega)+n_2|A|^2$$

where $n(\omega)$ is the linear contribution of the refractive index and n_2 is the nonlinear-index coefficient.

The nonlinear coefficient γ [1/(W·km)] is due to the Kerr effect and its relation to the fiber nonlinear index n_2 [m^2/W] is the following:

$$\gamma = \frac{2\pi n_2}{\lambda_0 A_{\text{eff}}} \tag{1.6}$$

where A_{eff} [μm^2] is the fiber effective area. To the nonlinear coefficient one can associate the ***nonlinear length*** $L_{\text{NL}} = 1/(\gamma \cdot P)$, being P a reference power, usually the transmitted signal peak power. A direct comparison between the nonlinear length and the dispersion length allows to deduce the propagation regime inside the optical fiber. $L_{\text{NL}} \gg L_D$ implies propagation in the dispersion-limited, or purely linear, regime; on the opposite with $L_{\text{NL}} \ll L_D$ the propagation is in the nonlinear regime.

Non-linear effects are often categorized into two sets of effects: those resulting from the propagation of a single channel as Self Phase Modulation (SPM) and those resulting from the interactions between WDM channels as Cross Phase Modulation (XPM), Cross Polarization Modulation (XPolM) and Four Wave Mixing (FWM). In the following sections we briely discuss the SPM and XPM, while XPolM will be analyzed in Ch. 3. See [4] for an introduction to FWM and [6, 7] for advanced studies of FWM.

1.1.2.1 SPM

SPM refers to the self-induced phase shift experienced by an optical field during its propagation in optical fibers. Its magnitude can be obtained by noting that the phase of an optical field changes as

$$\phi = \tilde{n} k_0 L = (n + n_2 |A|^2) k_0 L$$

where $k_0 = 2\pi/\lambda_0$ and L is the fiber length. The intensity-dependent nonlinear phase shift $\phi_{NL} = n_2 k_0 L |A|^2$ is due to SPM. Among other things, SPM is responsible for spectral broadening of ultrashort pulses.

The nonlinear part of NLSE in a single channel transmission (neglecting $\beta_{1,2,3}$) is:

$$\frac{\partial A}{\partial z} = -\frac{\alpha}{2} A - j\gamma |A|^2 A$$

whose solution is:

$$A(z,t) = A(0,t) e^{-\frac{\alpha}{2} z - j\gamma |A(0,t)|^2 L_{\text{eff}}(z)} = A(0,t) e^{-\frac{\alpha}{2} z} e^{-j\Phi_{SPM}(z,t)} \tag{1.7}$$

where $\Phi_{SPM}(z,t)$ is the SPM nonlinear phase rotation, while $L_{\text{eff}}(z)$ is the ***effective length*** up to coordinate z and is equal to:

$$L_{\text{eff}}(z) = \frac{1 - \exp(-\alpha z)}{\alpha}.$$

For $z \ll 1/\alpha$ $L_{\text{eff}}(z) \simeq z$, while for $z \gg 1/\alpha$ $L_{\text{eff}}(z) \simeq L_A$. It turns out that L_{eff} is a measure of the distance over which the nonlinear effect is significant.

Note from (1.7) that the solution is memoryless so that what happens at time t depends only on the input at the same time. Assuming zero loss, for the energy conservation principle the energy carried by time t, being the system memoryless, must remain unaltered, i.e., $|A(z,t)|^2 = |A(0,t)|^2$, so that SPM is a pure phase rotation in the time domain.

1.1.2.2 XPM

In a WDM simplified scenario in which two scalar optical fields (single polarization) A_a and A_b at frequencies ω_a and ω_b respectively, propagate simultaneously inside the fiber, the nonlinear phase shift for the field at ω_a is given by

$$\phi_{NL} = n_2 k_0 L(|A_a|^2 + 2|A_b|^2), \tag{1.8}$$

where we have neglected all terms that generate power at frequencies other than ω_a and ω_b. The two intensity-dependent terms on the right-hand side of eq. (1.8) are due to SPM and XPM, respectively. XPM refers to the nonlinear phase shift of an optical field induced by another field having a different wavelength. An important feature of XPM is that, for equally intense optical fields of different wavelengths, the contribution of XPM to the nonlinear phase shift is twice that of SPM. Among other things, XPM is responsible of spectral broadening of copropagating optical pulses.

In a WDM transmission the electrical field $A(z,t)$ can be expressed as:

$$A(z,t) = \sum_{k=1}^{M} A_k(z,t) e^{j\Delta\omega_k t} \tag{1.9}$$

where $A_k(z,t)$ is the lowpass envelope of channel k, M is the number of channels and $\Delta\omega_k = \omega_k - \omega_0$ is the difference between the central frequency of channel k and the central frequency ω_0 of the WDM comb. The nonlinear part of (1.1) for $A_k(z,t)$ for $k = 1,\ldots,M$ is:

$$\sum_k \frac{\partial A_k}{\partial z} e^{j\Delta\omega_k t} = -j\gamma \sum_{n,l,m} A_n A_l A_m^* e^{j(\Delta\omega_k + \Delta\omega_l - \Delta\omega_m)t}. \tag{1.10}$$

We can rewrite the (1.10) as:

$$\frac{\partial A_k}{\partial z} = -j\gamma \sum_{\substack{n,l,m \\ \omega_n+\omega_l-\omega_m=\omega_k}}^{M} A_n A_l A_m^* \qquad k = 1,\ldots,M \tag{1.11}$$

where n, l, m can range from 1 to M but must satisfy $\omega_n + \omega_l - \omega_m = \omega_k$. Hence in (1.11) we discarded all terms of the sum falling outside the bandwidth of $A_k(z,t)$, i.e., all n, l, m

such that $\omega_n + \omega_l - \omega_m$ cannot be associated with a frequency ω_k, $k = 1, \ldots, M$ of the comb. However, such terms are usually of small energy for weakly nonlinear systems.

By varying the indexes n, l, m we can identify the following terms for SPM and XPM:

- SPM: $\omega_n = \omega_l = \omega_m = \omega_k \Rightarrow |A_k|^2 A_k$
- XPM: $(\omega_n = \omega_m) \neq (\omega_l = \omega_k)$ or $(\omega_n = \omega_k) \neq (\omega_l = \omega_m) \Rightarrow \sum_{m \neq k} |A_m|^2 A_k$

Exploiting these terms in (1.11) yields:

$$\frac{\partial A_k}{\partial z} = -j\gamma \left(|A_k|^2 A_k + 2 \sum_{m \neq k} |A_m|^2 A_k \right). \tag{1.12}$$

1.2 Coupled-NLSE (CNLSE) and Manakov-PMD equation

The coupled nonlinear Schrödinger equation (CNLSE) is the vectorial version of the scalar NLSE and it is used whenever polarization effects are to be considered. Indeed, it is used to consider polarization effects, both linear as polarization mode dispersion (PMD) (Sec. 1.2.1) and nonlinear as XPolM (Ch. 3) and to study the propagation of polarization division multiplexing PDM signals (Sec. 2.1).

Being the optical propagating field represented by a complex vector with 2 elements (Jones vector) $\overrightarrow{A}(z,\tau) = [A_x, A_y]^T$, the CNLSE, in a general form, has the following expression [8, 9]:

$$\begin{array}{cccccc} \frac{\partial \overrightarrow{A}}{\partial z} + & \frac{\alpha}{2}\overrightarrow{A} - & \frac{i}{2}\beta_2 \frac{\partial^2 \overrightarrow{A}}{\partial \tau^2} + & i\frac{\Delta\beta_0(z)}{2}\left(\vec{l}(z)\cdot\vec{\sigma}\right)\overrightarrow{A} + \frac{\Delta\beta_1(z)}{2}\left(\vec{l}(z)\cdot\vec{\sigma}\right)\frac{\partial \overrightarrow{A}}{\partial \tau} & = \\ & & = & -i\gamma\left[\left|\overrightarrow{A}\right|^2 \overrightarrow{A} - \frac{1}{3}\left(\overrightarrow{A}^{\dagger}\sigma_3\overrightarrow{A}\right)\sigma_3\overrightarrow{A}\right] & \end{array} \tag{1.13}$$

where $\tau = t - \frac{z}{v_g}$ is the retarded time frame, discussed in Sec. 1.1.1.2, while † means transpose-conjugate. Comparing (1.13) with the *scalar* NLSE (1.1), two new terms (those with $\Delta\beta_{0,1}$) appear in the linear part, while the nonlinear part of the equation is modified by an extra term (that with σ_3).

The linear terms are due to the fiber ***Birefringence***, i.e., the dependence of the refractive index on the signal polarization (Sec. 1.2.1). $\Delta\beta_0(z)$ and $\vec{l}(z)$ are the birefringence strength and orientation in Stokes space and $\Delta\beta_1(z)$ is the differential group delay (DGD) or ***first order PMD***.

The symbol $\vec{\sigma}$, appearing in (1.13), is the so-called spin-vector, whose elements are the three Pauli matrices, shown in App. A (hence, $\vec{\sigma}$ is actually a tensor). The scalar product $(\hat{l}(z)\cdot\vec{\sigma})$ yields a unitary Jones matrix, i.e., a 2×2 complex matrix with unit determinant. Of the three Pauli matrices, only the third (σ_3) appears in the nonlinear term in (1.13). Despite many different (but equivalent) expressions are possible for the nonlinear term in (1.13), the concept is that the circular component of the signal polarization (associated with the third Stokes component, whose mathematical expression is $\left(\overrightarrow{A}^{\dagger}\sigma_3\overrightarrow{A}\right)$) plays a special role in the CNLSE. This peculiarity is not always remarked in the literature, since another alternative and simplified form of the CNLSE is implemented for its numerical solution, namely the Manakov-PDM equation (1.14).

1.2.1 Birefringence and Polarization Mode Dispersion

Polarization mode dispersion (PMD) [10] arises from the birefringence of the fiber, i.e., the dependence of the refractive index on the signal polarization. In an ideal optical fiber, in which the core has a perfectly circular cross-section, the fundamental mode has two orthogonal polarizations that travel at the same speed. In a realistic fiber, however, there are random imperfections that break the circular symmetry, causing the two polarizations to propagate at different speeds. The symmetry-breaking random imperfections, e.g., slightly elliptical cores, can stem from either imperfections in manufacturing or from thermal and mechanical stresses imposed on the fiber in the field, that generally vary over time.

Therefore, birefringence is stochastic and fluctuates both in time (slowly compared with the symbol period, hence it is assumed constant in τ in the CNLSE) and along fiber length, depending on the characteristics of the fiber and on local condition, such as temperature. At each position z, the fiber is characterized by an *eigenmode* $\hat{l}(z)$, corresponding to the field polarization with slowest propagation constant $\beta_s(\omega)$ and, at the same time, the orthogonal eigenmode $\hat{l}_o(z)$ is the field polarization with fastest propagation constant $\beta_f(\omega)$. The difference $\Delta\beta(\omega)=\beta_s(\omega)-\beta_f(\omega)$ between these propagation constants is the strength of the birefringence, while $\hat{l}(z)$ represents the birefringence orientation [11]. It is generally believed that silica fibers are characterized by linear birefringence [12]: this implies that the third component of $\hat{l}(z)$ is null.

In (1.13), birefringence is modeled by a Stokes vector $\vec{W}(z,\omega)=(\Delta\beta_0+\Delta\beta_1\omega)\hat{l}(z)$ with a linear frequency dependence, although other models are possible. The variations in z of its orientation cause "random mode coupling", i.e., the exchange of energy between the field components parallel or perpendicular to $\hat{l}(z)$, eventually leading to PMD. If the frequency dependent term $\Delta\beta_1$ is set to zero, there is no pulse distortion, and the overall result of bire-

fringence is just a rotation of the signal State Of Polarization (SOP) on the Poincaré sphere (App. A). If $\hat{l}$ is constant along z, the fiber is called Polarization Maintaining Fiber (PMF): its input-output behavior amounts to splitting each input pulse into two "shadow pulses" arriving at the fiber output with a mutual delay equal to $\Delta\beta_1 z$. This effect, known as first-order PMD, causes pulse broadening and consequently intersymbol interference. In the more general case, mode coupling produces PMD at all orders, hence a pulse is not only split in two (first-order PMD), but each of the shadow pulses suffers a different amount of linear distortion, including GVD, that differently affects the polarized components of the signal.

The PMD value of a fiber is defined as the the root-mean-square (RMS) value of the DGD (Δt in [ps]), i.e., the time separation between the two principal polarization modes of the transmission link at the receiver. Since DGD is an instantaneous event and varies randomly with wavelength and time, the PMD is the average value of a distribution of a large number of independent DGD measurements.

For a fiber of length L, Δt is given by

$$\Delta t = \left| \frac{L}{v_{gs}} - \frac{L}{v_{gf}} \right| = L \left| \beta_{1,s} - \beta_{1,f} \right| = L \left| \Delta\beta_1 \right|$$

where $1/\beta_{1,s}$ and $1/\beta_{1,f}$ are the group velocities of the slow and fast modes, respectively.

Contrary to chromatic dispersion, polarization mode dispersion changes quickly with time (second to millisecond range) [13]. Furthermore the impact of PMD evolves linearly with the symbol duration and its constraints are challenging at 40 Gb/s and 100 Gb/s.

Some techniques have been proposed to mitigate this impairment, such as replacing fibers having poor PMD properties (very expensive) or to electrically regenerate the optical signal. One of the most effective techniques is to select a modulation format or a detection method which is more tolerant to PMD impairments. Several works demonstrated that the association of multilevel modulation format with coherent detection and digital signal processing can be extremely effective for compensating PMD distortions of high speed signals [14, 15, 16].

1.2.2 Manakov-PDM equation

The birefringence term with $\Delta\beta_0$ is purely imaginary and causes a differential phase rotation in the signal components (hence a change of its state of polarization). Such a phase rotation is frequency-independent and does not cause signal distortion, but affects the nonlinear term in the CNLSE. Hence, the numerical integration of (1.13) requires choosing a sufficiently small step (in z). An important parameter of transmission fibers is the ***beat length*** $L_B = \frac{2\pi}{\Delta\beta_0}$: one must then choose an integration step Δz such that the phase rotation $\Delta\beta_0 \Delta z$ is small compared

with 2π. Since L_B is typically of the order of meters (or tens of meters), for standard fibers, the integration of (1.13) is extremely time-consuming.

An alternative approach, requiring much smaller computation times, is that of averaging the impact of signal polarization over the nonlinear term in (1.13). If L_B is small enough, the rapid variations of the state of polarization of $\overrightarrow{A}$ are such that the term $\frac{1}{3}\left(\overrightarrow{A}^{\dagger}\sigma_3\overrightarrow{A}\right)\sigma_3\overrightarrow{A}$ undergoes a *complete mixing* and reduces to $\frac{1}{9}\left|\overrightarrow{A}\right|^2\overrightarrow{A}$, on average.

Over sufficiently long distances to ensure complete mixing (averaging the nonlinear effects over the random polarization changes that uniformly cover the Poincaré sphere), the CNLSE can be then simplified to

$$\frac{\partial\overrightarrow{A}}{\partial z}+\frac{\alpha}{2}\overrightarrow{A}-\frac{i}{2}\beta_2\frac{\partial^2\overrightarrow{A}}{\partial\tau^2}+i\frac{\Delta\beta_0(z)}{2}(\hat{l}(z)\cdot\vec{\sigma})\overrightarrow{A}+\frac{\Delta\beta_1(z)}{2}(\hat{l}(z)\cdot\vec{\sigma})\frac{\partial\overrightarrow{A}}{\partial\tau}=-i\gamma\frac{8}{9}\left|\overrightarrow{A}\right|^2\overrightarrow{A} \tag{1.14}$$

which is known as the ***Manakov-PMD equation*** [8]. The Manakov-PMD equation is generally regarded as a simple and reliable way to model optical fibers affected both by Kerr effects and PMD. Note that although the birefringence term (that with $\Delta\beta_0$) is still there, there is no need for integration steps smaller than L_B. In other terms, the Manakov equation amounts to the propagation in a PMF fiber (in a frame of reference aligned with the birefringence axes) in which the nonlinear term is reduced by a factor 8/9. The factor of 8/9 in the nonlinear coefficient has been verified experimentally [17, 18].

1.2.3 Numerical solution: SSFM

The ***split-step Fourier Method*** (SSFM) is an efficient algorithm for the numerical solution of the NLSE. It is a special application of the splitting method for solving a PDE. Generally speaking, the method is useful solve out the problem [19]:

$$\frac{\partial A(z,t)}{\partial z}=\mathscr{D}A(z,t) \tag{1.15}$$

where $\mathscr{D}$ is a differential operator that can be written in the form $\mathscr{D}=\mathscr{L}+\mathscr{N}$, being $\mathscr{L}$ and $\mathscr{N}$ differential operators as well, such that:

$$\begin{aligned}\frac{\partial A}{\partial z} &= \mathscr{L}A\\ \frac{\partial A}{\partial z} &= \mathscr{N}A\end{aligned}$$

are two easy to solve differential equations. This is the case of the NLSE where closed form solutions exist with only dispersion or with only nonlinearity. The SSFM consists therefore of subdividing the fiber in small steps and by applying at each step the two operators separately (splitting). Since in the NLSE the operators $\mathscr{L}$ and $\mathscr{N}$ do not commute, i.e., $\mathscr{L}\mathscr{N} \neq \mathscr{N}\mathscr{L}$, applying separately the two operators leaves an error. Such an error decreases quadratically for decreasing z. Hence, if the steps are sufficiently small, the local truncation error is "small" as well and hopefully $A(L,t)$ is very close to the exact solution. The linear operator L is efficiently evaluated in the frequency domain while the nonlinear operator N in the time domain. Such approach calls for the FFT and IFFT algorithms for switching between the two domains efficiently.

The choice of the SSFM step size is a hard task for which it is difficult to give a universal answer, suitable for any optical system. The most accurate approach for choosing the SSFM step size is to run many simulations for decreasing step sizes until some convergence is observed. In App. D.1 we will focus on the nonlinear steps choice, whose length has to be sufficiently long to take in account the walk-off between the two edge channels in a WDM comb.

Chapter 2

Techniques to increase transmission capacity

In the communication systems, the process of sharing a common physical channel by several transmitted signals is called multiplexing and it is used to increase the total transmission system capacity. An optical signal presents different orthogonality to be exploited for multiplexing operations:

- Orthogonality in time is used to generate time division multiplexed (**TDM**) signal where the same channel is used from different users within different time slots.
- Orthogonality in wavelength introduced the concept of wavelength division multiplexing (**WDM**) where different users use the same channel at the same time but exploiting disjoint spectral regions.
- Orthogonality in polarization typical means for optical systems the possibility to transmit different data streams in the same WDM frequency band but with orthogonal polarizations, polarization division multiplexing (**PDM**).
- Orthogonality in mode introduce the concept of mode division multiplexing (**MDM**), thanks to development of multi-mode fiber with very low loss.

The optical transmission systems performance (system capacity and system reach) drastically improved in the last years by using new technologies such as wide band optical amplification and wavelength division multiplexing (WDM).

The recent development of digital signal processing (DSP)-based coherent detection in optical networks expedited the use of polarization division multiplexing (PDM) as a cost-effective way of doubling system capacity. Thanks to a high spectral efficiency and a remarkable resilience against Polarization Mode Dispersion (PMD) and Group Velocity Dispersion (GVD), the joint use of PDM and coherent detection offers the advantage of increasing the total system capacity without incurring performance degradation due to a wider spectrum.

Furthermore, very recent transmission experiments have shown the possibility of further increasing the transmission capacity by using a new technique called mode division multiplexing (MDM), which can be used jointly with PDM and coherent detection.

This chapter briefly discusses the PDM technique, showing a PDM transmitter scheme (Sec. 2.1), the coherent detection, explaining the advantages and the receiver functionality (Sec. 2.2) and finally, the MDM technique, discussing the key components needed to this approach, namely multi-mode fiber, mode converter and spatial multiplexer/demultiplexer.

2.1 Polarization Division Multiplexing

Polarization division multiplexing (PDM) consists of encoding one independent tributary in each of the two orthogonal polarizations of the optical field. Since both polarizations are transmitted simultaneously, PDM allows doubling the spectral efficiency. For example PDM-QPSK achieves a spectral efficiency of 4 bit/s/Hz per wavelength, because it transmits two QPSK signals at the same time over the same spectral components.

Fig. 2.1 illustrates the PDM transmitter scheme.

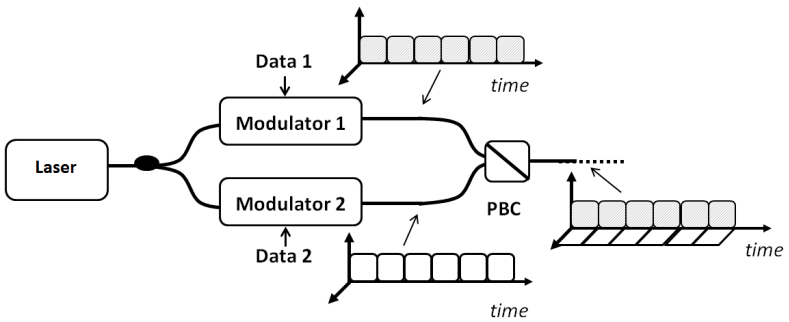


Figure 2.1: PDM transmitter scheme.

The light from the CW laser is split into two copies through a 3 dB coupler. Each copy is

sent into a distinct modulator driven with different electrical data. Finally, a polarization beam combiner (PBC) recombines the output of both modulators onto the two polarizations of the PDM channel. Polarization tributaries are commonly referred to as either parallel/orthogonal or X/Y tributaries. All the components used in the PDM generation process must be polarization maintaining to ensure correct operation.

PDM can be combined with different modulation formats such as BPSK [20], QPSK [15] or 16QAM [21]. In this thesis we focused on two of these options, namely PDM-QPSK and PDM-BPSK, whose performance will be evaluated in ch. 4.

The bit rate of each polarization tributary is half the total bit rate of the PDM signal. Hence, the symbol rate of PDM signals is half the symbol rate of singly polarized signals for a given total bit rate. Consequently, the required bandwidth of opto-electronic devices as well as the drive frequency of modulators is also halved with respect to single polarization transmitters, which is a remarkable property of PDM transmitters with respect to single polarization transmitters. On the other hand, PDM generation requires almost twice as many components.

The optical bandwidth of PDM signals is halved with respect to signals employing only one polarization (at the same bit rate). This fact enhances the tolerance of PDM signals against linear effects, namely chromatic dispersion, PMD and narrow filtering, since the tolerance to these effects decreases with the spectral width, and hence with the symbol rate [22]. Furthermore, narrow signal spectra is important in the context of dense WDM transmission systems.

From another point of view, when operating at the same symbol rate, PDM allows doubling the bit rate while keeping almost unchanged the tolerance to linear impairments [23]. On the contrary, tolerance against nonlinearities decreases, in general, with the application of PDM due to nonlinear interactions between both polarizations [24] and to the lower symbol rate which makes signals more sensitive against cross nonlinearities [25], that we will evaluate in Ch. 4.

The OSNR requirements for a PDM signal is 3 dB higher than a single polarization signal (at same symbol rate), this because to total signal power is divided in two orthogonal components that carry different information.

The major downside of PDM transmission systems is the receiver complexity. Nevertheless, recent advances in electronics allow robust reception thanks to coherent detection (Sec. 2.2) and advanced signal processing. Paired with coherent detection and digital signal processing, PDM opens a wide range of possibilities to increase the robustness against linear effects and the spectral efficiency of optical transmission systems.

2.2 Coherent Detection

Coherent detection relies on detecting a signal through the beating with a reference frequency carrier, commonly supplied by a local oscillator (LO) much more powerful than the signal.

This technology has been investigated at the end of the 1980s in order to achieve better receiver sensitivity and longer unregenerated distance [26, 27, 28]. Indeed, by beating with the LO laser to enhance the signal, the receiver sensitivity can be improved by up to 20 dB compared to direct detection without pre-amplification [29]. In some sense, the mixing with LO laser serves as a signal amplifier.

With the introduction of the Erbium Doped Fiber Amplifier (EDFA) at the beginning of the 1990s, longer unregenerated distance can be achieved by periodically amplifying the optical signal and better sensitivity can be achieved by pre-amplifying the received signal. Furthermore compared with coherent detection, direct detection in conjunction with EDFA pre-amplification gives very similar performance without most of the technical issues of coherent detection. For this reason, further research activities in coherent optical communications have almost been interrupted for twenty years.

Coherent technologies have restarted to attract a large interest over the recent years [30]. The motivation lies in finding methods of meeting the ever-increasing bandwidth demand with multi-level modulation formats based on coherent technologies.

Furthermore, the recent development of high-speed digital electronics for signal processing and, especially, of analog-to-digital converters (ADC), has offered the possibility of treating the electrical signal in a digital signal processing (DSP) core.

The electrical treatment of the signals opened the possibility of using powerful algorithms able to compensate for signal distortions incurred along the fiber-optic transmission, mainly chromatic dispersion and PMD [14, 31]. It should be stressed that although PMD is time varying, the digital coherent receiver can equalize it in an adaptive manner.

Thus, today's coherent receivers enjoy the high sensitivity of coherent detection with the added benefits of digital signal processing and are a very likely the solution for the next generation high bit-rate transponders.

2.2.1 Coherent receiver description

In the coherent receiver (Fig. 2.2), the optical signal is first mixed with the light of a local oscillator (LO) laser to down-convert the signal from the optical carrier frequency to microwave carrier frequency in the range of GHz or tens of GHz. When the received signal is mixed with LO laser, an optical beat signal is generated at the photodiode a frequency equal to an inter-

mediate frequency (ω_{IF}) that is the frequency difference between the received signal and the LO laser. If the optical frequency of the signal is the same as that of the LO laser, the system is called **homodyne**. If the optical frequency of the signal differs with that of the LO laser, the system is called **heterodyne** with an ω_{IF} of:

$$\omega_{IF} = \omega_s - \omega_{ol}$$

where ω_s and ω_{ol} are the optical frequency of the carrier signal and LO laser respectively. In homodyne systems, $\omega_{IF} = 0$.

Homodyne reception requires the local-oscillator frequency to be strictly locked in frequency and phase to the received signal and gives optimal receiver sensitivity [29]. The main constraint is therefore that the frequency and phase of the local oscillator must be controlled and adjusted continuously, which is traditionally done with a phase-locked loop (PLL).

Heterodyne reception presents the advantage of relaxing the constraints on the linewidth of the lasers, which makes it easier to implement than homodyne detection. Nonetheless, heterodyne detection requires a receiver bandwidth at least twice the symbol rate [32], i.e., double compared to homodyne detection. The implementation for high bit rate operation is therefore challenging. Moreover, the sensitivity of a heterodyne detector is at least 3 dB worse compared to homodyne detection, since the effective energy of a heterodyne-detected signal is half of the signal effective energy with homodyne detection [33, 34].

When the frequency of the local oscillator is *approximately* the same as the frequency of the signal, the system is called **intradyne**. The frequency mismatch between the signal and the local oscillator is recovered in the digital domain by processing the baseband signal so as there is no need of adjusting the local-oscillator frequency and phase. The maximum tolerable frequency-difference is therefore determined by the signal processing. Besides, a phase-diversity receiver is essential as the local oscillator is not phase-locked. Intradyne detection aims to detect the in-phase and quadrature components of the signal and therefore a 90º hybrid is required. Both components of the (baseband) optical field can be transferred to the electrical domain by analog-to-digital converters (ADC); thus, the same receiver can operate with any kind of optical modulation format. Moreover, any phase-drift can be compensated through digital signal processing.

Moreover, polarization-diversity receivers offer the potential of detecting polarization division multiplexed signals without any additional component. Thus polarization-diversity receivers can be considered as multiple-input multiple-output receivers (MIMO). Okoshi et al. were also the firsts, in 1987, to demonstrate a receiver combining intradyne detection with polarization diversity in which all the characteristics of the optical field (without polarization

multiplexing) were translated to the electrical domain [35].

Today's coherent receivers, use on one hand, intradyne polarization-diversity detection to convert the full optical field (i.e., amplitude, phase and polarization) to the electrical domain [28] and on the other hand advanced algorithms to compensate for transmission impairments [14, 31]. This requires the detection and digitalization of four signals, i.e., the in-phase (I) and quadrature (Q) components for two arbitrary, but orthogonal, polarization states (Fig. 2.2).

2.2.1.1 Coherent receiver scheme

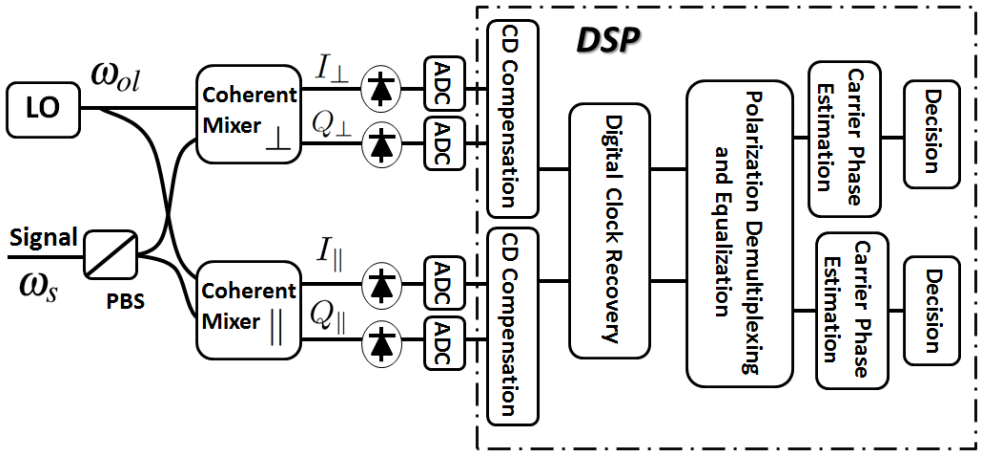


Figure 2.2: Coherent Receiver scheme.

In order to recover the full electrical field of the signal, polarization diversity scheme is used as shown in Fig. 2.2. The signal is first split in two by a PBS which sends each one of the incoming polarization, into two coherent mixers. The four optical interference signals are then sampled, and digitalized by Analog to Digital Converters (ADC). In order to respect the Shannon–Nyquist criteria, the sampling rate of the ADCs has to be at least twice as large as the largest frequency of the signal. In practice, the ADC 3-dB bandwidth is roughly 0.5 to 0.8 times the symbol rate and ADC sampling rate is twice the symbol rate. The two electrical fields are then reconstructed by digital signal processing (DSP) and symbols are then identified by a simple threshold method. The DSP is done in several steps as described in Fig. 2.2.

First, chromatic dispersion (CD) can be compensated, either by using Finite Impulse Response (FIR) filters, or by using Fast Fourier Transform (FFT) method as described in [36].

These filters may be very long (150 taps or more depending on the amount of chromatic dispersion to be compensated for) but are static and do not required fast adaptive algorithms. A digital clock recovery is then required for the other parts of the processing. A key part of the DSP is to demultiplex the two initial signals sent along two orthogonal polarizations, and to equalize simultaneously the two signals. This can be done by using Constant Modulus Algorithm (CMA) as proposed in [37]. The filters used within this part, have to adapt themselves continuously to the incoming signal, to follow polarization fluctuation and PMD variations [31]. The last main part of the digital signal processor is the Carrier Phase Estimation (CPE) process. This process is required to recover and cancel the frequency offset $(\omega_s - \omega_{ol})$ between the local oscillator and the carrier frequency of the signal as described in [37, 38].

More details about the coherent receiver and DPS algorithms are in [14, 31, 39].

2.3 Mode Division Multiplexing

Increasing the constellation complexity has been the most frequently-reported approach to increase the transmission capacity in the past few years, but it can have a detrimental impact on reach, e.g. a reduction by a factor of five from 100Gb/s Polarization Division Multiplexed (PDM) Quaternary Phase-Shift Keying (QPSK) to 200Gb/s PDM-16-QAM (Quadrature Amplitude Modulation) [40].

The mode division multiplexing (MDM), which consists in the transmission of independent signals over the modes that can propagate in a multi mode fiber (MMF), represents a new approaches for increasing the transported capacity of an optical fiber without compromising the ultimate reach of the transmission systems. Recent transmission experiments have shown the possibility to employ few-mode fibers (FMF) and mode division multiplexing (MDM) [41, 42, 43, 44]. In these reports, two different kinds of FMFs are used, either with a large differential mode group delay (DMGD) between the modes [41, 42], or a DMGD close to zero [43]. As a rule of thumb, very large DMGDs prevent the system from mode coupling during the propagation and in the absence of other sources of inter-modal cross-talk, the system would not suffer from interference between the modes. However, compensating potential inter-modal cross-talk by joint digital signal processing (DSP) of several modes using multiple-input multiple-output (MIMO) techniques is challenging in this case because of the huge mode delays. On the contrary, small DMGDs increase the risk of distributed mode coupling along the line so that the modes have to be separated with MIMO techniques in the DSP. But at the same time, this is more practical here thanks to the small mode delays at the

receiver [43].

The pioneering experiments previously cited have employed different techniques for mode multiplexing and demultiplexing: long period fiber Bragg gratings [41], programmable or fixed phase masks [42, 43], or a prototype mode coupler [44]. At present, it is very difficult to understand if one technology for mode conversion and multiplexing is better suited than the others for satisfying the needs of high capacity MDM transmission systems. Practical implementations would require quantifying the limits of each of these technologies. Systems using FMFs with weak mode coupling would suffer from inter-modal cross-talk generated by the finite mode rejection ratio of the mode converters in the mode-multiplexer and –demultiplexer rather than from the one generated during the propagation.

In Ch. 6 we will summarize the experimental work made at Alcatel-Lucent Bell-Labs France on this topic. For our transmission experiments, we employed a LCOS-based mode converter spatial modulator and a prototype FMF [45]. This LCOS-based approach to mode conversion is attractive because of the possibility to reconfigure the phase plates to any desired mode conversion. The FMF has the advantage of exhibiting very large effective-index differences and very large group delays between different modes and thus low linear crosstalk between modes, with only 0.22dB/km loss.

Chapter 3

Cross-Polarization Modulation

The solution of the Manakov-PDM equation with multiple propagating fields is of paramount importance in analyzing WDM systems affected by both linear impairments (scattering loss, group velocity dispersion, GVD, polarization mode dispersion (PMD) and the Kerr effects[1]. It is becoming more and more interesting, especially in polarization-division multiplexed (PDM) systems, to identify the different contributions entailed in the Kerr term of the Manakov-PDM equation, so as to understand their relative impact on transmission using different modulation formats: e.g., to answer questions like "in a system using PDM-quadrature phase shift keying (QPSK) modulation, is it more impairing to have a strong cross phase modulation or to suffer nonlinear polarization rotation due to neighboring channels?". Hence, several papers in the literature address the phenomenon of cross-polarization modulation (XPolM) when referring to the nonlinear change of the state of polarization (SOP) suffered by a propagating channel due to the nonlinear interaction with neighboring channels. However, to our knowledge, there is not a common consensus of what term of the Manakov-PDM equation is to be called XPolM, while the identification of the XPM term (and of the SPM term) is well assessed, although it is sometimes questioned by recognizing that the amount of nonlinear cross-phase rotation depends on the SOP of the interfering channels.

3.1 Definition of XPolM

In this section we define the *XPolM* term and we analyze several case studies that show its impact on the received symbols, both theoretically and through simulations.

[1] We neglect here higher order non-linear effects, like self-steepening etc.

We rewrite the nonlinear term of the Manakov equation (1.14) expanded as in the case of multiple channels, labeling the terms that produce SPM, XPM and XPolM[2]

$$\frac{d\overrightarrow{A_n}}{dz} = -i\frac{8}{9}\gamma e^{-\alpha z}\left\{\underbrace{|A_n|^2\overrightarrow{A_n}}_{\text{SPM}} + \underbrace{\frac{3}{2}\sum_{k\neq n}|A_k|^2\overrightarrow{A_n}}_{\text{XPM}} + \underbrace{\frac{1}{2}\sum_{k\neq n}(\overrightarrow{a_k}\cdot\vec{\sigma})\overrightarrow{A_n}}_{\text{XPolM}}\right\} \tag{3.1}$$

where the term $e^{-\alpha z}$ accounts for the fiber scattering loss. In (3.1), we neglected the linear terms, thus the above equation represents a simplified model, that does not account for the different group velocities of the channels (i.e., the walk-off), assuming that the bit-patterns of the interfering channels are aligned in time. Chromatic dispersion and linear birefringence are also not considered in (3.1), since we wish here to concentrate our attention on the Kerr term. In (3.1), we expanded $(\overrightarrow{a_k}\cdot\vec{\sigma})$ by expliciting the elements of the Stokes field $\overrightarrow{a_k} = [a_{k1}; a_{k2}; a_{k3}]$, with magnitude $|A_k|^2 = \|\overrightarrow{a_k}\|$.

In (3.1), there are three terms, labeled *SPM*, *XPM*, *XPolM*, of which the first two produce the well-known scalar self- and cross-phase modulation impairments. We wish here to concentrate on the *XPolM* term, highlighting the impact of this polarization dependent term alone on the integration of (3.1). Neglecting SPM and XPM, eq. (3.1) becomes

$$\frac{d\overrightarrow{A_n}}{dz} = -i\frac{8}{9}\gamma e^{-\alpha z}\frac{1}{2}\sum_{k\neq n}\left(\overrightarrow{a_k}\cdot\overrightarrow{\sigma}\right)\overrightarrow{A_n} \tag{3.2}$$

which could be integrated in closed form if the direction $\sum_{k\neq n}\overrightarrow{a_k}$ did not vary in z. However $\sum_{k\neq n}\overrightarrow{a_k}$ is not constant along z, hence we manipulate eq. (3.2) by rewriting it as

$$\begin{aligned}\frac{d\overrightarrow{A_n}}{dz} &= -i\frac{8}{9}\gamma e^{-\alpha z}\frac{1}{2}\left[\sum_{k\neq n}\left(\overrightarrow{a_k}\cdot\overrightarrow{\sigma}\right)\overrightarrow{A_n} \pm \left(\overrightarrow{a_k}\cdot\overrightarrow{\sigma}\right)\overrightarrow{A_n}\right] \\ &= -i\frac{8}{9}\gamma e^{-\alpha z}\frac{1}{2}\left[\left(\overrightarrow{a}\cdot\overrightarrow{\sigma}\right)\overrightarrow{A_n} - |A_n|^2\overrightarrow{A_n}\right]\end{aligned} \tag{3.3}$$

being $(\overrightarrow{a_n}\cdot\vec{\sigma})\overrightarrow{A_n} = a_0\overrightarrow{A_n} = |A_n|^2\overrightarrow{A_n}$ and defining the Pivot vector as $\overrightarrow{a} = \sum_k\overrightarrow{a_k} = a_0\widehat{a}$ (the summation extends to all channels, including $\vec{a}_n$). Since the Pivot does not vary[2] in z, eq. (3.3) can now be integrated. If the sum of the interfering channels' power is much larger than

[2] In the absence of walk-off, the Pivot is constant along the fiber, and the Stokes vector $\vec{a}_k$ of each channel rotates around it in Stokes space [46]. We shall refer to this picture and assume that, at least in a nonlinear length L_{NL}, GVD is negligible and a "local" pivot vector $\vec{a}(z)$ exists, which is constant within the nonlinear length.

the power of the channel under test ($\vec{a}_n$), we can assume that $\vec{a}_n$ does not influence the Pivot orientation. In this case eq. (3.3) is particularly simple to deal with the case where all the interfering channels are aligned with a specific SOP $\widehat{a}$[3]. In this case the pivot $\overrightarrow{a} = \left(\sum_k |A_k|^2\right)\widehat{a}$ has the same z-invariant SOP as the interfering channels. We can integrate the above equation by defining the effective length $L_{eff}(\Delta z) = (1-\exp\{-\alpha\Delta z\})\alpha^{-1}$ of the integration step. The resulting field, at $z+\Delta z$, is

$$\overrightarrow{A_n}(z+\Delta z) = \mathrm{e}^{-\mathrm{i}\frac{8}{9}\gamma \mathrm{L}_{\mathrm{eff}}(\Delta z)|\mathrm{A_n}|^2}\mathrm{e}^{-\frac{\mathrm{i}\frac{8}{9}\gamma \mathrm{L}_{\mathrm{eff}}(\Delta z)}{2}\left[\overrightarrow{a}\cdot\overrightarrow{\sigma}\right]}\overrightarrow{A_n}(\mathrm{z})\,. \tag{3.4}$$

Eq. (3.4) it the solution of (3.1) for the sole XPolM effect [47].

3.2 XPolM impact on the received symbols – Theory

Neglecting the scalar phasor multiplier (3.4) angle rotation, which does not induce any SOP rotation and defining a nonlinear rotation angle $\theta = \frac{8}{9}\gamma L_{eff}(\Delta z)\frac{1}{2}a_0$, we can rewrite eq. (3.4) as[4]

$$\overrightarrow{A_n}(z+\Delta z) = \mathrm{e}^{-\mathrm{i}\theta\left[\hat{\mathrm{a}}\cdot\overrightarrow{\sigma}\right]}\overrightarrow{A_n}(\mathrm{z}) = \left\{\cos(\theta)\sigma_0 - \mathrm{i}\sin(\theta)\left[\hat{\mathrm{a}}\cdot\overrightarrow{\sigma}\right]\right\}\overrightarrow{A_n}(\mathrm{z})\,. \tag{3.5}$$

Using the technique in Appendix C.1, the unitary transformation (3.5) is mapped in Stokes space as [11]

$$\overrightarrow{a_n}(z+\Delta z) = \mathrm{e}^{+2\theta[\hat{\mathrm{a}}\times]}\overrightarrow{a_n}(\mathrm{z}) \tag{3.6}$$

whose geometrical interpretation is that $\overrightarrow{a_n}(z)$ is subject to a positive (i.e., counter-clockwise) rotation around $\widehat{a}$, by an angle 2θ. The cross-product operator $[\widehat{a}\times]$ should be interpreted as the skew-symmetric matrix defined in (vi), Appendix C.2. We wish now to better understand the nonlinear polarization-dependent distortions suffered by the field $\overrightarrow{A_n}$ (in Jones space), whose elements carry the QPSK constellations. Let us now assume, without loss of generality, that the PDM-QPSK signal $\overrightarrow{A_n}$ employs the canonical X and Y polarizations (i.e., linear horizontal and vertical polarizations, $\pm\widehat{S_1}$, in Stokes space) to multiplex the PDM subcarriers. Representing the unit magnitude pivot SOP

$$\widehat{a} = [\cos(2\chi); \sin(2\chi)\cos(\phi); \sin(2\chi)\sin(\phi)]$$

[3] In the following sections, we will show numerical results obtained both under this hypothesis and with randomly variable SOPs.

[4] From Appendix C.1 we obtained, for a matrix $W = \mathrm{e}^{\mathrm{D}}$ where $D = -i\theta\left[\widehat{a}\cdot\overrightarrow{\sigma}\right]$, whose eigenvalues are $\pm i\theta$:

$$\begin{cases}\alpha_0 = \frac{-i\theta e^{i\theta} - (i\theta)e^{(-i\theta)}}{-i2\theta} = \cos(\theta) \\ \alpha_1 = \frac{e^{-i\theta} - e^{i\theta}}{-i2\theta} = \frac{1}{\theta}\sin(\theta)\end{cases} \Rightarrow W = \alpha_0 D^0 + \alpha_1 D^1 = \cos(\theta)\sigma_0 + \frac{1}{\theta}\sin(\theta) - i\theta(\widehat{a}\cdot\overrightarrow{\sigma})$$

by the two angular coordinates $(2\chi,\phi)$ ($2\chi \in [0;\pi]$, $\phi \in [-\pi;\pi]$) (see Appendix A) it is then useful to express the matrix in (3.5) as the sum of two contributions[5]

$$\begin{aligned}\left[\widehat{a}\cdot\overrightarrow{\sigma}\right] &= \cos(2\chi)\sigma_1+\sin(2\chi)[\cos(\phi)\sigma_2+\sin(\phi)\sigma_3] \\ &= \cos(2\chi)\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}+\sin(2\chi)\begin{bmatrix}0 & e^{-i\phi}\\ e^{+i\phi} & 0\end{bmatrix}\end{aligned} \tag{3.7}$$

which have conceptually different implications. The first term induces fading while the second term induces subcarrier cross-talk, as will be discussed hereafter.

We can express eq. (3.5), resorting to the definition of θ and $\left[\widehat{a}\cdot\overrightarrow{\sigma}\right]$ given above, as

$$\overrightarrow{A_n}(z+\Delta z)=\begin{bmatrix}\cos(\theta)-i\sin(\theta)\cos(2\chi) & -i\sin(\theta)\sin(2\chi)e^{-i\phi}\\ -i\sin(\theta)\sin(2\chi)e^{+i\phi} & \cos(\theta)+i\sin(\theta)\cos(2\chi)\end{bmatrix}\overrightarrow{A_n}(z) \tag{3.8}$$

which is valid for any value of $2\chi,\phi$ and which highligths the terms that induce fading and cross-talk as a function of the pivot orientation $\widehat{a}$.

The general rotation matrix expression in (3.8) clearly highlights that there is a cross-talk between the PDM sub-channels $A_n^{(x,y)}(z)$, such that the symbol constellations at the receiver are linear combinations of the originally transmitted QPSK constellations, as shown in Fig. 3.1.

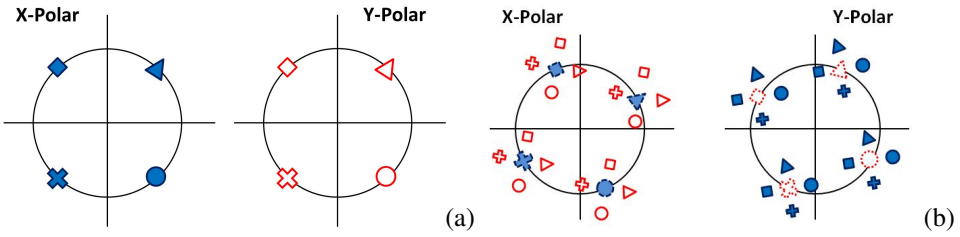


Figure 3.1: X-polar and Y-polar at the transmitter side (a) and after propagating in presence of Pivot oriented along a general SOP with angular coordinates $(2\chi,\phi)$ (b). Dotted symbols in (b) represent the original symbols' positions in (a), reported as a reference.

Eq. (3.8) which coincides with (3.5), can be explicitly written in terms of the orthogonal

[5]Remember that: $\left[\widehat{a}\cdot\overrightarrow{\sigma}\right]=\begin{bmatrix}a_1 & a_2-ia_3\\ a_2+ia_3 & -a_1\end{bmatrix}$, where $\widehat{a}=[a_1;a_2;a_3]$

X,Y polarization components of the field

$$\begin{bmatrix} A_n^{(x)}(z+\Delta z) \\ A_n^{(y)}(z+\Delta z) \end{bmatrix} = H \begin{bmatrix} A_n^{(x)}(z) \\ A_n^{(y)}(z) \end{bmatrix} = \begin{bmatrix} h_{xx}A_n^{(x)}(z) & +h_{xy}A_n^{(y)}(z) \\ h_{yx}A_n^{(x)}(z) & +h_{yy}A_n^{(y)}(z) \end{bmatrix} \tag{3.9}$$

where the four complex entries $h_{i,j}$ are the elements of the unitary unit-determinant matrix $H = \begin{bmatrix} h_{xx} & h_{xy} \\ h_{yx} & h_{yy} \end{bmatrix} \triangleq \mathrm{e}^{-\mathrm{i}\theta[\hat{\mathrm{a}}\cdot\vec{\sigma}]}$ with $h_{yy} = h_{xx}^*$, $h_{yx} = -h_{xy}^*$ and $|h_{xx}|^2 + |h_{yy}|^2 = 1$.

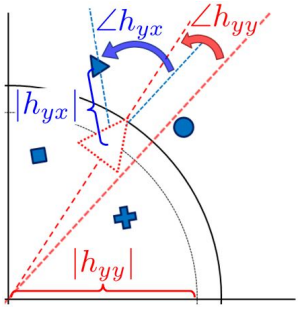


Figure 3.2: Upper-right quadrant of Y-polar after propagating in presence of Pivot oriented along a general SOP with angular coordinates $(2\chi,\phi)$.

Fig. 3.2 is a detailed view of the upper-right quadrant of the received Y-constellation, which shows how the magnitudes and arguments of the complex coefficients h_{ij} affect the received symbols. For instance, $|h_{yx}| = |h_{xy}| = |\sin(\theta)\sin(2\chi)|$ determines the radius of the four crosstalk-patterns, while $\arg[h_{yx}] = -\pi - \arg[h_{xy}] = \phi - \frac{\pi}{2}$ determines their angular offset. Such crosstalk-patterns may collapse to a single value ($|h_{yx}| = 0$) if $2\chi = 0$ or π, i.e., if the Pivot is parallel to a PDM subcarrier, which is the first case that we will analyze next.

- Pivot parallel to the PDM polarization subcarriers:

in this case, suppose that the direction of the Pivot coincides with that of the PDM subcarrier x ($\widehat{S_1}$ in Stokes space), so that $2\chi = 0$ and (3.7) reduces to the first term only. The general solution (3.8) becomes

$$\vec{A_n}(z+\Delta z) = \begin{bmatrix} \cos(\theta) - i\sin(\theta) & 0 \\ 0 & \cos(\theta) + i\sin(\theta) \end{bmatrix} \vec{A_n}(z) = \begin{bmatrix} \exp\{-i\theta\}A_n^{(x)}(z) \\ \exp\{+i\theta\}A_n^{(y)}(z) \end{bmatrix} \tag{3.10}$$

The result that we just found implies that each of the polarizations in $\vec{A_n}$ is independently rotated by opposite angles $\pm\theta$: the impact on the QPSK symbols is visible in Fig. 3.3.

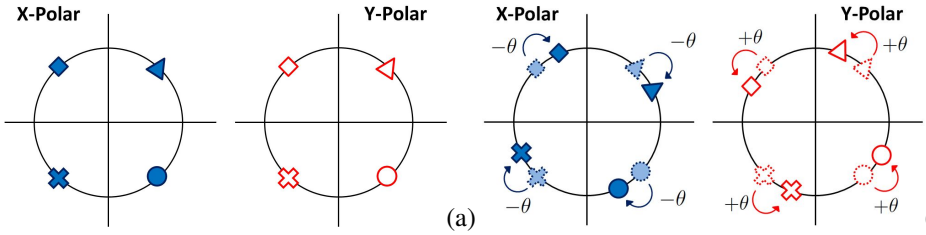


Figure 3.3: X-polar and Y-polar at the transmitter side (a) and after propagating in presence of Pivot oriented along S_1 axis (b).

Until now we supposed that the Pivot has a constant magnitude, but in general it is not always true. If we consider for example a WDM comb in which the interfering channels are OOKs we obtain that the nonlinear rotation angle θ depends on the number of "active" OOK channels, i.e., those transmitting a *mark*, whose power contributes to the summation appearing in the definition of θ. Thus, especially in the case of many interfering OOK channels, the QPSK patterns are not simply doubled, as shown in Fig. 3.3 for the sake of illustration, but rather spread on a range of angular values whose maximum is simply obtained by counting all channels as active. The picture is further enriched by the presence of walk-off, which varies the number of active OOK channels during the integration of (3.3). The effect of walk-off and modulation of interfering channels will be shown in Sec. 3.3, when the PDM-QPSK propagates with OOKs. Moreover, looking back at eq. (3.1), it is clear that both received constellations in Fig.3.3 will be rotated by the same amount, due to the scalar nonlinear phase rotation brought about by the *XPM* term. As a matter of fact, in the present special case, where the Pivot is aligned with the first PDM subcarrier, the *XPolM* term in (3.3) algebraically adds to the *XPM* term in (3.1):

$$\begin{aligned}\frac{d\overrightarrow{A_n}}{dz} &= -i\frac{8}{9}\gamma\left\{\underbrace{|A_n|^2\overrightarrow{A_n}}_{SPM}+\underbrace{\frac{3}{2}\sum_{k\neq n}|A_k|^2\overrightarrow{A_n}}_{XPM}+\underbrace{\frac{1}{2}\sum_{k\neq n}|A_k|^2\begin{bmatrix}1&0\\0&-1\end{bmatrix}\overrightarrow{A_n}}_{XPolM}\right\}\\ &= -i\frac{8}{9}\gamma\left\{\underbrace{|A_n|^2\overrightarrow{A_n}}_{SPM}+\underbrace{\sum_{k\neq n}|A_k|^2\begin{bmatrix}2A_n^{(x)}\\A_n^{(y)}\end{bmatrix}}_{XPM+XPolM}\right\}\end{aligned}$$

hence *XPolM* behaves as a ***differential XPM***, making the XPM impact on the (x)-polarization

twice as large as the impact on the (y)-polarization[6]. Such an angular distortion may be subsequently compensated by a phase recovery algorithm, provided that phase recovery acts on the two polarization subcarriers separately, as we will show in Sec. 3.3.

- Pivot orthogonal to the PDM polarization subcarriers:

in this case, suppose that the direction of the interfering fields $\overrightarrow{a_k}$ is orthogonal (in Stokes space) to that of the PDM subcarriers (x,y) ($\pm\widehat{S_1}$ in Stokes space), so that $2\chi = \frac{\pi}{2}$ and (3.7) reduces to the second term only. The result of the integration over a step Δz is again the one reported in (3.8), with the imaginary parts of the diagonal terms set to zero:

$$\overrightarrow{A_n}(z+\Delta z) = \begin{bmatrix} \cos(\theta) & -i\sin(\theta)\mathrm{e}^{-\mathrm{i}\phi} \\ -i\sin(\theta)\mathrm{e}^{+\mathrm{i}\phi} & \cos(\theta) \end{bmatrix} \overrightarrow{A_n}(z) \tag{3.11}$$

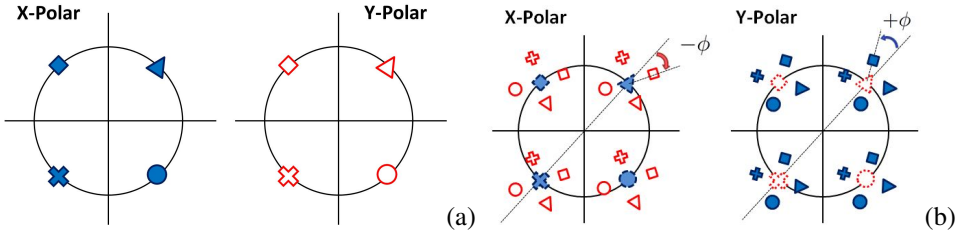


Figure 3.4: X-polar and Y-polar at the transmitter side (a) and after propagating in presence of a Pivot orthogonal to the PDM polarization subcarriers (b).

The impact on the QPSK symbols is visible in Fig. 3.4, where, contrary to the previous case of Pivot aligned with PDM subcarriers, *XPolM* introduces not only a rotation of the angular position of the symbols, but also a variation of their radial position: a phenomenon which was already evident from Fig. 3.1. We will call such a phenomenon ***cross-amplitude modulation (XAM)***.

All of the observations made in the previous case (Pivot aligned with PDM), regarding the role of bit-patterns and walk-off, still apply. To simplify the discussion, we are assuming in Fig. 3.4 that the Pivot amplitude is constant, so that θ is a constant angle.

The dashed symbols in Fig.3.4(b) help identify the cross-talk patterns associated with the transmission of a specific QPSK symbol on the corresponding polarization in Fig 3.4(a).

[6]By a similar argument, it follows that the XPolM will double the impact of XPM on the (y)-polarization, if the Pivot is aligned with the second PDM subcarrier (y), in the standard reference frame).

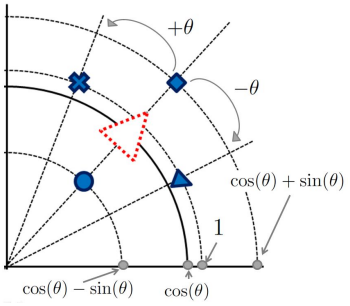


Figure 3.5: Upper-right quadrant of Y-polar after propagating in presence of orthogonally oriented Pivot . The figure is obtained for a choice of $\phi = 0$, which maximizes the spread among the symbols' magnitudes.

For instance if the symbol $\exp\{i\frac{\pi}{4}\}$ is transmitted on the Y-polarization (marked by a red triangle in Fig. 3.4a), one of the four blue symbols in the upper-right quadrant in Fig. 3.4(b) (surrounding the dashed red triangle) will appear in the Y-polarization.

To quantify the amount of XAM, one can calculate the radius of the four symbols in each cross-talk subpattern, by taking the magnitude of the (x,y) elements in (3.11):

$$A_n^{(x,y)}(z+\Delta z) = \cos(\theta)A_n^{(x,y)}(z) - i\sin(\theta)\mathrm{e}^{\mp \mathrm{i}\phi}\mathrm{A}_\mathrm{n}^{(\mathrm{y,x})}(\mathrm{z})$$

where the minus (plus) sign applies to ϕ in the x (y) element. The squared magnitude is thus

$$\begin{aligned}\left|A_n^{(x,y)}(z+\Delta z)\right|^2 &= \cos^2(\theta)\left|A_n^{(x,y)}(z)\right|^2 + \sin^2(\theta)\left|A_n^{(y,x)}(z)\right|^2 \\ &\quad +2\Re\left\{\sin(\theta)\cos(\theta)(+i)\mathrm{e}^{\pm \mathrm{i}\phi}\mathrm{A}_\mathrm{n}^{(\mathrm{x,y})}(\mathrm{z})\left(\mathrm{A}_\mathrm{n}^{(\mathrm{y,x})}(\mathrm{z})\right)^*\right\} \\ &= 1+\sin(2\theta)\Re\{(+i)\mathrm{e}^{\pm \mathrm{i}\phi}\mathrm{e}^{\pm \mathrm{ik}\frac{\pi}{2}}\} \\ &= 1-\sin(2\theta)\sin\left(\pm\phi \pm k\frac{\pi}{2}\right) \end{aligned} \tag{3.12}$$

where we assumed that the original constellations lie on the unit-circle $\left(\left|A_n^{(x,y)}(z)\right|^2 = 1\right)$, so that the inner product of the transmitted symbols $A_n^{(x,y)}(z)(A_n^{(y,x)}(z))^* = \mathrm{e}^{\pm \mathrm{ik}\frac{\pi}{2}}$ has an argument multiple of 90°. Thus, the variation in the intensity of the received symbols can be as large as $2\sin(2\theta)$. From (3.12), such maximal variation in symbols' intensity is reached if ϕ is a multiple of 90°: assuming, for instance, that $\phi = 0$, we can check that $\left|A_n^{(x,y)}(z+\Delta z)\right| = \cos(\theta) \pm \sin(\theta)$, if the two symbols transmitted on $A_n^{(x)}$ and $A_n^{(y)}$ are in quadrature (k odd),

while $\left|A_n^{(x,y)}(z+\Delta z)\right| = 1$ if the same symbol ($k=0$) or two opposite symbols ($k=2$) are transmitted. This is the case depicted in Fig. 3.5, showing the upper-right quadrant of the Y-polarization after propagation; similar figures (with 90° rotations) are obtained for $\phi = m\frac{\pi}{2}$. In Fig. 3.5, the four blue symbols form a square, aligned with the complex plane axes, and centered at the red-dashed triangle, whose position is $\cos(\theta)\exp\{i\frac{\pi}{4}\}$, as evident from eq. (3.11). Two corners of the square lie on the unit circle, hence two out of the four possibly received symbols are only subject to angular distortions, while the other two are subject to amplitude distortion (XAM).

The pictures given above assume a well-determined value of θ, but, as we already noted, the nonlinear rotation angle θ can vary in time, due to the interfering channels' bit-patterns, that walk past each other because of chromatic dispersion. Thus, a polar plot of the received symbols will entail many possible values of θ, including very small or null values for which the matrix in (3.8) reduces to the identity, hence the received symbol coincides with the transmitted one. In order to describe the full received symbol patterns, we must then let θ vary between zero and its maximum theoretical value (all mark bits, on OOK channels, i.e., continuous wave interfering) and see how, for given choices of symbols on $A_n^{(x,y)}(z)$, the field $A_n^{(x)}(z+\Delta z)$ (or $A_n^{(y)}(z+\Delta z)$) moves on the complex plane. We shall now demonstrate that the curve described by a received symbol, as a function of θ, is an arc of ellipse. Assume we transmit unit-amplitude QPSK symbols on the PDM subcarriers, such that their phase difference is $A_n^{(x)}(z)(A_n^{(y)}(z))^* = \mathrm{e}^{+\mathrm{i}k\frac{\pi}{2}}$, hence $A_n^{(y)}(z) = A_n^{(x)}(z)\mathrm{e}^{-\mathrm{i}k\frac{\pi}{2}}$ and the symbol received on (x) results[7] from (3.8)

$$\begin{aligned} A_n^{(x)}(z+\Delta z) &= A_n^{(x)}(z)\left\{[\cos(\theta) - i\sin(\theta)\cos(2\chi)]\right\} + \\ &- A_n^{(x)}(z)\left\{\left[i\sin(\theta)\sin(2\chi)\mathrm{e}^{-\mathrm{i}\phi}\right]\mathrm{e}^{-\mathrm{i}k\frac{\pi}{2}}\right\}. \end{aligned} \tag{3.13}$$

Let us now define the unit-magnitude Stokes vector:

$\hat{k} = [k_1; k_2 k_3] = [\cos(2\chi); \sin(2\chi)\cos\left(\phi + k\frac{\pi}{2}\right); \sin(2\chi)\sin\left(\phi + k\frac{\pi}{2}\right)]$, which can be easily located on the Poincarè sphere with respect to the SOP of the interfering OOK channels (which we assumed oriented as the Pivot). Its first component is the same as $\hat{a}$, $k_1 = a_1$, and the remaining two components are rotated by $k\frac{\pi}{2}$ with respect to $\hat{a}$. The normalized $k_{1,2,3}$ ($k_1^2 + k_2^2 + k_3^2 = 1$) are useful to express the received symbol in the complex plane as

$$A_n^{(x)}(z+\Delta z) = A_n^{(x)}(z)\left\{[\cos(\theta) - k_3\sin(\theta)] + i[-(k_1+k_2)\sin(\theta)]\right\}. \tag{3.14}$$

[7] A similar result holds for $A_n^{(y)}(z+\Delta z)$, since, from (3.8),

$$A_n^{(y)}(z+\Delta z) = A_n^{(y)}(z)\left\{[\cos(\theta) + i\sin(\theta)\cos(2\chi)] - \left[i\sin(\theta)\sin(2\chi)\mathrm{e}^{+\mathrm{i}\phi}\right]\mathrm{e}^{+\mathrm{i}k\frac{\pi}{2}}\right\}$$

where the only differences from (3.13) are in the signs of ϕ, k and the imaginary part of the first term.

In fact, besides the (unit-magnitude) transmitted symbol $A_n^{(x)}(z)$, which just accounts for an angular rotation in the complex plane, the real and imaginary parts $(R(\theta), I(\theta))$ of (3.14) are

$$\begin{cases} R(\theta) = \cos(\theta) & -k_3 \sin(\theta) \\ I(\theta) = & -(k_1 + k_2)\sin(\theta) \end{cases} \tag{3.15}$$

These coordinates can be compared with the general parametric form of the Cartesian coordinates $(R_e(\theta), I_e(\theta))$ of an ellipse

$$\begin{cases} R_e(\theta) = & a\cos(\lambda)\cos(\theta+\theta_0) - b\sin(\lambda)\sin(\theta+\theta_0) \\ I_e(\theta) = & a\sin(\lambda)\cos(\theta+\theta_0) + b\cos(\lambda)\sin(\theta+\theta_0) \end{cases} \tag{3.16}$$

centered at the origin, with azimuth λ, major and minor axes a and b; the starting point of the ellipse, at $\theta = 0$, is determined by the angular offset θ_0 with respect to the major axis' tip. A comparison of the symbols' trajectories in (3.15) with the parametric ellipse in (3.16) yields the following set of four non linear equations

$$\begin{cases} a\cos(\lambda)\cos(\theta_0) - b\sin(\lambda)\sin(\theta_0) & = 1 \\ a\sin(\lambda)\cos(\theta_0) + b\cos(\lambda)\sin(\theta_0) & = 0 \\ -a\cos(\lambda)\sin(\theta_0) - b\sin(\lambda)\cos(\theta_0) & = -k_3 \\ -a\sin(\lambda)\sin(\theta_0) + b\cos(\lambda)\cos(\theta_0) & -(k_1+k_2) \end{cases} \tag{3.17}$$

that must be solved to find the four parameters a, b, λ, θ_0. Algebraic manipulation of the above equations reveals that the squared major/minor axes, a^2 and b^2, are such that their sum is $a^2 + b^2 = 2(1 + k_1 k_2) > 0$ and their product is $a^2 b^2 = (k_1 + k_2)^2$, hence the squared axes[8] are

$$\begin{aligned} a^2 &= (1 + k_1 k_2) + \sqrt{k_3^2 + k_1^2 k_2^2} \\ b^2 &= (1 + k_1 k_2) - \sqrt{k_3^2 + k_1^2 k_2^2} \end{aligned}$$

Once the axes lengths are calculated, the ellipse's azimuth λ can be obtained from (3.17) as

$$\tan^2(\lambda) = \frac{1 - a^2}{b^2 - 1} \cdot \frac{b^2}{a^2}$$

where $b^2/a^2 = \tan^2(\varepsilon)$ stems from the definition of the ellipticity ε. The initial offset θ_0 obeys (again from (3.17))

$$\tan(\theta_0) = -\frac{a}{b}\tan(\lambda)$$

[8] Recall that a and b are the lengths of the ellipse's axes, hence they are intrinsically positive.

hence has always an opposite sign with respect to λ. The elliptical shapes of the received symbol patterns is thus determined, once the interfering signals' SOP is given and the transmitted symbols' offset k in (3.13) is specified.

Since k in (3.13) can take four integer values $(0,1,2,3)$ for each of the four possible QPSK symbols $A_n^{(x)}(z)$, this would imply a total of 16 possible elliptical trajectories in the complex plane, for the (x) polarization. Nonetheless, changing the symbol $A_n^{(x)}(z)$ to its opposite in (3.13) has the same effect as changing the nonlinear rotation angle by adding π to θ: the possible elliptical trajectories are thus 8. In addition, some of the 8 ellipses may coincide in pairs (e.g., if $2\chi = \frac{\pi}{2}, \phi = \frac{\pi}{4}$) or in quartets (e.g., if $2\chi = \frac{\pi}{2}, \phi = 0$) or even all 8 can coincide on the unit circle, as in the case of OOK channels SOP (Pivot) parallel to the PDM subcarriers ($2\chi = 0$).

The general result can be stated as follows: **in the case of Pivot with fixed orientation and θ depending on the bits transmitted on the interfering channels, the patterns received on each PDM subcarrier describe arcs of** 8 **possible ellipses. The overall power of the interferers determines the maximum angular span of such arcs**[9]. This noteworthy result will be verified by simulations in Sec. 3.3.1.1.

It is important to remark once again that the effects described so far and caused by XPolM (both angular and magnitude variations) co-exist with the time varying common-phase variations caused by the SPM and XPM term in eq. (3.1). Being XPM a scalar phenomenon, such phase rotations will affect both polarizations in the same way. In particular, the geometrical layouts of the received (x,y) constellations, described so far under various configurations of the interfering channels' SOP, will be subject to a common angular jitter, brought about by a time-varying XPM.

3.3 XPolM impact on the received symbols – Simulation Results

In eq. (3.8) we supposed a Pivot with both fixed SOP direction and modulus. Consequently, the SOP of each channel simply rotates around it in Stokes space, as discussed in [46]. However, in general this assumption is not true, since in the presence of walk-off and channels' modulation the Pivot changes continuously its modulus along the propagation and, if channels have different polarizations, also its SOP direction.

[9] In general, complete ellipses can be seen in the scattering plots when the OOK power is such that the nonlinear rotation angle approaches π.

In this section we investigate the impact of *XPolM* on the received PDM-QPSK symbols in the presence of a time-varying Pivot, showing many examples of constellations after propagation in the presence of OOK signals. We reproduced the special case in which the Pivot direction is fixed during the propagation, transmitting OOK channels aligned in polarization and with a power much larger than the PDM-QPSK channel. We also analyzed a real case in which the Pivot SOP changes along the fiber, for example using a WDM comb composed of a large number of channels with randomly oriented SOPs. It is important to recall that, in both cases, the Pivot modulus is not constant, due to modulation of interfering channels and walk-off.

3.3.1 DM link

3.3.1.1 Pivot with fixed orientation

In this section we focus on the case in which the Pivot SOP remains fixed along the propagation. We transmitted the OOK channels all aligned in polarization and with a power much larger than the central PDM-QPSK channel (i.e., the PDM-QPSK channel does not influence the Pivot SOP). This way the Pivot SOP coincides in direction with the OOK SOPs, while the magnitude depends on the OOK bits at the time of interest. We changed the OOK SOPs to analyze the special cases in which the Pivot is parallel or orthogonal to the PDM-QPSK subcarriers. In Fig. 3.6(a) we report some SOPs that will be analyzed in the following: parallel (along $\widehat{S_1}$ axis) or orthogonal (along $\widehat{S_2}$ axis and with orientation labeled $\widehat{S_A}$) to the PDM-QPSK subcarriers and equidistant (labeled $\widehat{S_B}$) from the main axes of the Poincaré sphere, i.e., with an azimuth of $\frac{\pi}{8}$ and an ellipticity of $\frac{\pi}{8}$ (we will analyze it in Sec. 3.3.3). $\widehat{S_A}$ is coplanar with the PDM-QPSK symbols' SOP, since it lies in the ($\widehat{S_2}$, $\widehat{S_3}$) plane with an angle of $\frac{\pi}{4}$ from both $\widehat{S_2}$ and $\widehat{S_3}$ axes.

System Setup We transmitted a WDM comb composed of a central 112 Gb/s RZ(50%)-PDM-QPSK (-10 dBm) surrounded by two 10 Gb/s NRZ-OOK channels (5 dBm each), with the same SOP and with a channel spacing of 2 nm. To avoid extra inter-symbol interference (ISI), neither optical nor electrical filter were applied.

The optical link is dispersion managed, composed of 20×100 km spans of standard single mode fiber (SMF) ($D = 17$ ps/nm/km @1550 nm, $A_{eff} = 80\ \mu m^2$, $n_2 = 2.5 \cdot 10^{-20}$ m^2/W, $\alpha = 0.2$ dB/km, $D^{'} = 0$ ps/nm^2/km,$\gamma = 1.3$ W^{-1}km^{-1}). The DM link had a pre-compensation whose dispersion was chosen according to the straight line rule [48], an in-line residual dispersion of 30 [ps/nm/span] and a post-compensating fiber whose value sets the

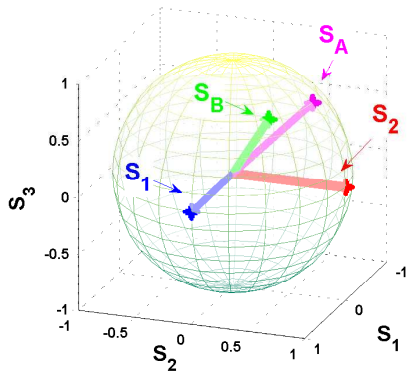


Figure 3.6: OOK SOPs under analysis: along $\widehat{S_1}$ (blue), along $\widehat{S_2}$ (red), along $\widehat{S_A} = \left[0, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right]^T$ (magenta) and along $\widehat{S_B} = \left[\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right]^T$ (green). When the OOK power is much larger than the PDM-QPSK power, such orientations coincide with the Pivot direction.

overall dispersion to zero. All compensating fibers were simulated as purely linear. Fiber birefringence was modelled by 50 discrete random waveplates per span, with zero PMD. The propagation, through the Manakov-PDM equation, was solved with the Split Step Fourier Method (SSFM), neglecting FWM [22]. All the optical amplifiers had a gain perfectly flat in frequency and were noiseless.

We used two different phase recovery algorithms: the first, which we call data-aided in the following, simply recovers the time-average nonlinear phase rotation calculated as

$$\overline{\phi} = \arg\left\langle \overrightarrow{A}_{RX} \cdot \overrightarrow{A}^{*}_{TX} \right\rangle$$

where $\overrightarrow{A}_{RX}$ and $\overrightarrow{A}_{TX}$ are the received/transmitted signal, respectively, while $\langle . \rangle$ indicates time-averaging over the entire sequence length. The second method is based on the Viterbi and Viterbi algorithm [38], whose estimated phase is

$$\phi_{VV}(n) = \frac{1}{4} unwrap\left\{ \arg\left[\sum_{k=n-L}^{n+L} |A_{RX}(k)|^2 \exp\{+i4\phi_{RX}(k)\} \right] \right\}$$

where $N = 2L+1$ is the number of taps of the moving average, while ϕ_{RX} is the phase of $\overrightarrow{A}_{RX}$. The *unwrap(P)* function corrects the radian phase angles in a vector P by adding multiples of $\pm 2\pi$ when absolute jumps between consecutive elements of P are greater than or equal to a

jump of π radians. Note that we use $|A_{RX}(k)|^2$ instead of $|A_{RX}(k)|^4$, since it is more robust at small signal-to-noise ratio [38].

Once the phase offset was estimated and the received constellation was centered as shown in Fig. 3.7 (left), we folded the constellation first horizontally and then vertically thus obtaining a one-symbol constellation, as indicated in Fig. 3.7 (right). Then we analyzed the Probability Mass Function (PMF) of the real and imaginary components of the symbol in the reference system shown in Fig. 3.7 (right). From these PMFs we can directly get the probability of exceeding the decision threshold hence making a symbol error, which is twice the probability of bit error in absence of differential decoding and using Gray coding.

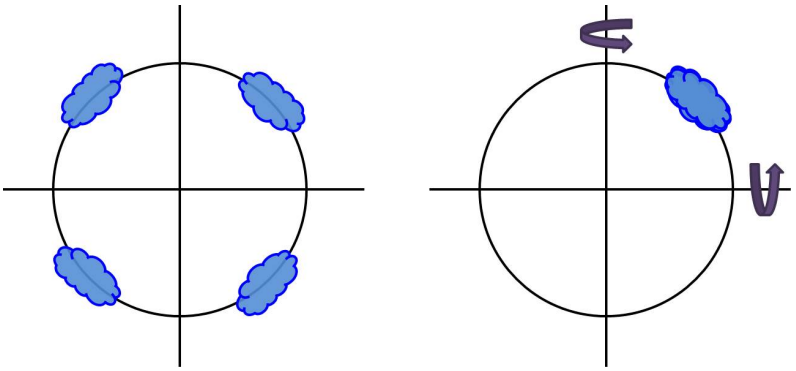


Figure 3.7: (Left) Initial constellation and (right) one-symbol constellation, after folding. The PMFs of the real and imaginary components have been evaluated on the one-symbol folded constellation.

Pivot parallel to the PDM polarization subcarriers We started orienting the OOK channels along $\widehat{S_1}$, co-polarized with one of the PDM-QPSK polarization subcarriers. As seen in Sec. 3.2, in this case the *XPolM* term adds to the *XPM* term in changing the relative phase of the two subcarriers. In particular, the PDM-QPSK subcarrier parallel to the OOKs ($\widehat{S_1}$) acquires a phase shift twice as large as the orthogonal component ($-\widehat{S_1}$). It turns out that *XPolM* does not change the polarization of the individual PDM-QPSK subcarriers, since it does not mix the two polarizations.

The simulation results are depicted in Fig. 3.8. This figure reports the constellation of the two PDM-QPSK subcarriers, obtained from the model in eq. (3.8) (see Fig. 3.8(a)) or by simulation (Fig. 3.8(b)). Fig. 3.8(c) shows the PDM-QPSK stokes vectors before and after optical propagation. Fig. 3.8(d) finally shows the PMFs of the real and imaginary components

of X-polarization after folding the symbols as discussed in Fig. 3.7(right). For the sake of completeness, in Fig. 3.8 (d) we also report the Gaussian fit of the real and imaginary PMF, using the mean and the variance of the numerical PMFs, respectively. Note that the real and imaginary components are here superposed, making them indistinguishable in the figure.

Note that Fig. 3.8(c) shows the PDM-QPSK stokes vectors of the overall waveform, while in Fig. 3.8(b) and (d) only the sample at the center of each bit is reported.

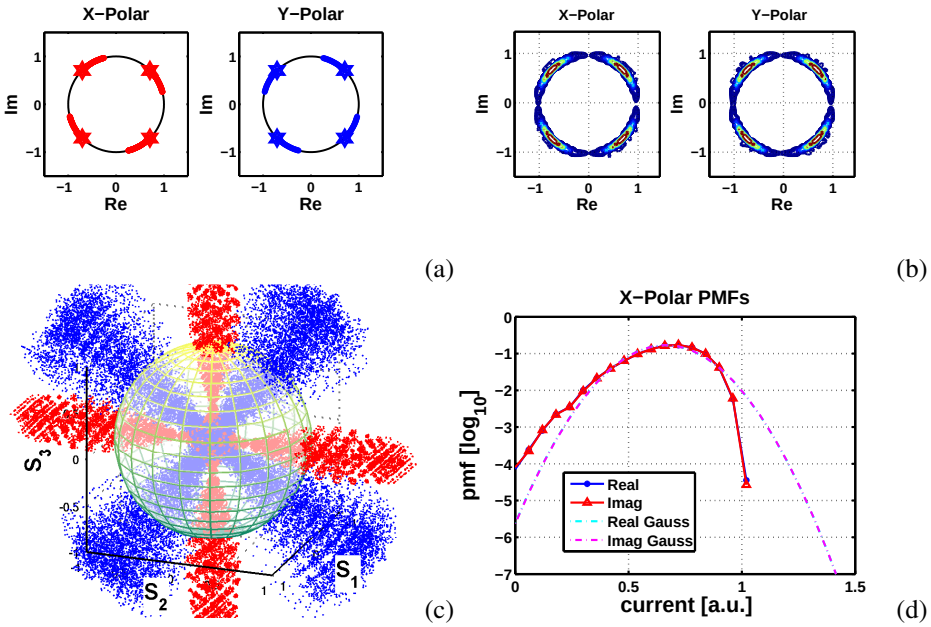


Figure 3.8: OOK channels and Pivot oriented along $\widehat{S_1}$. PDM-QPSK constellations applying eq. (3.8) (a) and after propagation along link (b). PDM-QPSK stokes vector before (red) and after propagation (blue) (c). PMFs of real (blue-circle) and imaginary components (red-triangle) of the one-symbol constellation (right-side Fig. 3.7) of X-polarization (d). The two PMFs are superposed.

Fig. 3.8(b) confirms eq. (3.8), showing that in this setup XPolM shows up as phase noise. Note that Fig. 3.8(a) shows a negative correlation between the phases of the two subcarriers, so that the transmitted symbols (indicated by stars) rotate in opposite directions. Such an information is not visible in Fig. 3.8(b) because the phase estimator recovered the average

nonlinear phase rotation. Fig. 3.8(c) shows that, as predicted by eq. (3.6), the PDM-QPSK stokes vector rotates around the $\hat{S_1}$ axis on average by an angle of 2θ, i.e., proportional to the induced nonlinear phase rotation. Note that, at the transmitter side, the PDM-QPSK signal has the $\widehat{S_1}$ component equal to zero, hence its SOP, in absence of distortions, lies in the plane $\left(\widehat{S_2}, \widehat{S_3}\right)$. At any given time, the SOP is one of the points shown in gray in Fig. 3.8(c), so that over a sufficiently long observation time the plot takes the characteristic form of a cross in 3D Stokes space (App. B). Comparing the solid and dotted curves of Fig. 3.8(d) we note that the PMF of the real and imaginary components do not match well the Gaussian distributions and have higher probability to exceed the decision threshold.

Pivot orthogonal to the PDM polarization subcarriers We now move to a case in which the Pivot is orthogonal (in Stokes space) to the PDM-QPSK subcarriers. We chose OOK channels aligned with $\widehat{S_2}$, obtaining the results shown in Fig. 3.9.

Fig. 3.9(a) and (b) confirm the scenario depicted in Fig. 3.5 for the symbol in the upper-right quadrant. Strictly speaking, two of the four «interfering-symbols» lie on the circle with unit-magnitude radius, while the other two symbols lie on the radius connecting the origin to the point $\exp\left\{i\frac{\pi}{4}\right\}$. The walk-off and the OOK modulation move the two pairs of symbols along the circle and the radius respectively, creating the continuous clouds depicted by the constellations. The effect of rotation of the PDM-QPSK stokes vector around the $\widehat{S_2}$ axis is shown in Fig. 3.9(c), in which the two clouds on $\pm\widehat{S_3}$ are rotated on average by an angle 2θ. The PMFs of the real and imaginary components are similar to the previous case although the constellation clouds are very different. Also the probability to exceed the decision threshold is similar to OOK $\widehat{S_1}$ case.

Note that a Pivot oriented along $\widehat{S_3}$ produces the same constellations as in Fig. 3.9(b) and the same amount of SOP rotation as Fig. 3.9(c). The only difference is that the PDM-QPSK SOP at the fiber output is rotated around the $\widehat{S_3}$ axis, hence the two clouds in $\pm\widehat{S_3}$ remain fixed, while the two clouds in $\pm\widehat{S_2}$ are rotated on average by an angle 2θ.

In Fig. 3.10 we analyze now another case in which the Pivot is orthogonal (in Stokes space) to the PDM-QPSK subcarriers. The OOK SOPs are now oriented as $\widehat{S_A}$ in Fig. 3.6, i.e., in the ($\widehat{S_2}$, $\widehat{S_3}$) plane with an angle of $\frac{\pi}{4}$ with respect to both $\widehat{S_2}$ and $\widehat{S_3}$ axes.

Fig. 3.10 (c) shows that all the PDM-QPSK clouds rotate around the Pivot thus yielding a SOP lying in a new plane that is not ($\widehat{S_2}$,$\widehat{S_3}$). The effect is similar to the one observed in Fig. 3.8 and Fig. 3.9, except that now an $\widehat{S_1}$ component of the SOP exists for any cloud. It turns out that we always observe a power imbalance between the two subcarriers that takes away the symbols from radius one, as confirmed by the constellation in the figure. Since the

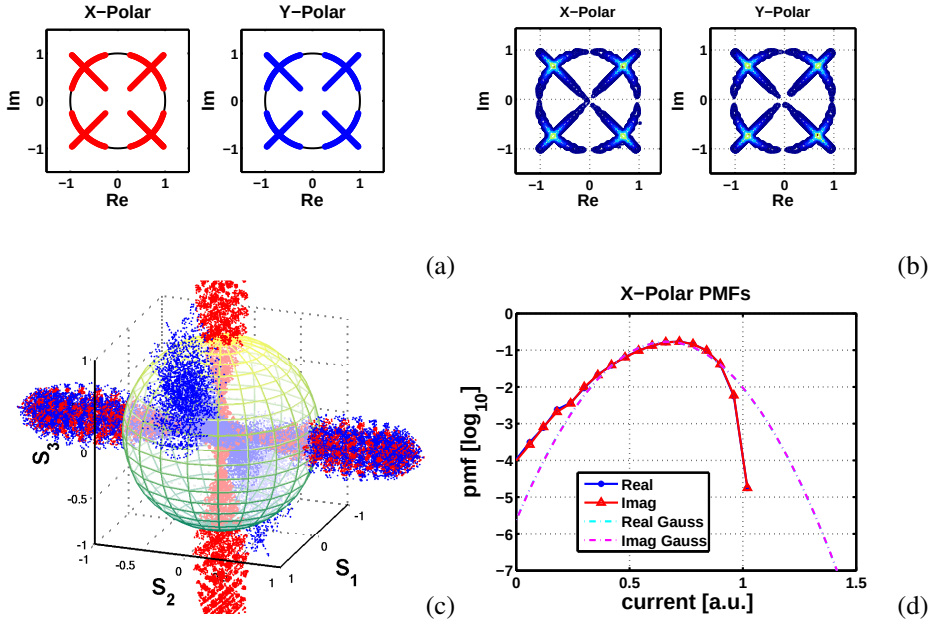


Figure 3.9: OOK channels and Pivot oriented along $\widehat{S_2}$. PDM-QPSK constellations applying eq. (3.8) (a) and after propagation along link (b). PDM-QPSK stokes vector before (red) and after propagation (blue) (c). PMFs of real (blue-circle) and imaginary components (red-triangle) of the one symbol constellation (right-side Fig. 3.7) of X-polarization (d). The two PMFs are superposed.

overall energy is conserved, these shifts are negatively-correlated between the two subcarriers. Observing Fig. 3.10 (d) we note that the PMFs do not match the Gaussian shape and have a higher peak centered at their center. The probability to exceed the decision threshold is higher than both OOK $\widehat{S_1}$ case and OOK $\widehat{S_2}$ case by almost one order of magnitude.

We simulated the same system as in Fig. 3.10 but now increasing the OOK power to 10 dBm. We obtained the constellation in Fig. 3.11, which confirms that the constellations become ellipses if the interfering channels induce a sufficiently high nonlinear phase rotation.

Phase recovery based on Viterbi and Viterbi algorithm – Pivot parallel to the PDM polarization subcarriers In real DSP-based coherent receivers[49] (Sec. 2.2.1) the phase

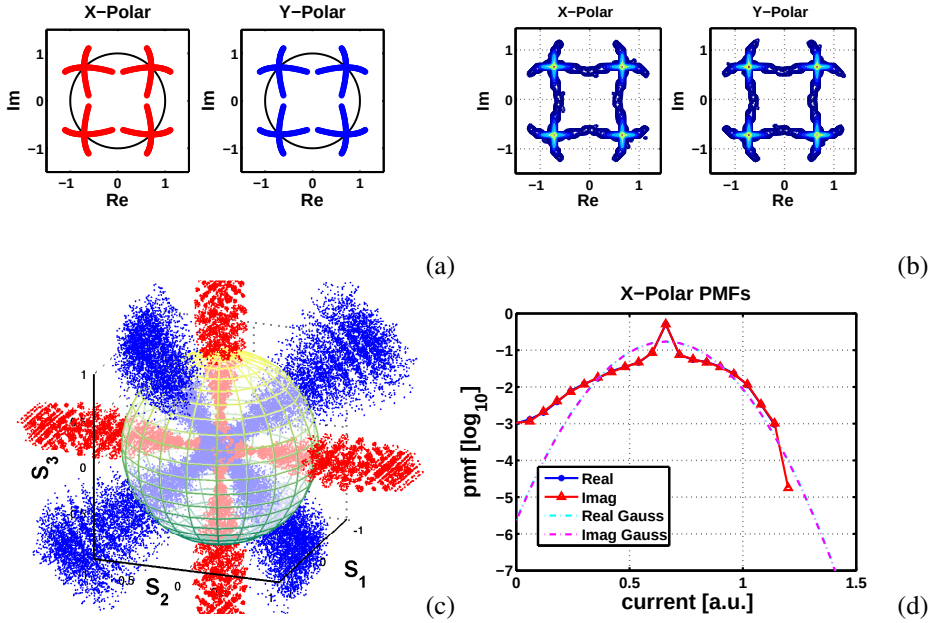
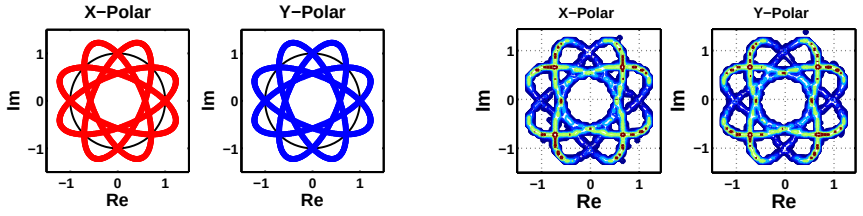


Figure 3.10: OOK channels and Pivot oriented as $\widehat{S_A}$. PDM-QPSK constellations applying eq. (3.8) (a) and after propagation along link (b). PDM-QPSK stokes vector before (gray) and after propagation (black) (c). PMFs of real (blue-circle) and imaginary components (red-triangle) of the one symbol constellation (right-side Fig. 3.7) of X-polarization (d). The two PMFs are superimposed.

recovery is usually accomplished by the Viterbi and Viterbi algorithm [38]. In this section we investigate the impact of the Viterbi and Viterbi algorithm on the previous results.

Starting form the case of OOKs oriented along $\widehat{S_1}$, we obtained the constellations depicted in Fig. 3.12. The figure reports the case of a V&V with 7, 21 and 41 taps, respectively.

Generally speaking, Fig. 3.12 shows that Viterbi and Viterbi algorithm, working on a window shorter than the sequence length, better recovers the nonlinear phase distortion and reduces the constellation spread observed in Fig. 3.8(b). Furthermore, we note that the constellation spread is weaker for shorter number of taps, like for scalar XPM effects [50], in agreement with the previous explanation that described *XPolM* as an additional *XPM* term in this special setup.



(a) (b)

Figure 3.11: OOK channels (10 dBm) and Pivot oriented as $\widehat{S_A}$. PDM-QPSK constellations applying eq. (3.8) (a) and after propagation along link (b).

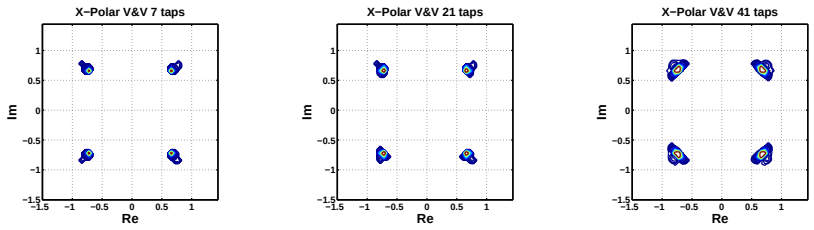


Figure 3.12: X-polar: OOKs and Pivot along $\widehat{S_1}$. Phase recovery based on Viterbi and Viterbi algorithm with different number of taps (N): 7 (left), 21 (center) and 41 taps (right).

Phase recovery based on Viterbi and Viterbi algorithm – Pivot orthogonal to the PDM polarization subcarriers Fig. 3.13 shows the constellations obtained when the OOKs are oriented along $\widehat{S_2}$ for different number of Viterbi and Viterbi taps.

Unlike the case analyzed before, the Viterbi and Viterbi does not reduce the constellation spread, which increases for decreasing number of taps. This difference of Viterbi and Viterbi effectiveness suggests that the *XPolM* term induces a different distortion in this case. We will discuss in detail this observation in Sec. 3.3.3. More details on the impact of the Viterbi and Viterbi number of taps on the system performance can be found in Sec. 4.1.2.

3.3.1.2 Randomly oriented Pivot

In this section we analyze the more realistic case in which both the modulus of the Pivot and its direction change along the fiber.

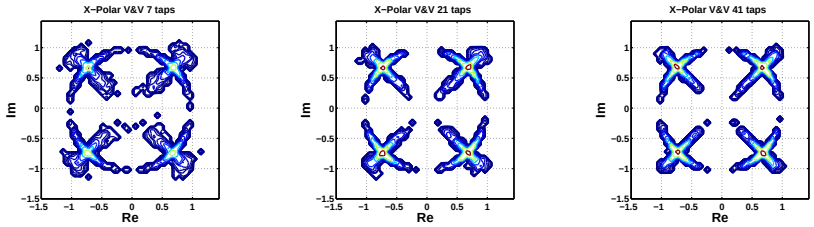


Figure 3.13: X-polar: OOKs and Pivot along $\widehat{S_2}$. Phase recovery based on Viterbi and Viterbi algorithm with different number of taps (N): 7 (left), 21 (center) and 41 taps (right).

System Setup Starting from the setup of Sec. 3.3.1.1, we analyzed both a 3-channel case (fast simulation) and 19-channel case (slow simulation) with randomly oriented OOK SOPs and random patterns. The fast setup has been analyzed for a total of 100 different random generator seeds, corresponding to a selection of different WDM random data patterns and SOPs. In the slow setup the 18 OOK had 0 dBm of power and were 50 GHz-spaced. Such a setup has been analyzed for 10 different random seeds, corresponding to different data patterns. The results obtained using the data-aided phase recovery algorithm will be compared those of the Viterbi and Viterbi algorithm.

Two OOKs with random SOPs The constellations and the PMFs obtained with the three channel-case are reported in Fig. 3.14.

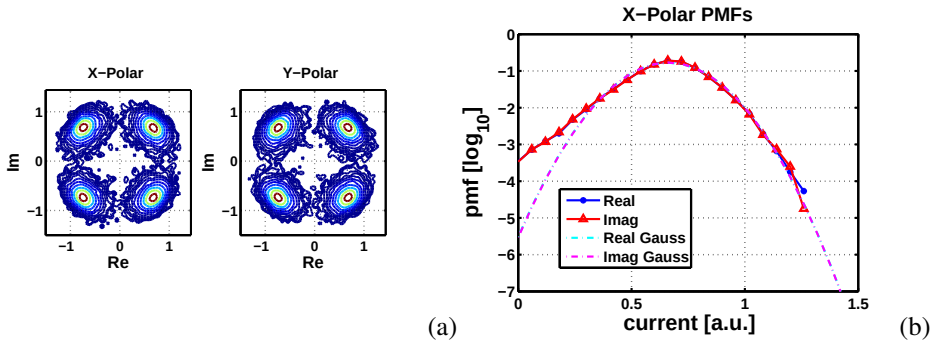


Figure 3.14: Two OOK channels' SOP random and data-aided phase recovery algorithm. PDM-QPSK constellations after propagation along link (a). PMFs of real (blue-circle) and imaginary components (red-triangle) of the one symbol constellation (right-side Fig. 3.7) of X-polarization (b). The two PMFs are superposed.

Fig. 3.14(a) shows elliptical constellations, that seem a sort of a weighted average of the very different effects highlighted in Sec. 3.3.1.1, with the phase noise induced by XPolM more likely to appear than the amplitude noise. The previous theory agrees with this observation, since we showed that XPolM can manifest as crosstalk or phase noise: while in the first case the interfering term induces both amplitude and phase noise, in the second case we do not have amplitude noise. From Fig. 3.14(b) we observe that the PMFs are generally not-Gaussian, despite the elliptical appearance of Fig. 3.14(a) . Note that the probability of exceeding the decision threshold (left-side) is almost two orders of magnitude higher w.r.t. the Gaussian approximation.

18 OOK 50GHz-spaced with random SOPs and random pattern By transmitting the 19 channel-WDM comb we obtained the constellations and the PMFs reported in Fig. 3.15.

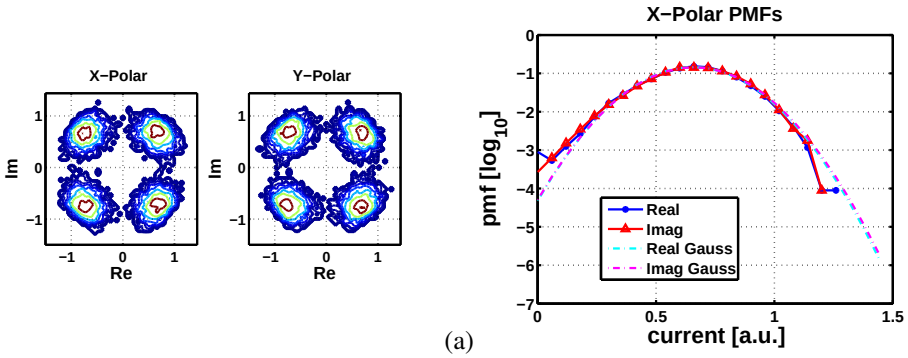


Figure 3.15: 18 OOK channels' SOP random and data-aided phase recovery algorithm. PDM-QPSK constellations after propagation along link (a). PMFs of real (blue-circle) and imaginary components (red-triangle) of the one symbol constellation (right-side Fig. 3.7) of X-polarization (b). The two PMFs are superposed.

The constellations in Fig. 3.15(a) are very similar to Fig. 3.14(a) suggesting that the simplified 3-channel case captures the main properties of the 'more realistic' scenario with 19 channels. Nevertheless the PMFs in this case better match the Gaussian distribution, in particular in the right side of the PMFs. Indeed the probability of exceeding the thresholds is almost the same, even if comparing the Gaussian approximations we observe a little larger variance in the 19-channel case.

Replacing the data-aided phase recovery with the Viterbi and Viterbi algorithm, we obtained the constellations reported in Fig. 3.16, for three different Viterbi and Viterbi number

of taps.

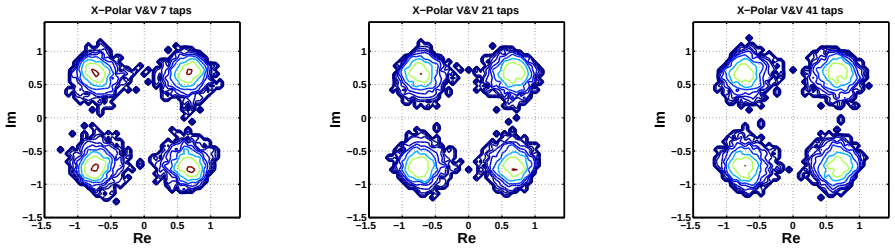


Figure 3.16: 18 OOK channels' SOP random: PDM-QPSK constellations after propagation along link. Phase recovery based on Viterbi and Viterbi algorithm with different number of taps (N): 7(left), 21 (center) and 41 taps (right).

The Viterbi and Viterbi algorithm partially removes the phase noise, yielding more circular constellation distributions. Such a phase noise reduction in more evident for the 41-taps Viterbi and Viterbi.

PDM-QPSK and OOK with same power Until now we set the Pivot SOP only through the OOK SOPs. However if the PDM-QPSK power is comparable with the OOK channels power the Pivot SOP changes both orientation and amplitude along propagation not just because of the temporal variation of the OOK SOPs, but also because of temporal variation of the PDM-QPSK SOP.

We tested the previous 3 channel setups with the same power of 5 dBm for both the PDM-QPSK and the OOK channels. In Fig. 3.17 we report the corresponding received constellations for three special OOK orientations.

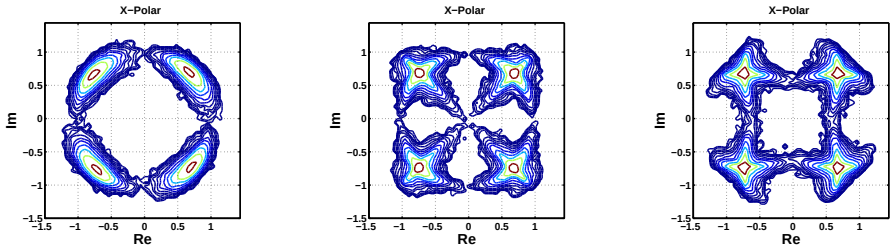


Figure 3.17: Received constellations for OOKs along $\widehat{S_1}$ (left); OOKs along $\widehat{S_2}$ (center); OOKs oriented as Fig. 3.10 (right). All channels have 5 dBm equal power.

It is worth noting that, even thought the Pivot is randomly varying along the Poincaré sphere, the constellations closely resemble those analyzed in Sec. 3.3.1.1, where the Pivot orientation was substantially constant in direction. It turns out that the 5 dBm of the PDM-QPSK channel only slightly perturbs the Pivot (and hence XPolM) since the two OOK channels carry a total of 10 dBm.

3.3.2 NDM link

In this section we repeated the same test of Sec. 3.3.1, for an uncompensated link (NDM). The NDM link had no pre and in-line compensation, while the post-compensating fiber set the overall dispersion to zero.

3.3.2.1 Pivot with fixed orientation

In the 3-channel case, by setting the OOK SOPs along $\widehat{S_1}$ and $\widehat{S_2}$ we obtained the constellations and PMFs reported in Fig. 3.18 and Fig. 3.19, respectively. Both data-aided and 41-taps Viterbi and Viterbi based phase recovery algorithms are analyzed. Since in NDM links the distortion is smaller than in DM links, we increased the OOK power to 15 dBm, while the PDM-QPSK power was always -10 dBm.

Fig. 3.18 (left-column) shows constellations with a bean-like two-dimensional PMF, whose marginal PMFs do not well match the Gaussian distribution, see Fig. 3.18 (right-column). The Viterbi and Viterbi algorithm is able to recover part of the phase distortion, making the clouds more circular and the PMFs matching with Gaussian distribution improves consequently. Furthermore, the Viterbi and Viterbi strongly reduces the probability of exceeding the decision thresholds.

We will discuss more in detail the choice of V&V number of taps in Sec. 3.3.3.

In Fig. 3.19, where the OOKs are aligned with S2, the 41-taps V&V does not introduce any significant improvement, an indication of a weakly correlated phase noise. For both phase recovery methods we can not approximate the PMFs with Gaussian distributions.

3.3.2.2 Randomly oriented Pivot

Using the 19-channels WDM comb, in which the PDM-QPSK probe channel (-10 dBm) is transmitted with 18 randomly oriented OOKs (9 on each side) with a power of 5 dBm each, we obtained the constellations and the PMFs depicted in Fig. 3.20.

In this case the strong channel walk-off induces the Pivot to experience very fast changes along propagation, thus covering the Poincaré sphere like a Brownian motion [2]. Not sur-

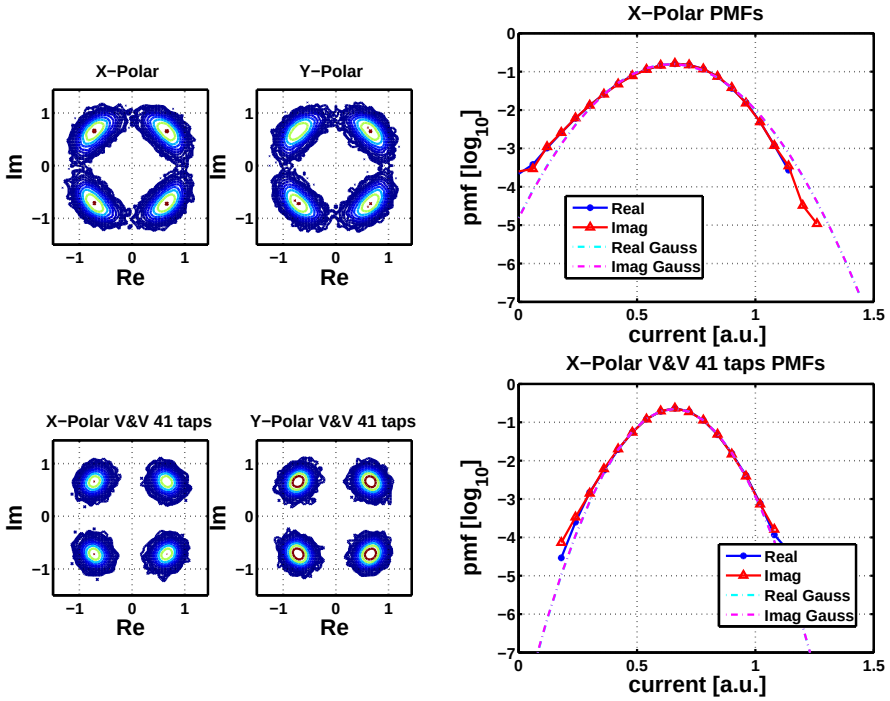


Figure 3.18: NDM link, OOK channels along $\widehat{S_1}$. Data-aided phase recovery (top) and 41-taps Viterbi and Viterbi phase recovery (bottom). PDM-QPSK constellations after propagation along link (left). PMFs of real (blue-circle) and imaginary components (red-triangle) of the one symbol constellation (right-side Fig. 3.7) of X-polarization (right).

prisingly, the PMFs well match a Gaussian distribution, both with data-aided and Viterbi and Viterbi phase recovery. In presence of Viterbi and Viterbi algorithm the probability of exceeding the decision threshold decreases by almost one order of magnitude.

3.3.3 Covariance Functions

In the previous sections we observed that channel walk-off affects XPolM, hence we expect that the received symbols are correlated in time. This information is of great interest, because a temporal correlation can be removed through a properly designed estimator. For instance, the Viterbi & Viterbi algorithm tries to remove as much as possible covariance on the received phase distortion, leading to a white phase-spectrum at its output [50, 51].

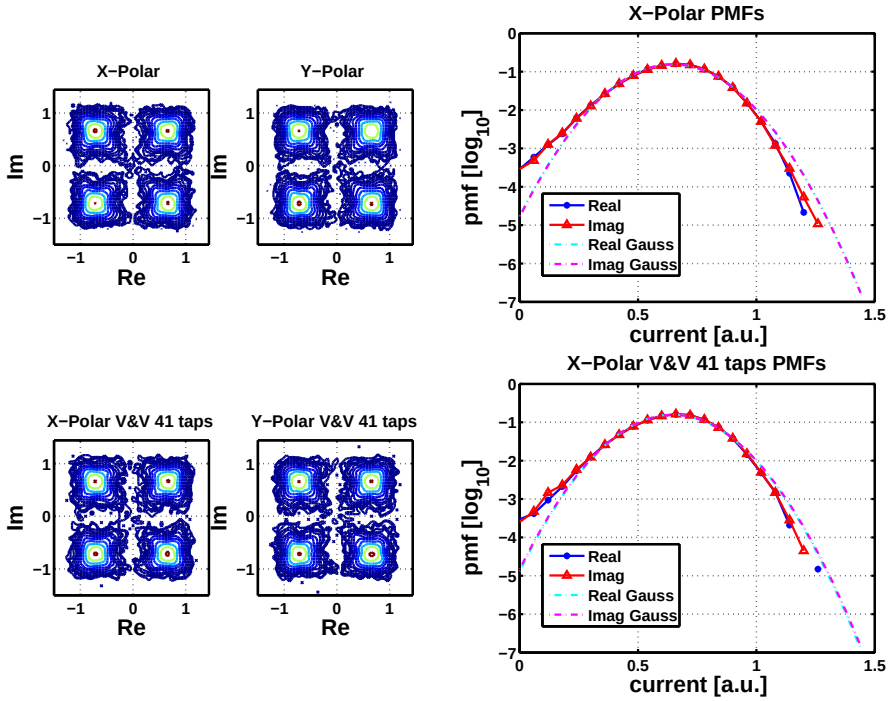


Figure 3.19: NDM link, OOK channels along $\widehat{S_2}$. Data-aided phase recovery (top) and 41-taps Viterbi and Viterbi phase recovery (bottom). PDM-QPSK constellations after propagation along link (left). PMFs of real (blue-circle) and imaginary components (red-triangle) of the one symbol constellation (right-side Fig. 3.7) of X-polarization (right).

Aim of this section is to evaluate the time-covariance of XPolM. The received signal after applying a data-aided or Viterbi and Viterbi phase recovery to $\overrightarrow{A}_{RX} = \left|\overrightarrow{A}_{RX}\right| \exp\{i\phi_{RX}\}$ is, respectively :

$$\begin{aligned} \overrightarrow{A}_{VV} &= \left|\overrightarrow{A}_{RX}\right| \exp\{+i(\phi_{RX} - \phi_{VV})\} \\ \overrightarrow{A}_{DA} &= \left|\overrightarrow{A}_{RX}\right| \exp\left\{+i(\phi_{RX} - \overline{\phi})\right\}. \end{aligned} \tag{3.18}$$

In order to evaluate the XPolM covariance we first data-aided remove the transmitted signal phase ϕ_{TX} from (3.18) and then we remove the average value of the remaining field in order to have a zero mean signal. Summarizing, we measure the covariance of the real and imaginary components of the signal:

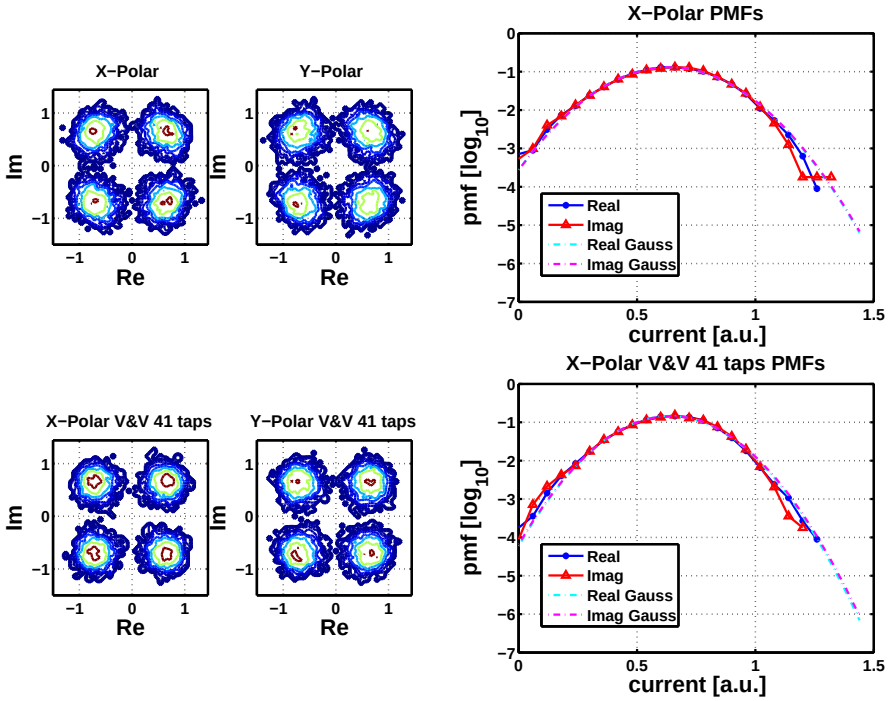


Figure 3.20: NDM link, 18 OOK channels' SOP random. Data-aided (top) and 41-taps Viterbi and Viterbi phase recovery (bottom). PDM-QPSK constellations after propagation along link (left). PMFs of real (blue-circle) and imaginary components (red-triangle) of the one symbol constellation (right-side Fig. 3.7) of X-polarization (right).

$$\overrightarrow{U}_{DA,VV} = \overrightarrow{A}_{DA,VV} \exp\{-i\phi_{TX}\} - \left\langle \overrightarrow{A}_{DA,VV} \exp\{-i\phi_{TX}\} \right\rangle \tag{3.19}$$

which are practically very close to the covariance of the amplitude and phase[10]. This approach corresponds to obtaining the one-symbol constellation reported on the right-side of Fig. 3.21.

We tested the impact of a 19 channels WDM scenario with randomly oriented OOK SOPs, and of a 3 channel scenario with fixed OOK SOP orientations.

[10]This way we are operating in a reference system rotated by 45° w.r.t . the one used in Fig. 3.7

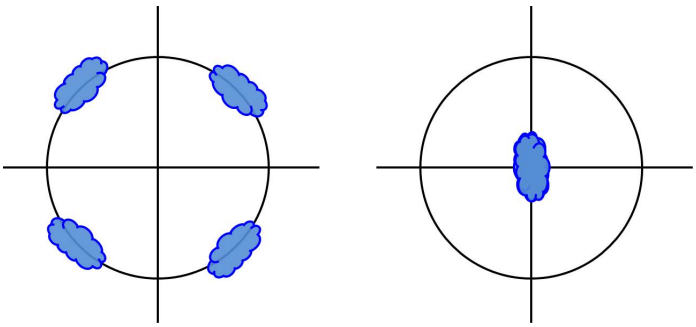


Figure 3.21: Initial constellation (left-side) and one-symbol constellation ($\overrightarrow{U}_{DA,VV}$), after removing the transmitted signal phase and the average value. Hence the real part is the amplitude distortion and the imaginary part is the phase distortion of the received signal.

3.3.3.1 DM link

Fig. 3.22 reports the normalized temporal auto- and cross-covariance functions of the real and imaginary components of $\overrightarrow{U}_{DA}$, for the DM link analyzed in Sec. 3.3.1. In these simulations data-aided phase recovery was performed. Note that each covariance function is normalized to unit peak value one and is discretized with one sample per symbol.

From Fig. 3.22(a) we observe that the auto-covariance temporal width of the real component is almost 100 symbols when OOK SOPs and Pivot are oriented along $\widehat{S}_1$, while in all the other cases the auto-covariance of the real component is similar to a Dirac's delta as for white noise, whatever the orientation of OOKs (a).

Also the imaginary component (Fig. 3.22(b)) takes a non-white behavior when OOK SOPs are oriented along $\widehat{S}_1$. This observation is in agreement with the discussion in Sec. 3.2, since, when the Pivot is oriented as one of the PDM-QPSK subcarriers, the *XPolM* term algebraically adds to the *XPM* term yielding a pure phase noise whose «banana»-like PMF (see Fig. 3.8) correlates the real/imaginary components of the distortion. Another proof of the phase noise correlation in this case is provided by Fig. 3.12, which shows a decreasing constellation spread for decreasing Viterbi and Viterbi number of taps. The covariance of the imaginary component is almost white only when the OOK SOPs are oriented along $\widehat{S}_2$ (red curve in Fig.3.22 (b)). It turns out that, for symmetry, any Pivot in the plane orthogonal to the PDM-QPSK subcarriers, i.e., the ($\widehat{S}_2, \widehat{S}_3$ plane), will induce a white auto-covariance. In this case the XPolM matrix is (3.11), whose real h_{xx} and h_{yy} coefficients do not induce any phase rotation on each subcarrier, while the h_{xy} and h_{yx} coefficients induce a crosstalk

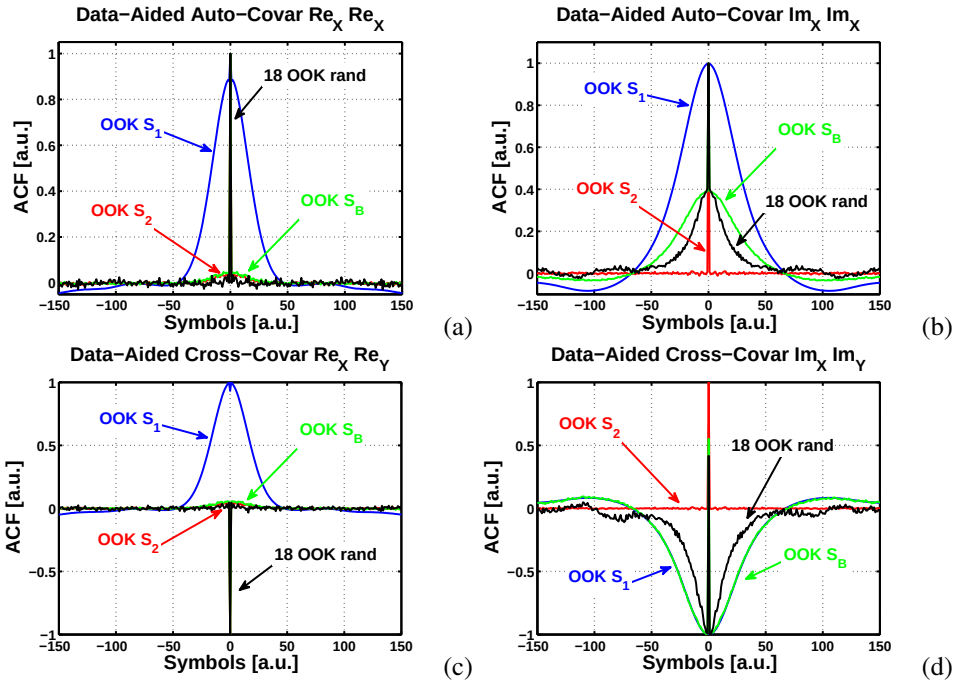


Figure 3.22: DM link, data-aided phase recovery: Auto-covariance function of Real (a) and Imaginary components (b) of X-polarization, for different OOK SOP orientations. Cross-covariance between the Real components of X- and Y-polarization (c) and between Imaginary components of X- and Y-polarization (d), for different OOK SOP orientations.

term that appears as white noise for time shifts larger than the symbol duration T because of independent information symbols on X and Y. The white behavior of XPolM is confirmed by Fig. 3.13, where we observe that the Viterbi and Viterbi algorithm requires a large number of taps to estimate the carrier phase. When the Pivot SOP is not in these two special cases and for example is oriented as $\widehat{S_B}$ in Fig. 3.6, we still have a white contribution to the auto-covariance (green curve in Fig. 3.22(b)) as discussed for the $\widehat{S_2}$ case, while the now complex h_{xx} and h_{yy} coefficients induce phase rotations that correlate the imaginary components. The case marked '18 OOK rand', since it is a combination of random OOK SOPs, has an auto-covariance that is a mixture of the previously analyzed cases.

Fig. 3.22(c) shows a positive cross-covariance with a memory of almost 100 symbols when the Pivot is along $\widehat{S_1}$. Recalling (3.10) a positive phase rotation on X corresponds to a

negative rotation on Y. In both cases, the real component decreases for X and Y, an indication of a positive correlation. In all the remaining cases, the real components of X and Y are strongly negatively correlated at the same time epoch. For instance in the $\widehat{S_2}$ case, see Fig. 3.4 with $\phi = 0$, we observe that the amplitude of the X- and Y-polarizations move in opposite directions. If we transmit the square/triangle symbols on X/Y polarizations, respectively, at the output the square symbol of X decreases its amplitude (see red triangle in the same quadrant) while the triangle symbol magnitude on Y increases (see blue square in the same quadrant).

Fig. 3.23(d) shows a negative cross-covariance of the imaginary components, when the Pivot is along $\widehat{S_1}$ (blue curve). This is in line with eq. (3.10) in which a negative phase rotation on the X-polarization is associated with a positive phase rotation on the Y-polarization. In the $\widehat{S_2}$ case, see Fig. 3.4 with $\phi = 0$, we observe that the phases of the X- and Y-polarizations move in the same directions. If we transmit the cross/triangle symbols on X/Y polarizations, respectively, at the output the cross symbol on X experiences a positive phase shift (see red triangle in the same quadrant), like the triangle symbol on Y (see blue cross in the same quadrant). In the $\widehat{S_B}$ case (green curve in Fig. 3.22(d)), at all time shifts but zero the cross-talk of X is independent of the one of Y after data-aided phase recovery. It turns out that the only contribution to the cross-covariance is due to the phase rotation induced by h_{xx} and h_{yy} as for the $\widehat{S_1}$ case. Please note that the blue and the green curves are superposed at all times but zero. At $t = 0$, since XPolM operates through a unitary matrix, the X- and Y- cross-talks are not independent anymore after data-aided phase recovery, thus impacting significantly the cross-variance, as in the $\widehat{S_2}$ case. Recalling (3.19) and expressing the unitary matrix H (3.9) as

$$H = \mathrm{e}^{\mathrm{i}\psi} \begin{bmatrix} \cos(\zeta)\mathrm{e}^{+\mathrm{i}\psi_1} & \sin(\zeta)\mathrm{e}^{+\mathrm{i}\psi_2} \\ -\sin(\zeta)\mathrm{e}^{-\mathrm{i}\psi_2} & \cos(\zeta)\mathrm{e}^{-\mathrm{i}\psi_1} \end{bmatrix}, \tag{3.20}$$

after data-aided phase recovery, we can write the signal as

$$\begin{aligned} X_{out}(t) &= \mathrm{e}^{+\mathrm{i}\psi_1(\mathrm{t})} \left(\cos(\zeta(\mathrm{t}))\,|\mathrm{X_{in}(t)}|^2 + \sin(\zeta(\mathrm{t}))\mathrm{e}^{+\mathrm{i}(\psi_2(\mathrm{t})-\psi_1(\mathrm{t}))}\mathrm{X_{in}^*(t)Y_{in}(t)} \right) \\ Y_{out}(t) &= \mathrm{e}^{-\mathrm{i}\psi_1(\mathrm{t})} \left(\cos(\zeta(\mathrm{t}))\,|\mathrm{X_{in}(t)}|^2 + \sin(\zeta(\mathrm{t}))\mathrm{e}^{+\mathrm{i}(\psi_2(\mathrm{t})-\psi_1(\mathrm{t}))}\mathrm{Y_{in}^*(t)X_{in}(t)} \right) \end{aligned} \tag{3.21}$$

where $X_{in}(t)$ and $Y_{in}(t)$ are the transmitted signals on X- and Y-polarizations, respectively, while the corresponding received signal components are $X_{out}(t)$ and $Y_{out}(t)$. We have that $Im\{X_{out}(t)\}$ and $Im\{Y_{out}(t-\tau)\}$ are independent at $\tau \neq 0$, since $X_{in}^*(t-\tau)Y_{in}(t-\tau)$ and $Y_{in}^*(t-\tau)X_{in}(t-\tau)$ are uncorrelated because of independent information symbols on X and Y. On the other hand, $Im\{X_{out}(t)\}$ and $Im\{Y_{out}(t-\tau)\}$ are correlated at $\tau = 0$, since $X_{in}^*(t)Y_{in}(t)$

and $Y^*_{in}(t)X_{in}(t)$ are strongly correlated since they are due to the same data symbols. Since the '18 OOK rand' case (black curve) is a combination of random OOK SOPs, its cross-covariance is a mixture of the previously analyzed cases.

Replacing the data-aided phase recovery with the Viterbi and Viterbi, we obtained the normalized temporal auto-covariance functions of the imaginary components of $\overrightarrow{U}_{VV}$ reported in Fig. 3.23, for three number (N) of Viterbi and Viterbi taps: 7, 21 and 41 respectively. The Data-Aided auto-covariance function is reported for a comparison.

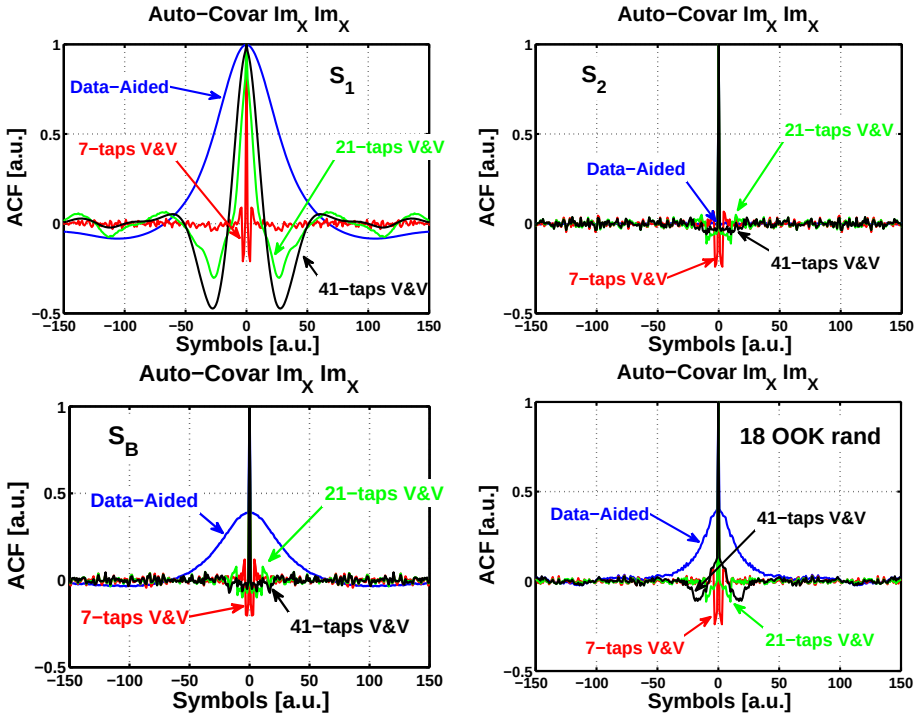


Figure 3.23: DM link, N-taps Viterbi and Viterbi phase recovery: Auto-covariance function of Imaginary components of X-polarization for OOK $\widehat{S_1}$ case, OOK $\widehat{S_2}$ case, OOK $\widehat{S_B}$ case and 18 OOK rand case.

Fig. 3.23 shows that the Viterbi and Viterbi algorithm, working on a window shorter than the transmitted sequence length, shrinks as much as possible the auto-covariance of the imaginary component of the received signal (associated to its phase), leading to a white phase-spectrum at its output. Furthermore we observe that the correlation length of the residual 'colored part' increases for increasing number of taps, like scalar XPM effects, as clearly

shown by the $\widehat{S_1}$ case (top-left figure). Note that the same behavior has been observed for the cross-covariance functions.

3.3.3.2 NDM link

We move now to investigate the NDM link, by reporting in Fig. 3.24 the auto-covariance and the cross-covariance functions obtained for the different OOK orientations.

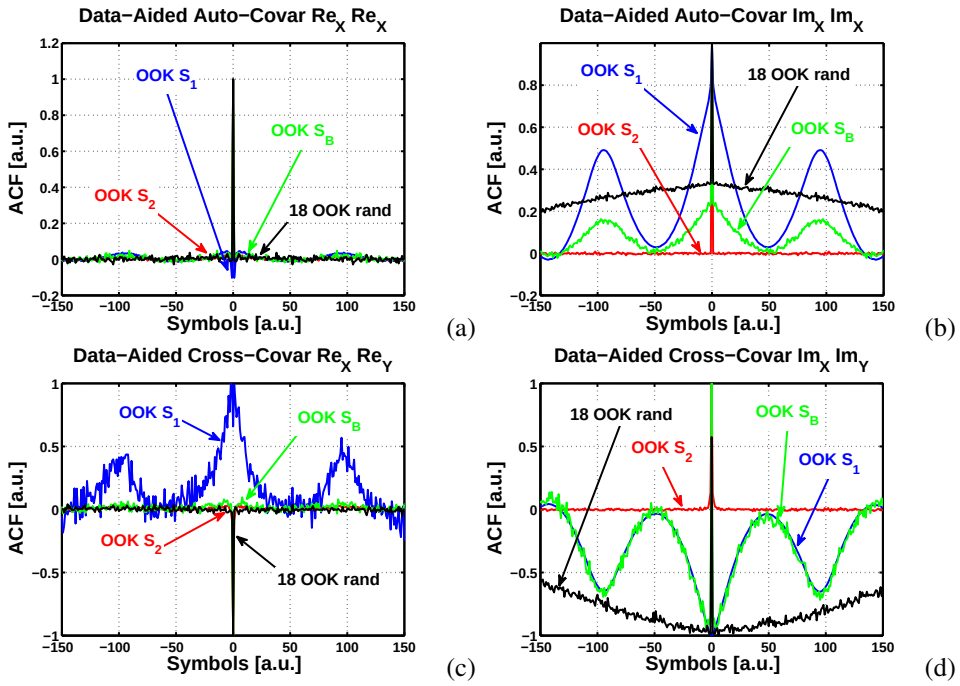


Figure 3.24: NDM link, data-aided phase recovery: Auto-covariance function of Real (a) and Imaginary components (b) of X-polarization. Cross-covariance between the Real components of X- and Y-polarization (c) and between Imaginary components of X- and Y-polarization (d).

Fig. 3.24(a) shows that the auto-covariance of the real components is similar to Dirac's delta for all Pivot (OOK) orientations, thus even with OOKs oriented along $\widehat{S_1}$, which represented an exception in the DM link. All the considerations drawn from Fig. 3.22(b-d) are still true in Fig. 3.24(b-d); the main difference, due to the different transmission link, is the shape of the correlated part.

This shape is strictly connected to the expression of the IM-PM XPM filter ($H_{XPM}(\omega)$), defined in [52]. Fig. 3.25 shows the absolute value of the XPM filter time response ($h_{XPM}(t)$) for the DM link and the NDM link, at channel spacing of 2 nm and 0.4 nm. In the OOK $\widehat{S_1}$ case, where XPolM manifests as an additive XPM effect (see Sec. 3.1), the auto-covariances of the imaginary components in Fig. 3.22(b) and Fig. 3.24(b) are in agreement with the expected covariance functions using the filter h_{XPM}.

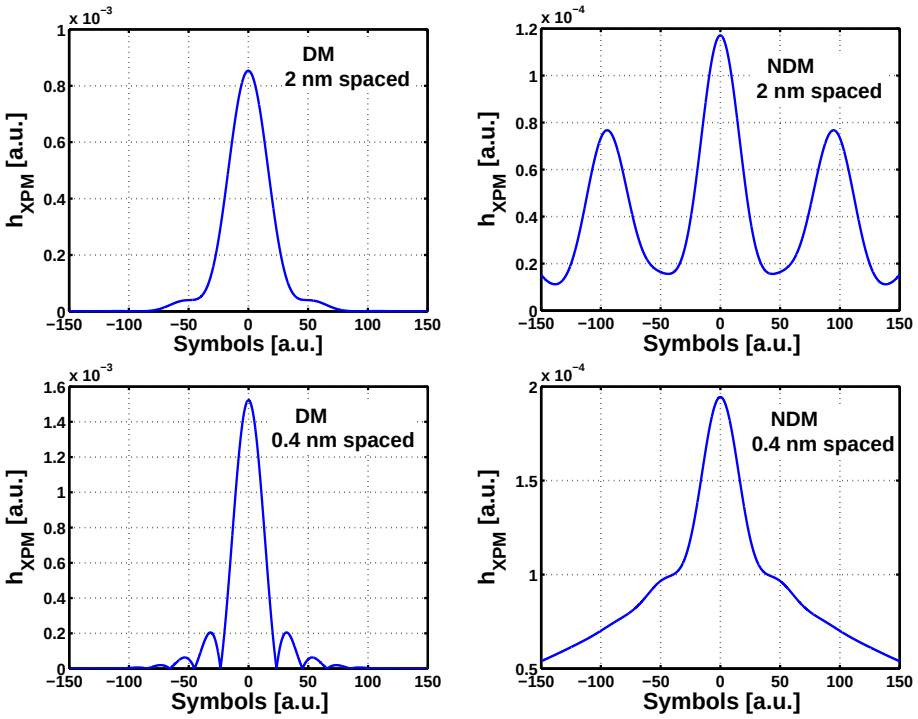


Figure 3.25: Time response of H_{XPM} filter[52] for DM link (right) and NDM link (left), with channel spacing of 2 nm (top) and 0.4 nm (bottom).

The black curves ('18 OOK rand' case) in Fig. 3.22(b) and Fig. 3.24(b) are not a simple combination of the other cases, because each channel experiences a different walk-off w.r.t. the central one, because of a different frequency spacing, as shown in Fig. 3.25. In the assumption of independent channels, the overall auto-covariance induced by the XPM filters is the sum over the autocovariance induced by each channel operating alone. Each contribution follows the same rules as before but using its channel spacing from the reference central

channel. Fig. 3.26 depicts the resulting auto-covariance function for DM link (left) and for NDM link (right) in presence of 18 OOK interfering channels. Despite the shape of Fig. 3.24 (black curve) differs from Fig. 3.26, mainly because the linear XPM filter theory neglects the OOK SOP's impact, the auto-covariance lengths are in agreement both for DM and NDM link.

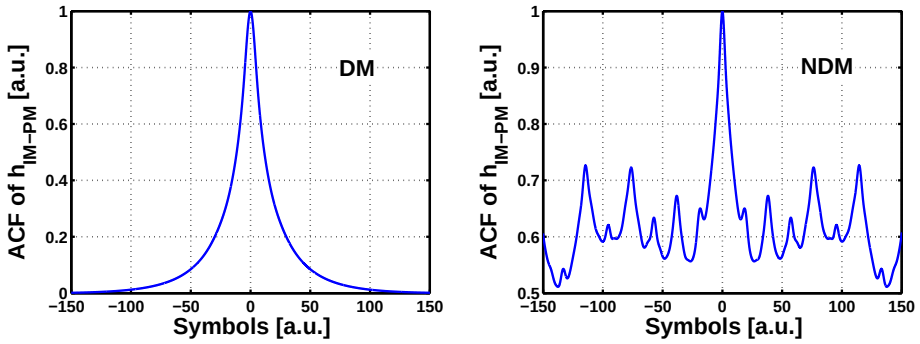


Figure 3.26: Auto-correlation function of h_{XPM} filter[52] for DM link (right) and NDM link (left), for a 19-channel WDM 0.4 nm-spaced.

Similarly to the DM link, replacing data-aided phase recovery with Viterbi and Viterbi, we observed that it reduces the auto-covariance, with an increasing effectiveness for decreasing number of taps, as confirmed by Fig. 3.27, except for the OOK $\widehat{S_2}$ case, which is always a Dirac's delta.

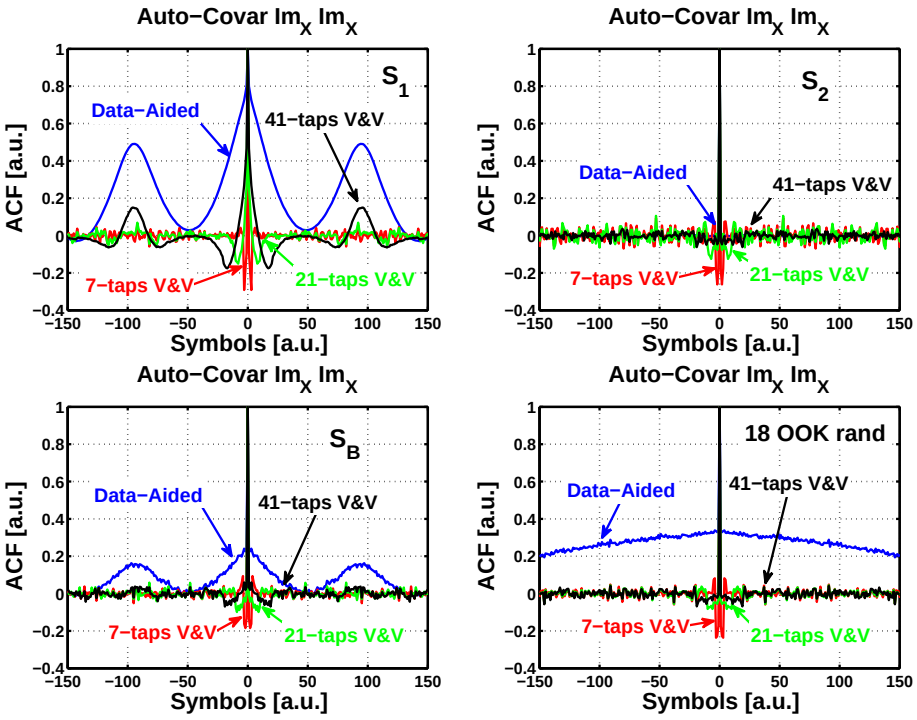


Figure 3.27: NDM link, N-taps Viterbi and Viterbi phase recovery: Auto-covariance function of Imaginary components of X-polarization for OOK $\widehat{S_1}$ case, OOK $\widehat{S_2}$ case, OOK $\widehat{S_B}$ case and 18 OOK rand case.

Chapter 4

Nonlinearities assessment in Long-haul Coherent Transmissions

As discussed in Ch. 2, the combination of DSP-based coherent detection and Polarization Division Mltiplexing (PDM), thanks to a high spectral efficiency and a remarkable resilience against linear impairments (PMD and GVD), increase the total system capacity without incurring performance degradation due to a wider spectrum. However, optical cross-channel nonlinear effects can strongly impair PDM-QPSK channels, as shown by both experiments [53, 54] and simulations [47, 55]. With wavelength division multiplexing (WDM), the cross-nonlinearities make neighboring channels interact depending on their power and state of polarization (SOP). The last is of particular concern in PDM systems since they are more sensitive to cross-polarization modulation (XPolM), defined in Sec. 3.1. Experiments showed that XPolM depends on a larger number of channels than cross-phase modulation (XPM) [56], while Winter *et al.*[2] provided a model that successfully measured the degree of polarization degradation in presence of XPolM. Despite these results, it is still not clear when the bit error rate (BER) is dominated by XPolM and how XPolM relates to the other relevant nonlinear effects, such as XPM and self-phase modulation (SPM).

The aim of this chapter is to fill the gap, by providing a systematic simulation study of system performance where each nonlinear effect acts individually. Our goal is to determine for the polarization-division-multiplexed quadrature-phase-shift-keying (PDM-QPSK)

and polarization-division-multiplexed binary-phase-shift keying (PDM-BPSK) modulation formats which is the dominant nonlinearity. We address such a goal by massive use of simulation in a wide range of system setups, by varying

- pulse shaping: non-return-to-zero (NRZ), aligned- return-to-zero (aRZ) and interleaved-return-to-zero (iRZ);
- transmission link: dispersion-managed (DM) with different RDPS values and dispersion un-managed (NDM);
- transmission fiber: single-mode-fiber (SMF) and LEAF fiber,

in both hybrid (PDM-QPSK – OOK) and homogeneous systems (PDM-QPSK and PDM-BPSK). In iRZ pulse shape the polarization tributaries are 50%-RZ shaped and delayed by half a symbol time [57].

4.1 PDM-QPSK / OOK hybrid systems

The increasing demand for capacity requires upgrading the current 10 Gb/s WDM networks to 100 Gb/s. The coherent PDM-QPSK solution is a promising candidate to this aim. One possible upgrading scenario can be the progressive insertion of several 100 Gb/s PDM-QPSK "upgraded" channels into an infrastructure originally designed for "legacy" NRZ OOK channels at 10 Gb/s, with a dense channel spacing (50 GHz), relying on dispersion managed (DM) links. In such a hybrid system, the performance of the upgraded channel is degraded by the penalties induced by legacy channels through cross-channel nonlinear effects, both scalar (XPM) and vectorial (XPolM). In the last few years, several works have appeared in the literature demonstrating how cross-channel effects are enhanced in such hybrid systems [54, 58, 59]. These studies demonstrated that the scalar XPM is the dominant nonlinear effect in such systems. Nevertheless other experimental works highlighted that in PDM systems a larger number of interfering channels than in single polarization transmission systems are necessary to correctly evaluate the cross-nonlinearities [53, 56]. The authors suggested that cross-polarization modulation is the most likely candidate to induce such penalties.

In this section, we numerically investigate cross-channel nonlinear effects induced by legacy 10 Gb/s OOK channels on an upgraded 112 Gb/s aRZ-PDM-QPSK channel. We show, using the XPolM definition in Sec. 3.1, the channel spacing value beyond which the XPolM induces stronger impairments than XPM. We also clarify the rule of the Viterbi and Viterbi

algorithm (V&V) against both XPM and XPolM. Finally we wish to investigate the dependence of XPolM on the OOK interfering channels SOPs. In particular, we want to find the best and worst SOPs, i.e., those for which the dependence of performance on channel spacing, as pointed out in [53, 56], is minimized.

4.1.1 Numerical setup

For our analysis we used the system setup in Fig. 4.1, in which we transmit a reference 112 Gb/s (100 Gb/s plus 12% of FEC overhead) aRZ-PDM-QPSK channel, with 50% duty cycle, surrounded by 10 Gb/s NRZ-OOK channels synchronous in time. During multiplexing, each channel was filtered by a 3^{rd}-order Supergaussian optical filter of 50 GHz of bandwidth. All measurements were performed over the central PDM-QPSK channel at 1550 nm. The central channel was discretized with 1120 symbols, 30 points per symbol, and modulated using random quaternary sequences with different seeds. The OOK channels were discretized with 400 symbols, 84 points per symbol, and modulated using random binary sequences. We chose such values in order to have the same size of the Fast Fourier Transform (FFT) window.

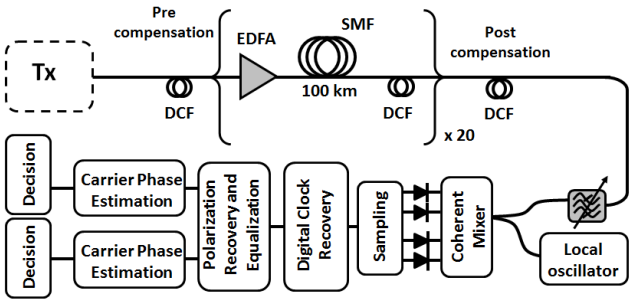


Figure 4.1: Numerical setup: hybrid PDM-QPSK — OOK transmission system.

We used two different transmitters, in which a PDM-QPSK central probe channel is surrounded by: i) a variable number of 50-GHz spaced OOK channels or ii) two OOK channels with variable channel spacing. We control the OOK SOPs to evaluate their impact on the PDM-QPSK Q-factor.

The optical link is dispersion managed, composed of 20×100 km spans of standard SMF ($D = 17$ ps/nm/km @1550 nm, $A_{eff} = 80\ \mu m^2$, $n_2 = 2.5 \cdot 10^{-20}$ m^2/W, $\alpha = 0.2$ dB/km, $D^{'} = 0$ ps/nm^2/km,$\gamma = 1.3$ W^{-1}/km^{-1}). The DM link, which is a typical link for direct detection systems, had a pre-compensation whose dispersion was chosen according to the straight line

rule [48], 30 [ps/nm/span] in-line residual dispersion and a post-compensating fiber whose value set the overall dispersion to zero. All compensating fibers were supposed to be purely linear. Fiber birefringence was modeled by 50 discrete random waveplates per span, with zero PMD. The propagation, through the Manakov equation, was solved with the Split Step Fourier Method, neglecting FWM [22]. Before detection, we perfectly compensated optical linear impairments, i.e., GVD and birefringence (PMD was set to zero) by applying the inverse Jones matrix of the optical line, which allows us to focus entirely on the extra penalty coming from the interplay of linear and nonlinear distortions along the link [60].

At the receiver side an optical 3^{rd}-order Supergaussian filter with a bandwidth of 50 GHz selected the central channel of the comb. The PDM-QPSK signal was detected by means of a DSP based coherent receiver [49] with a noise-figure of 19 dB. The received noise was assumed to be white, i.e., nonlinear phase noise was neglected[25]. The receiver includes: mixing with an ideal local oscillator (perfectly synchronized with the signal carrier), a 3^{rd}-order Supergaussian low pass filter of bandwidth 15 GHz, sampling, polarization demultiplexing, carrier phase estimation using the Viterbi and Viterbi algorithm (V&V) with a variable number of taps, decision and, finally, differential decoding. Two different algorithms for polarization demultiplexing were used: the Constant Modulus Algorithm (CMA) [61] with 13 taps and a least-mean-square algorithm, that we will call DARMAVG (Data Aided Removal of Average) in the following [62]. It is important to note that DARMAVG recovers the average polarization rotation without any kind of equalization, while CMA also performs linear impairments equalization. The length of V&V algorithm can be optimized to reduce the cross-channel impairments in the nonlinear regime, as we will show in the following.

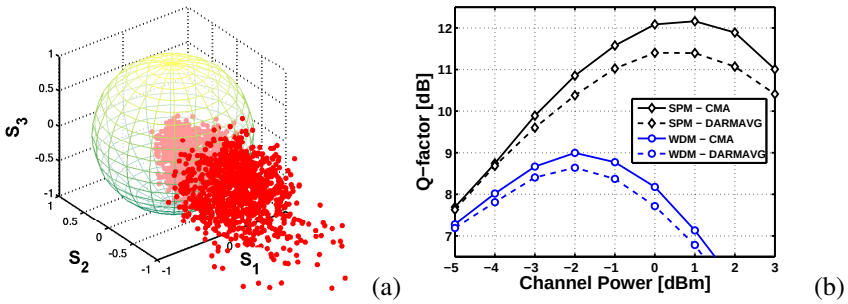
We measured the BER through the Monte Carlo algorithm by counting 100 errors, and then converting the estimated BER to Q-factor. Each simulation was repeated 25 times, in order to correctly sample the variability due to random symbol pattern and launched signal SOPs. All simulations were performed with the open source software Optilux, developed by our research group [3].

4.1.1.1 CMA and DARMAVG-based polarization recovery

Usually the standard coherent receiver performs the polarization recovery through the CMA algorithm[49], which tries to drive the output signal to one having a constant amplitude. This algorithm is simple to implement, but its main drawback is the possibility of convergence to a local minimum [63].

Fig. 4.2(a) reports the PDM-QPSK SOP when the CMA converges to a local minimum. We performed 100 learning loops, in absence of ASE, to bring the CMA to convergence.

Unfortunately sometimes, also for higher number of learning loops, CMA returns a SOP like Fig. 4.2(a), yielding a low Q-factor, and consequently the data were rejected.



(a) (b)

Figure 4.2: 19-channel (18 OOKs and central PDM-QPSK) WDM comb: PDM-QPSK SOP after CMA polarization recovery algorithm in presence of a mal-convergence to a local minimum (a). PDM-QPSK Q-factor vs channel power for both CMA after data rejection (solid line) and DARMAVG-based polarization recovery (dotted line).

The DARMAVG polarization recovery allows us to speed up the simulation and to avoid convergence issues. Fig. 4.2(b) shows the PDM-QPSK Q-factor versus channel power (OOK power is 1 dB lower) for both CMA (after data rejection) and DARMAVG- based polarization recovery. The aRZ-PDM-QPSK Q-factors for single channel transmission, (labeled SPM), and WDM transmission, in which a central PDM-QPSK channel is surrounded by 18 10Gb/s OOK channels (9 on each side), spaced by 50 GHz (la belled WDM) are reported. Since DARMAVG does not perform linear equalization, we observe some penalties w.r.t. CMA-based polarization recovery. The maximum difference is almost 0.7 dB in single channel and almost 0.4 dB in WDM transmission. In Fig. 4.2(b) the number of V&V taps is optimized, i.e., the best Q-factor for each PDM-QPSK is reported.

4.1.2 Impact of Viterbi and Viterbi number of taps on system performance

In this section we investigate the impact of the number of taps of V&V carrier phase estimation on the performance of a PDM-QPSK. We simulated a WDM comb, spaced by 50 GHz, formed by a central 112Gb/s aRZ-PDM-QPSK channel (-1 dBm) surrounded by nine 10Gb/s OOK channels (-2 dBm) on each side. The OOK SOPs were randomly oriented and the PDM-QPSK Q-factor was averaged over 25 pattern/SOP realizations.

Fig. 4.3 reports the Q-factor versus V&V number of taps, both for CMA and DARMAVG polarization recovery algorithm.

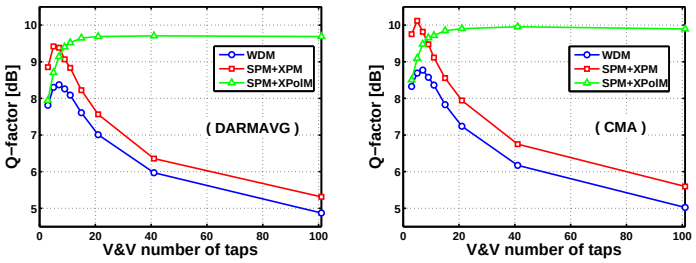


Figure 4.3: 19-channel WDM comb spaced by 50 GHz: aRZ-PDM-QPSK Q-factor versus V&V number of taps for two polarization recovery algorithms: DARMAVG (left) and CMA (right).

From Fig. 4.3 we note that the best number of V&V taps, in absence of XPolM (red-square), is almost 5 for both polarization recovery algorithms, yielding the highest Q-factor in presence of SPM and XPM. This is in agreement with the results of Sec. 3.3, which showed a decreasing polar spread for decreasing number of taps when XPolM acts as an additional XPM term. When SPM and XPolM (diamond) are turned ON the maximum Q-factor is reached for V&V at more than 21 taps. Indeed in Sec. 3.3 we observed a decreasing polar spread for increasing number of V&V taps when the OOK SOPs are randomly oriented (Fig. 3.16). In presence of all nonlinear effects a 7-tap V&V is the trade-off between XPM and XPolM that allows to obtain the best Q-factor. Such a value corresponds to the cross-point between the red and green curves, which means a comparable XPolM and XPM-induced penalty. Hence, reasoning on Fig. 4.3 we can not establish which is the dominant cross-nonlinear impairment in hybrid systems, since it depends on the number of taps of the phase recovery algorithm.

Fig. 4.4 shows the PDM-QPSK Q-factor versus PDM-QPSK power for three phase recovery filter lengths, confirming the previous assertion. The DARMAVG polarization recovery was used, but similar curves have been obtained also using CMA algorithm. The OOK power was again 1 dB lower than the PDM-QPSK one.

The results reported in Fig. 4.4 are in line with [58, 59], in which the authors showed that the scalar XPM is the dominant nonlinear effect (since they used 27-tap and 10-tap phase recovery algorithm, respectively, reducing XPolM at the expense of XPM).

We repeated the same test of Fig. 4.3 reducing the relative impact of XPM. To this aim we

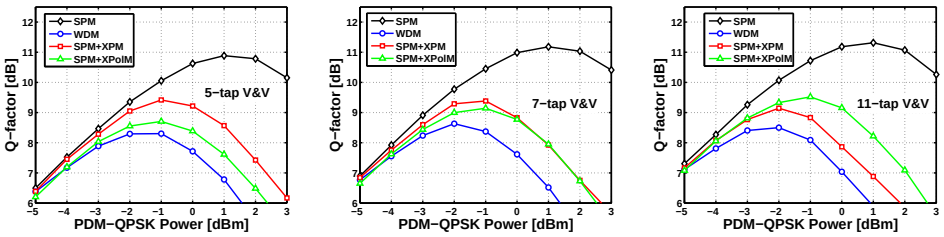


Figure 4.4: 19-channel WDM comb spaced by 0.4 nm: PDM-QPSK Q-factor versus power, for different V&V number of taps and DARMAVG polarization recovery.

used two OOK channels (one on each side) spaced 2nm from the central PDM-QPSK channel. To have significant nonlinear penalties, PDM-QPSK and OOK powers were increased to 0 dBm for PDM-QPSK and 1 dBm for OOKs. Curves of Q-factor versus V&V number of taps, both for CMA and DARMAVG, are reported in Fig. 4.5.

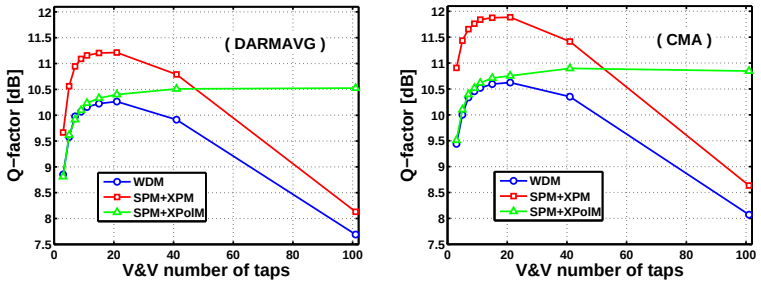


Figure 4.5: 3-channel WDM comb spaced by 2 nm: Q-factor versus V&V number of taps for two polarization recovery algorithms: DARMAVG (left) and CMA (right).

Unlike Fig. 4.3, we observe that XPolM is stronger than XPM until 50 taps of V&V algorithm. Since XPM is almost negligible at this channel spacing, the best Q-factor in presence of all nonlinearities is obtained for 21-tap V&V.

Fig. 4.6 shows that a long V&V filter can make XPM stronger than XPolM, hence a careful optimization of its number of taps is mandatory. Again, we obtained similar behaviors with the CMA polarization recovery (not reported here).

To speed up the simulation and to avoid convergence issues, for all the following analysis we apply the DARMAVG algorithm to perform the polarization recovery.

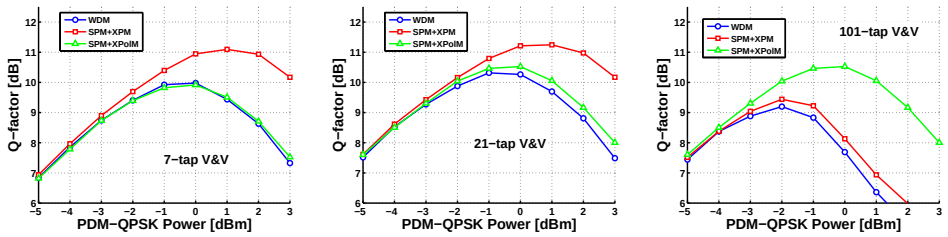


Figure 4.6: 3-channel WDM comb spaced by 2 nm: PDM-QPSK Q-factor vs power, for different number of V&V taps and DARMAVG polarization recovery.

4.1.3 Number of WDM channels/channel spacing to correctly assess the cross-nonlinearities

In this section we investigate the needed number of interfering channels to be taken into account to correctly evaluate XPolM (Sec. 3.1) and XPM. In addition we show the channel spacing value beyond which the OOK-induced XPolM becomes stronger than the OOK-induced XPM and the value beyond which the XPolM-induced penalty can be assumed negligible.

We start investigating XPM and XPolM-induced penalty coming from OOK channels on the central PDM-QPSK in a WDM grid spaced by 50 GHz by increasing the number of WDM channels. The OOK power was set to -1 dBm, while the PDM-QPSK power was set to 0 dBm. In Fig. 4.7 the PDM-QPSK Q-factor versus number of WDM channels is reported for two numbers of V&V taps: 7 and 21. We recall that a 7-tap V&V allows to maximize the PDM-QPSK Q-factor in a 0.4 nm-spaced grid (Fig. 4.3), while 21-tap are the optimum for a 2 nm-spaced grid (Fig. 4.5), respectively. Both graphs are obtained using the DARMAVG polarization recovery.

Fig. 4.7 shows that the XPM-induced penalty (red-square) still remains constant for a number of channels greater than 7 for both V&V lengths, in accord with the walk-off window theory [52]. On the contrary, XPolM-induced penalty (green-triangle) increases for increasing WDM channels until 35 WDM channels. We could not simulate more than 35 channels for time reasons. We can conclude that for the 7-channel WDM case all the penalties coming from XPM are taken into account, while to correctly take into account the XPolM penalty we must simulate a much larger number of channels, in agreement with [53, 56]. Even if the 7-tap V&V minimizes the XPM-induced penalty, we observe that XPM is the dominant cross-nonlinerity until almost 19 channels. When we minimize the XPolM impact (21-tap V&V) we observe that XPM is the dominant cross-nonlinearity up to 35 channels. It is worth to note

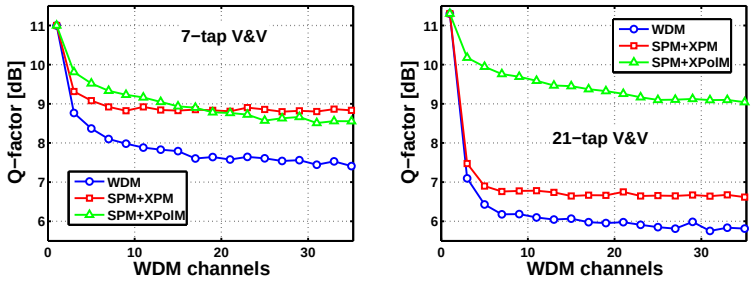


Figure 4.7: PDM-QPSK Q-factor versus number of WDM channels for 7-tap (left) and 21-tap V&V algorithm (right) in WDM case (blue-circle), SPM+XPM case (red-square) and SPM+XPolM case (green-triangle).

that the 7-tap V&V maximizes the 35-channel WDM performance (the Q-factor becomes almost 7.5 dB instead of almost 6 dB for the 21-tap case). This means that the PDM-QPSK performance in a hybrid system can be optimized by setting the V&V taps so as to reduce XPM.

Now we investigate the penalties of OOK-induced XPolM and XPM at each channel spacing. To this aim we emulate a WDM comb formed by a 112 Gb/s PDM-QPSK central probe channel surrounded by two 10 Gb/s OOK channels (one on each side), with randomly oriented SOPs and whose equal spacing $\Delta\lambda$ to central PDM-QPSK is varied from 0.4 nm to 11.6 nm. Note that to investigate the impact of the OOK at high channel spacing we adapted the nonlinear SSFM step accordingly to App. D.1. The PDM-QPSK power was set to 1 dBm, while OOK power was set to 0 dBm. In Fig. 4.8, we report the Q-factor versus channel spacing obtained when all NL effects are present (blue-circle), when SPM and XPM are turned on (red-square) and, finally, when SPM and XPolM are present (green-triangle). Again we perform the simulation for two numbers of V&V taps: 7 and 21. As shown in Sec. 4.1.2, the number of V&V taps that maximizes the system performance depends on the channel spacing and generally speaking it increases while increasing channel spacing, since at small channel spacing XPM is stronger than XPolM, and viceversa at large spacings.

Fig. 4.8 shows that, independently of V&V number of taps, the XPM-induced penalty become negligible at channel spacing larger than 1.2 nm, in agreement with Fig. 4.7; while XPolM-induced penalty is still present up to a channel spacing of almost 6 nm. Furthermore, this figure shows that the cross point between 'SPM+XPM' and 'SPM+XPolM' curves depends on the V&V taps: it is almost 0.6 nm at 7-tap and almost 1 nm at 21-tap, respectively.

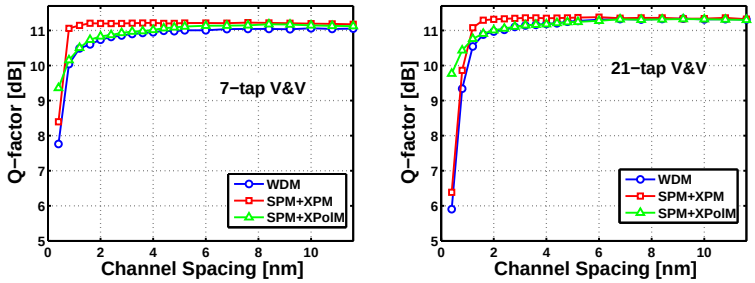


Figure 4.8: OOK random case: PDM-QPSK Q-factor versus channel spacing for 7-tap (left) and 21-tap V&V algorithm (right) in WDM case (blue-circle), SPM+XPM case (red-square) and SPM+XPolM case (green-triangle).

Generalizing, we can conclude that the penalties coming form far channels are mainly due to XPolM, as stated in [53], while the penalties coming from the neighboring channels are mainly due to XPM[58, 59]. Both cross-nonlinearities decrease with increasing walk-off, here represented in terms of channel spacing, but with a different speed. A deeper investigation on the scaling rules of XPM and XPolM with the walk-off will be given in Sec. 4.2.3.

4.1.4 Impact of OOK SOPs

In this section we investigate the impact of OOK SOPs on the PDM-QPSK Q-factor, i.e., we analyze how the different PDM-QPSK constellations shown in Sec. 3.3.1.1 are connected to the system performance. We start showing the PDM-QPSK Q-factor vs. channel spacing in OOK $\widehat{S_1}$, $\widehat{S_2}$ and $\widehat{S_A}$ cases, for both 7-tap and 21-tap V&V algorithm. We recall that the transmitted PDM-QPSK signal employs sub-carriers with linear horizontal and vertical polarizations ($\pm\hat{S_1}$), so that the SOPs of its time samples are those shown on the Poincarè sphere in App. B.

Fig. 4.9 reports the PDM-QPSK Q-factor versus channel spacing with OOK aligned with $\widehat{S_1}$ (top-line), $\widehat{S_2}$ (middle-line) or $\widehat{S_A}$ (bottom-line). The two figures in top-line show that in the $\widehat{S_1}$ case XPM is stronger than XPolM at all channels spacings and for both 7-tap and 21-tap V&V. This result is in agreement with the intuition since in this case XPolM acts as an additive XPM term (Sec. 3.2), but it is three times smaller than the XPM term (eq. 3.1). From middle and bottom lines of Fig. 4.9 we observe that in the $\widehat{S_2}$ and $\widehat{S_A}$ cases at high channel spacing the XPolM is stronger than the XPM, since the walk-off is more effective in reducing scalar (one-dimensional) effects like XPM than the (two-dimensional) diffusion over the

Poincaré sphere induced by XPolM. Summarizing, as stated in Sec. 4.1.3, the XPolM experiences a larger dependence on channel spacing w.r.t. XPM, with the remarkable exception of the $\widehat{S_1}$ case.

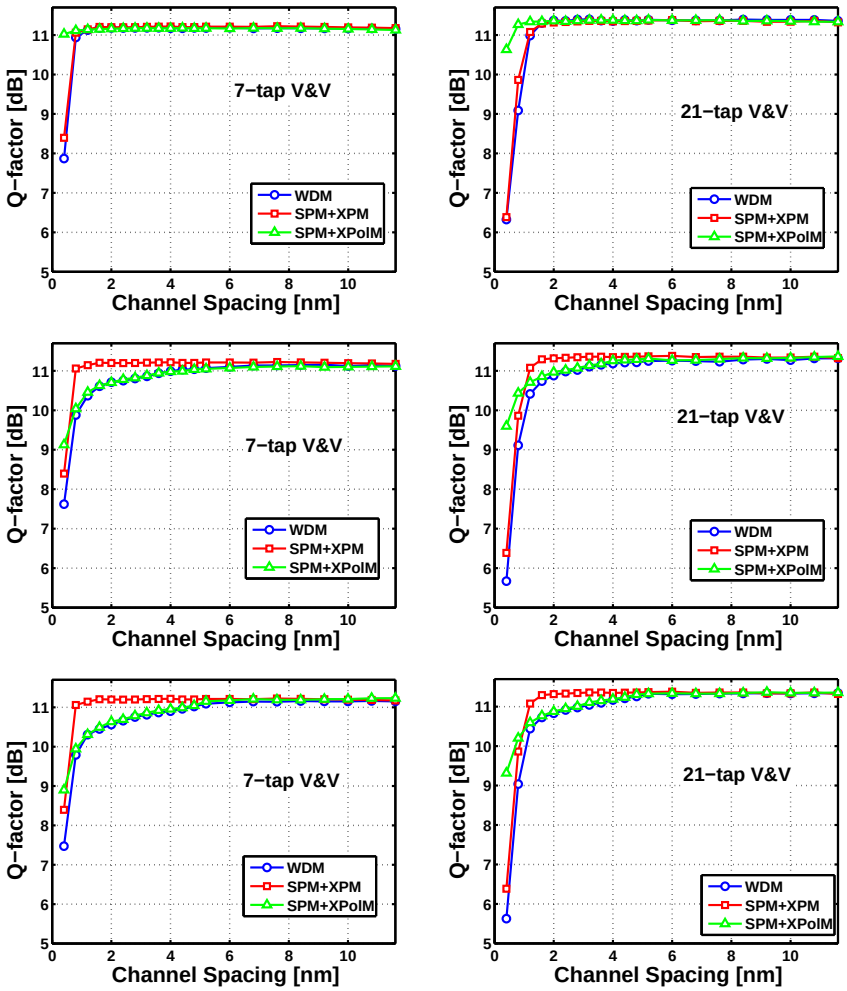


Figure 4.9: PDM-QPSK Q-factor versus channel spacing with OOK aligned with $\widehat{S_1}$ (top-line), $\widehat{S_2}$ (middle-line) or $\widehat{S_A}$ (bottom-line). 7-tap (left) and 21-tap V&V algorithm (right) in WDM case (blue-circle), SPM+XPM case (red-square) and SPM+XPolM case (green-triangle).

To better compare the impact of OOK SOPs on the XPolM-induced penalties we superpose the corresponding Q-factor versus channel spacing curves obtained for all the OOK SOPs under analysis in Fig. 4.10. The general random case is reported for comparison. We did not report the cases SPM+XPM since they are independent of interfering channel's SOPs.

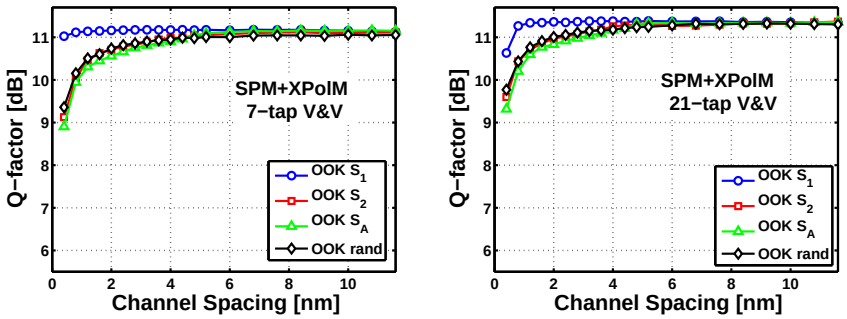


Figure 4.10: PDM-QPSK Q-factor versus channel spacing in SPM+XPolM case, for 7-tap (left) and 21-tap V&V (right).

Fig. 4.10 shows that all cases but the $\widehat{S_1}$ yield similar performance, an indication that XPolM diffused the PDM-QPSK SOP uniformly over the Poincaré sphere almost indipendently of the starting OOK SOP. The $\widehat{S_1}$ case is a special case for the reasons discussed previously.

Given the dependence of performance on the OOK SOPs, we now test all possible OOK SOP orientations on the Poincarè sphere. To this aim we simulate a WDM comb in which the reference PDM-QPSK channel (1 dBm) is surrounded by two OOK channels (3 dBm) having the same SOP (i.e., they are aligned in polarization) and equally spaced by $\Delta\lambda = 2nm$ from the central channel.

Fig. 4.11 reports the PDM-QPSK Q-factor (averaged over 5 different random patterns on the Poincaré sphere) as a function of the OOK SOP. Each point of the Poincarè sphere (right-side) indicates the PDM-QPSK Q-factor when the two aligned OOK have such a SOP; in the left side such a spherical plot is translated into a 2-dimensional (azimuth and ellipticity of OOK SOPs) graph. The three lines in the figure contain the results obtained for 7-tap, 21-tap and 101-tap V&V carrier phase estimation, respectively. For this analysis all nonlinearities are turned on (WDM case), even if at such channel spacing the XPM-induced penalty is negligible, as shown in Fig. 4.5.

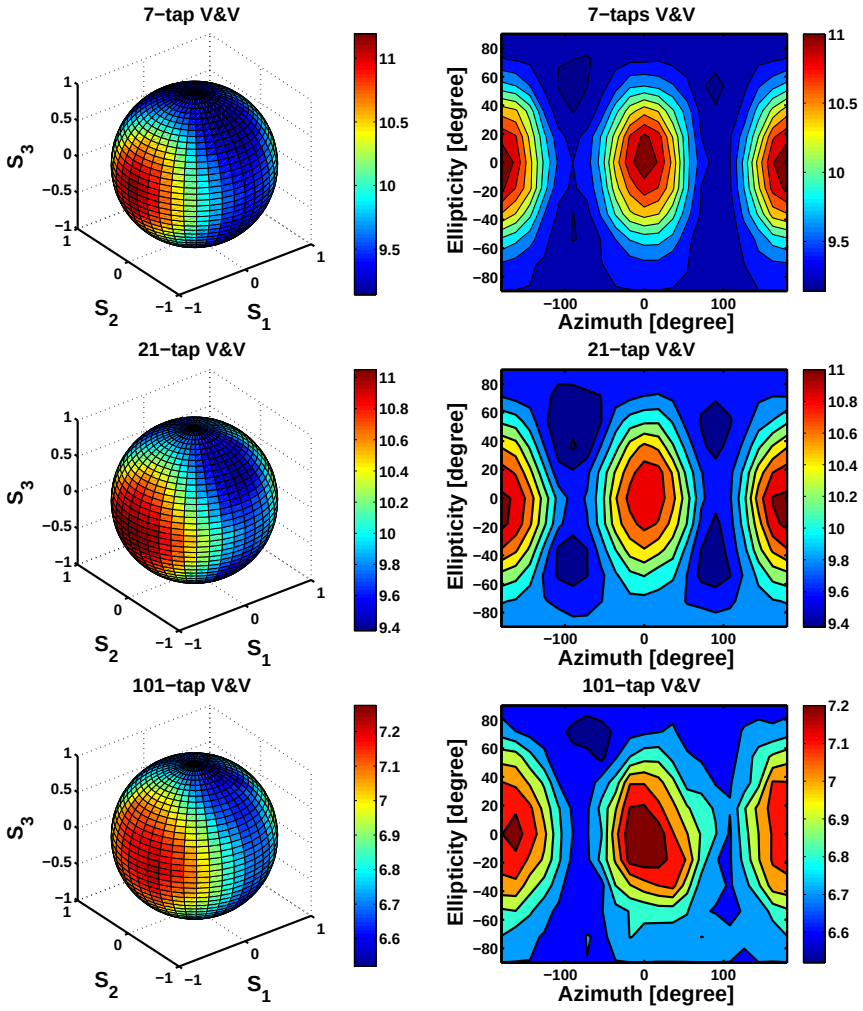


Figure 4.11: PDM-QPSK Q-factor versus common OOK SOPs, in presence of all nonlinearities. Each point of the Poincarè sphere (left-column) and of the azimuth/ellipticity plane (right-column) represents the PDM-QPSK Q-factor for that OOK SOP. Three numbers of V&V taps are investigated: 7, 21 and 101.

Fig. 4.11 shows that aligning the OOKs along $\widehat{S_1}$ is the best choice, since we obtained the highest Q-factor at every V&V number of taps. The XPolM-induced penalty is thus minimized when the interfering OOK channels have the same polarization as one of the two PDM

sub-channels ($\widehat{S_1}$), i.e., when the XPolM acts as an additional XPM term (Sec. 3.1). With this OOK orientation the relative phase of the $\widehat{S_1}$ (X) and $-\widehat{S_1}$ (Y) PDM-QPSK subcarriers have been changed, i.e., the component that is parallel with the OOK channels ($\widehat{S_1}$) experiences a phase shift that is twice as large as the orthogonal component ($-\widehat{S_1}$). Indeed, as shown in Fig. 3.8 the PDM-QPSK SOP still remains in the ($\widehat{S_2}$, $\widehat{S_3}$) plane, which means that there is no cross-talk between the two subcarriers.

When the OOK SOPs are in the ($\widehat{S_2}$, $\widehat{S_3}$) plane the PDM-QPSK experiences a constant Q-factor which is lower than the Q-factor observed with OOK aligned along $\widehat{S_1}$. The penalty is in agreement with the constellation shown in Fig. 3.9 and Fig. 3.10. Such an amplitude variation, which we called cross amplitude modulation (XAM) in Sec. 3.2, is the effect of the crosstalk between the two subcarriers. From Fig. 4.11 we can conclude that the subcarrier crosstalk induces a Q-factor penalty larger than the phase shift-induced one, as also observed in [64]. We can also observe a little smaller Q-factor (less than 0.5 dB) when the OOK has an azimuth of $\pm 90°$ and ellepticity of $\pm 45°$, which contains the OOK $\widehat{S_A}$ case (Fig. 3.6). The reason is connected to the PMFs in Fig. 3.9 and Fig. 3.10, that showed a slightly higher probability to exceed the decision threshold in OOK oriented as $\widehat{S_A}$ case, w.r.t. OOK along $\widehat{S_2}$ cases. The difference between the best and worst Q-factor decreases for increasing number of V&V taps. This is due to the higher impact of the XPM that, since it is independent of polarization, induces an homogeneous Q-factor decreasing. Finally we note that most of the sphere shows an average Q-factor similar to the worst case, in agreement with the fact that the average Q-factor in "OOK rand" case shown in Fig. 4.10 is very close to the worst cases ($\widehat{S_2}$ and $\widehat{S_A}$). This is more evident at small number of V&V taps.

4.2 112 Gb/s PDM-QPSK homogeneous systems

In this section, we numerically investigate the nonlinear penalty in a PDM-QPSK homogeneous WDM system, detailing the effect that sets the penalty in a wide range of system setups. We vary: pulse shaping (NRZ, aRZ and iRZ), transmission link (DM-link, with two different RDPS value and NDM-link) and transmission fiber (SMF and LEAF).

In Sec. 4.2.2 we analyze the dependence of the Q-factor on the number of neighboring channels/channel spacing for both SMF-based DM and NDM link. In Sec. 4.2.4 we investigate the impact of channel power on the central PDM-QPSK channel, in a 19-channel WDM transmission system.

4.2.1 Numerical Setup

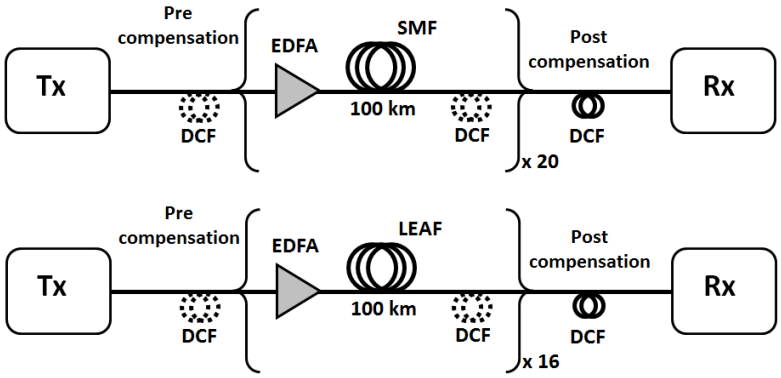


Figure 4.12: Simulation setup of 112 Gb/s PDM-QPSK homogeneous system: 20×100 km spans of SMF (a), 16×100 km spans of LEAF (b). The dotted line fibers are used only in DM links.

We simulated with the open-source software Optilux[3] the system setup reported in Fig. 4.12. All channels were 112 Gb/s (100 Gb/s plus 12% of FEC overhead) PDM-QPSK and they were synchronous in time. Three different pulse shapes were used: NRZ, aRZ and iRZ, respectively. All lasers had first their SOP independently randomized over the Poincaré sphere, and were then modulated by nested Mach-Zehnder modulators with independent random sequences of 1024 symbols each. During multiplexing, each channel was filtered by a 2$^{\text{nd}}$ order super-Gaussian optical filter of bandwidth 0.4 nm. We simulated two different transmission links:

- 20×100 km spans of SMF ($D = 17$ ps/nm/km @1550 nm, $A_{eff} = 80\ \mu m^2$, $n_2 = 2.5 \cdot 10^{-20}$ m^2/W, $\alpha = 0.2$ dB/km, $D^{'} = 0.057$ ps/nm^2/km)
- 16×100 km spans of LEAF ($D = 4$ ps/nm/km @1550 nm, $A_{eff} = 72\ \mu m^2$, $n_2 = 2.7 \cdot 10^{-20}$ m^2/W, $\alpha = 0.2$ dB/km, $D^{'} = 0.085$ ps/nm^2/km)

with two different maps:

- DM link with a residual dispersion per span (RDPS) of 30 ps/nm or 100 ps/nm
- NDM link (i.e., without in-line compensation).

In DM links, before transmission, we inserted a pre-compensating fiber whose dispersion was chosen according to the straight line rule [48], while no pre-compensating fiber was used in NDM link. In each link we set to zero the overall cumulated dispersion through a post-compensating fiber (this task is usually performed by an electrical GVD compensator in real coherent systems). All compensating fibers were supposed to be purely linear. Fiber propagation was obtained by solving the Manakov-PMD equation through the Split Step Fourier Method (SSFM, Sec. 1.2.3), neglecting FWM [22]. We assumed flat gain amplifiers with 6 dB noise figure at each span end, although the entire link noise was loaded as a unique noise source before detection. Such an approach neglects nonlinear phase noise, which is here negligible [25]. Before detection, we perfectly compensated optical linear impairments, i.e., GVD and birefringence by applying the inverse Jones matrix of the optical line, which allows us to focus entirely on the extra penalty coming from the interplay of linear and nonlinear distortions along the link [60].

The central channel was extracted by a 2^{nd} order super-Gaussian optical filter (bandwidth 0.25 nm (NRZ) and 0.22 nm (aRZ and iRZ) to have same Q-factor in the linear regime) and then detected with a DSP based coherent receiver including: mixing with an ideal local oscillator, low pass filtering over a bandwidth of 17 GHz, polarization recovery through a least-mean-square algorithm [62], sampling, phase-recovery with the Viterbi and Viterbi (V&V) algorithm using 7 taps, decision, and finally differential decoding [65]. We estimated the bit error rate (BER) of the central PDM-QPSK through the Monte Carlo algorithm by counting on average 100 errors, and then converting the estimated BER to Q-factor [25]. Each BER was averaged over 25 different runs with different random seeds, corresponding to selection of different WDM random data patterns, launched signals SOPs, and fiber birefringence. For a fair comparison, we used the same random realizations when testing the different pulse formats.

4.2.2 Number of WDM channels/channel spacing to correctly assess the cross-nonlinearities

In this section we investigate the Q-factor as a function of the WDM number of channels (spaced by $0.4\,nm$), for a SMF-based DM link, with RDPS of 30 ps/nm. We limit our analysis up to a 47-channel WDM comb because of time and memory restrictions. Fig. 4.13 shows the PDM-QPSK Q-factor versus WDM number of channels for NRZ-, aRZ and iRZ-PDM-QPSK homogeneous systems, respectively. The power of channels is set to -1 dBm for NRZ and aRZ and 2 dBm for iRZ, to have almost 5 dB of cross-nonlinear penalties at 40-channel WDM comb for all pulse formats.

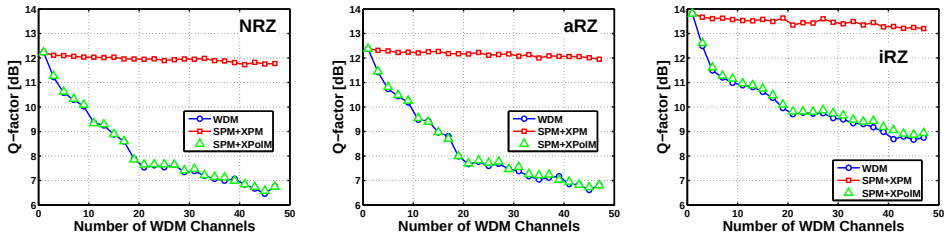


Figure 4.13: Q-factor versus number of WDM channels in a PDM-QPSK SMF-based DM link with RDPS of 30 ps/nm. Pulse formats: aRZ, iRZ and NRZ.

From Fig. 4.13 we note that in presence of XPolM the probe channel Q-factor decreases monotonically for increasing number of channels and we do not observe clear saturation up to 47 channels. A similar analysis was done by *Xia et al.* in [66] for an NDM link and several transmission fibers. They showed that the Q-factor penalty increases for increasing number of channels, although it becomes negligible for more than 40 channels in an SMF-based transmission link. Fig. 4.13 also shows that, since the blue and green curves are indistinguishable, the XPolM is the strongest cross-nonlinear effect in such a system, as confirmed by Sec. 4.2.4.

To investigate the contribution of PDM-QPSK-induced XPolM and XPM penalties at each channel spacing, we emulated a WDM comb composed of three PDM-QPSK channels whose equal spacing $\Delta\lambda$ was varied from 0.4 nm to 20 nm. We increased the channel spacing by moving the two interfering channels, with randomly oriented SOPs, away from the central one and then we measured the central channel Q-factor. We adapted the nonlinear SSFM step to correctly take into account the walk-off accordingly to the rule discussed in App. D.1. We emulated three SMF-based transmission links: a DM link with RDPS of 30 ps/nm, which we call DM30 in the following, a DM link with RDPS of 100 ps/nm (DM100) and a NDM link. The PDM-QPSK central probe channel power was set to -2 dBm, while the interfering PDM-QPSK powers were set to have a comparable central channel Q-factor for SPM+XPM and SPM+XPolM at 0.4 nm. For NDM link we used 4096-long sequences, according to App. D.2.

In Fig. 4.14 we report the Q-factor versus channel spacing obtained by propagating the WDM comb along the DM30, DM100 link and NDM link, respectively. The NRZ, aRZ and iRZ shapes are investigated in SPM+ XPM case (red-square) and SPM+XPolM case (green-triangle).

Fig. 4.14 confirms that even if XPM and XPolM are comparable at 0.4 nm their scal-

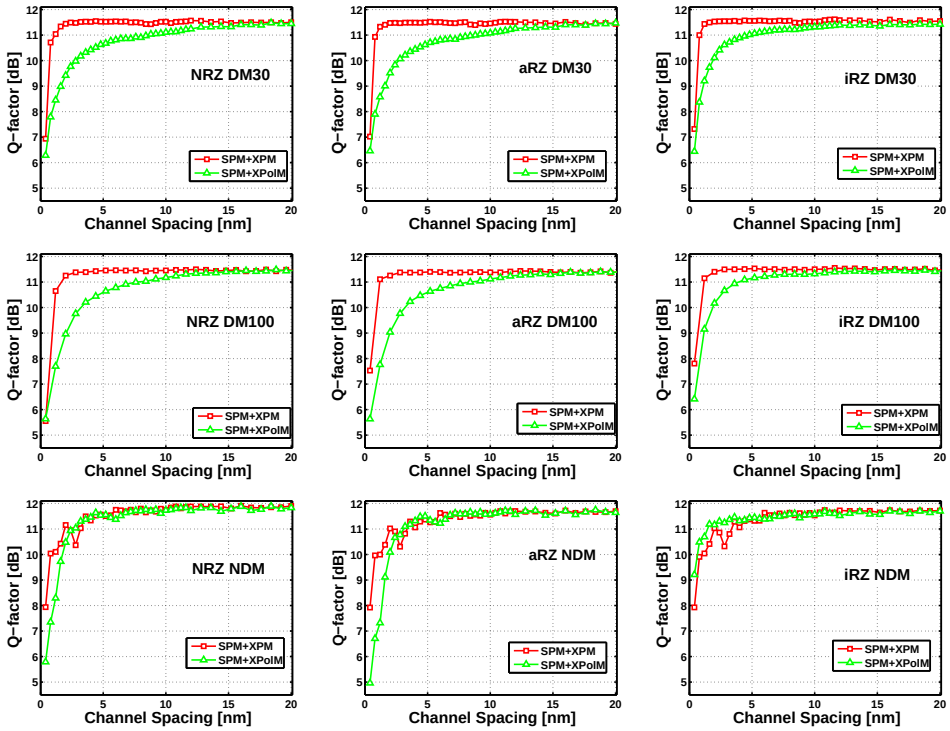


Figure 4.14: Q-factor versus channel spacing in 3 channel PDM-QPSK system relying on DM30 (top-line), DM100 (middle-line) and NDM link (bottom-line). SPM+XPM and SPM+XPolM case are reported for NRZ- (left-column) aRZ- (center-column) and iRZ-PDM-QPSK (right-column).

ing laws remain different, showing that the XPolM-induced penalty becomes negligible at a larger channel spacing respect to the XPM-induced penalty, as shown in Sec. 4.1.3 for hybrid systems. This figure also shows that as for hybrid systems, for both DM links the PDM-QPSK probe channel suffers the XPM-induced penalty by interfering channel spaced less than 1.2 nm, for all pulse formats. The saturation point for XPolM is different for the three pulse formats, being around 12 nm for aRZ and NRZ, while around 8 nm for iRZ. This difference is reasonably due to the weak XPolM induced by iRZ pulses [55, 57].

For the NDM link, Fig. 4.14 indicates that the Q-factor experiences resonances at increasing channel spacing thus making difficult to identify a saturation point, both for SPM+XPM and SPM+XPolM case. We associate the resonances to the special shape of the XPM filter in

NDM map, shown in Sec. 3.3.3, whose strongly irregular profile can induce strong variations of the performance.

4.2.3 XPM- and XPolM-induced penalty versus of walk-off

In this section we generalize the scaling rules of the XPM- and XPolM-induced penalty versus walk-off by expressing the walk-off in terms of symbols per span that walk past a probe symbol as

$$Walk-Off = \frac{D \cdot \Delta\lambda \cdot L_{span}}{T_0} \tag{4.1}$$

where $D\,[ps/nm/km]$ is the transmission fiber dispersion, $\Delta\lambda\,[nm]$ is the channel spacing, $L_{span}\,[km]$ is the span length and $T_0\,[ps]$ is the symbol period. The walk-off defined as in eq. (4.1) allows us to compare different transmission links by a unique parameter.

In this investigation, we sent three 112 Gb/s NRZ-PDM-QPSK channels in a resonant DM link, i.e., a link without pre-compensation, with zero RDPS and zero total cumulated dispersion. Since RDPS is zero in such a link, we take into account only the walk-off induced by the fiber dispersion. We measured the Q-factor of the central channel by varying the channel spacing. The central PDM-QPSK had a power of -2 dBm, while the two interfering channels had 4 dBm in the SPM+XPolM case and 12 dBm in the SPM+XPM case, in order to have comparable Q-factors. All states of polarization were random while the receiver was the same described in Sec. 4.2.1.

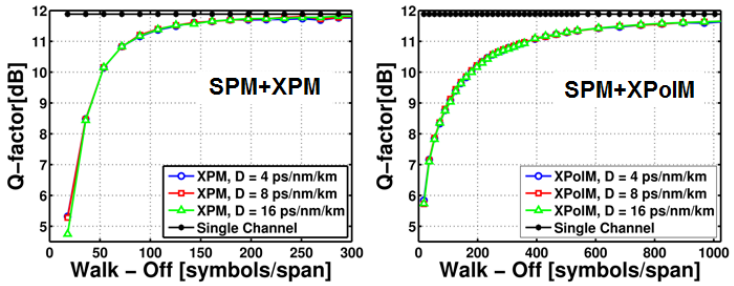


Figure 4.15: DM link with zero RDPS: Q-factor versus walk-off in presence of SPM+XPM (left) and SPM+XPolM (right), obtained for three fiber dispersion values (4, 8 and 16 [ps/nm/km]).

Fig. 4.15 depicts the Q-factor of the central PDM-QPSK versus the walk-off in the SPM+XPM

(left) and the SPM+XPolM (right) cases. In each figure three curves, obtained with different dispersion fiber values (4, 8 and 16 [ps/nm/km]) are reported. The perfect superposition of the three curves confirms that the performance really scale with the walk-off expressed in eq. (4.1).

Fig. 4.15 confirms that both cross-nonlinearities decrease for increasing walk-off, but with different "slopes". Indeed a walk-off larger than 150 symbols is sufficient to affirm that XPM-induced penalties are negligible, while a walk-off of almost 1000 symbols has to be taken into account to correctly evaluate the XPolM. We obtained similar scaling rules for aligned Return to Zero (aRZ) and interleaved Return to Zero (iRZ) PDM-QPSK.

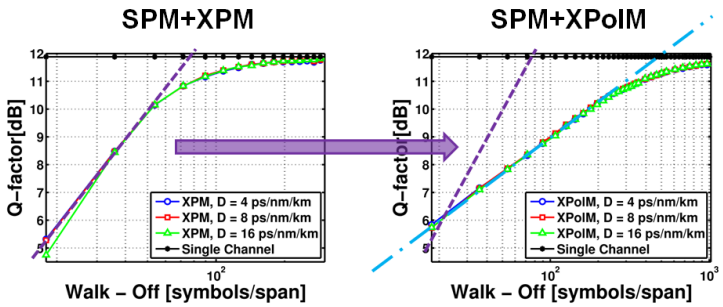


Figure 4.16: DM link with zero RDPS: Q-factor versus walk-off in log-scale in presence of SPM+XPM (left) and SPM+XPolM (right), obtained for three fiber dispersions (4, 8 and 16 [ps/nm/km]).

The difference between the slope of the curves reported in Fig. 4.15 are best viewed in a log-log scale, depicted in Fig. 4.16. In such a scale the Q-factor shows a Bode-plot behaviour, where the first ascending part is almost linear with the walk-off. For XPM this observation can be related to the linear dependence of the 3-dB bandwidth of the XPM filter with the walk-off[52]. For XPolM we do not have a simple explanation as for XPM, but these results indicate that the scaling law with the walk-off still shows a Bode like plot. Furthermore, reporting the XPM slope (purple) in XPolM graphs (please note the different scale of X axis) we note that the XPM slope (purple) is almost twice of XPolM one (cyan). Probably the reason is related to the XPolM dependence on the SOPs, but on this point further analysis is needed.

4.2.4 XPM and XPolM impairments: Q-factor vs Channel Power

In this section we analyze the impact of the channel power in WDM transmission. To this aim we simulated the propagation of a homogeneous WDM comb composed of 19 channels with identical modulation and we measured the central channel Q-factor in function of the channel power, which is the same for all channels. The Q-factor versus channel power is depicted in Fig. 4.17 for the 20 x 100 km-long SMF-based DM link with RDPS = 30 ps/nm (DM30), 100 ps/nm (DM100) and NDM link, respectively.

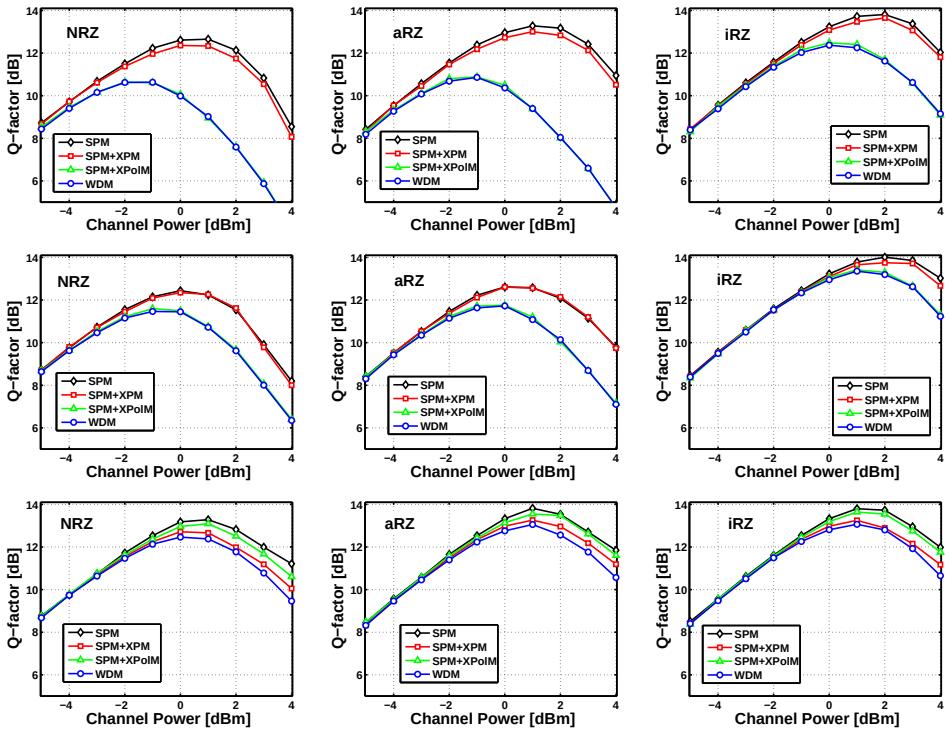


Figure 4.17: Q-factor versus channel power for 20x100 km SMF lynk of type: DM30 (top-line), DM100 (middle-line) and NDM link (bottom-line). NRZ- (left-column) aRZ- (center-column) and iRZ-PDM-QPSK (right-column).

Both DM maps show that almost the entire penalty is due to XPolM, while XPM adds a negligible contribution to the overall penalty; while in the NDM link the difference between XPolM and XPM is almost null. It turns out that in DM links the signal amplitudes are peri-

odical and the XPM-induced penalty is cancelled out by differential decoding, while in NDM links the amplitudes become aperiodic (amplitude distortion) and XPM-induced penalty becomes comparable to XPolM-induced one.

Furthermore, comparing the upper and the middle line of the figure, we note an increase of WDM-performance for increasing RDPS, for all the pulse formats. More details on the RDPS will be provided in Sec. 5.5.

Fig. 4.17 also shows that in DM link the iRZ-PDM-QPSK achieves much better performance than aRZ and NRZ. There are two main reasons: the first one is related to the iRZ-PDM-QPSK peak power, which is smaller in iRZ than aRZ (half) and NRZ generally reduces the nonlinear effects; the second is connected to the mapping between the information symbols and the SOP, which is pattern dependent in aRZ and NRZ but pattern independent in iRZ, yielding a deterministic and periodic SOP.

The power is relevant in systems with small walk-off, i.e., one-to-one symbol interaction, since the nonlinearities are proportional to peak power. The polarization is relevant in systems with large walk-off, i.e., one-to-many symbols interaction because of in this case all the pulse formats have the same average power but, while NRZ and aRZ induce nonlinear rotation around a fixed axis in a symbol time, the iRZ induces two half-rotations around opposite axes that compensate each other, as visible in Fig. 4.18.

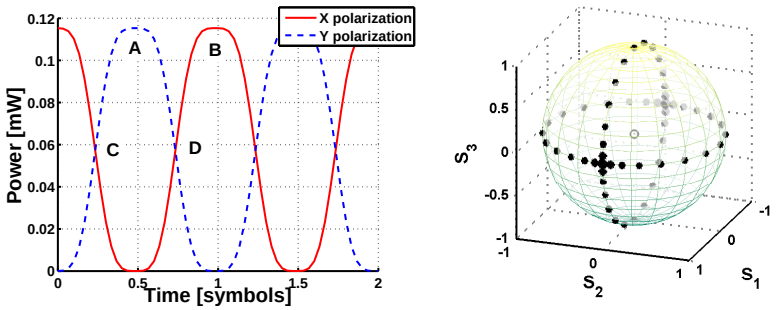


Figure 4.18: iRZ-PDM-QPSK: time waveform of X and Y polarization (left) and SOP on the Poincaré sphere (right).

Fig. 4.18 depicts the iRZ-PDM-QPSK SOP (right) and the time waveform of X and Y subcarriers of iRZ-PDM-QPSK (left). Since in iRZ the X and Y polarization signals are time shifted by half symbol, there is no overlap between the pulses in X and Y polarizations at the center of the symbols; hence, in A and B of Fig. 4.18(left) the SOP of the signal lies in $-\widehat{S_1}$ and $\widehat{S_1}$ on the Poincaré sphere, respectively. In C and D, since we have the same power on X

and Y subcarriers, the two circles in Fig. **4.18**(left) pass through the $\pm\widehat{S_2}$ and $\pm\widehat{S_3}$ in the same cases of App. (B), but with half of power.

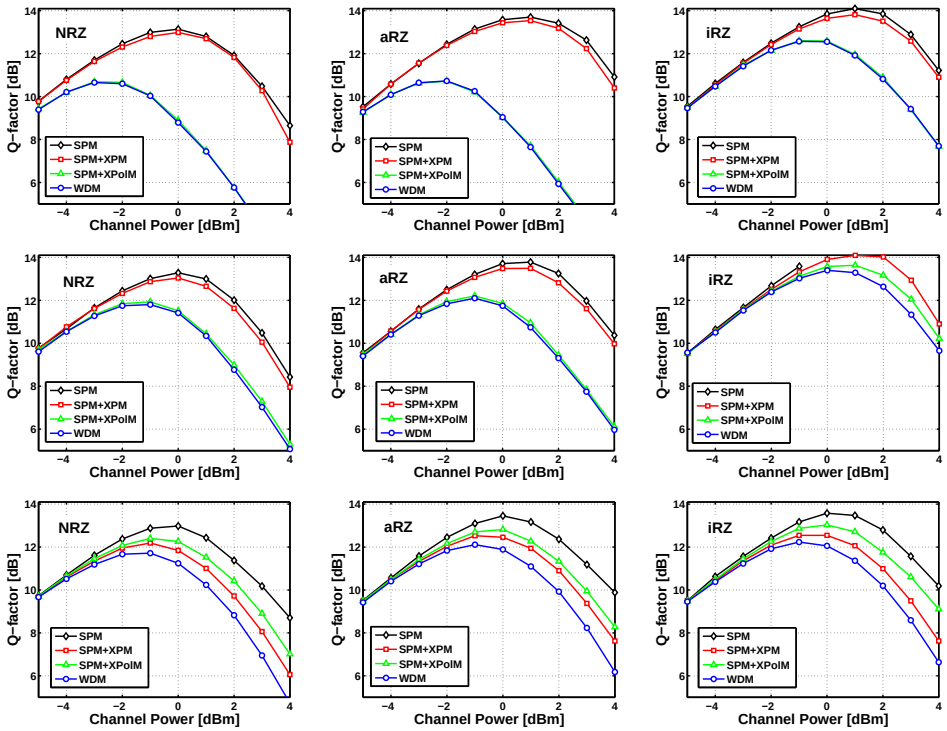


Figure 4.19: Q-factor versus channel power for LEAF-based 16x100 km DM link with RDPS = 30 ps/nm (upper-line), RDPS = 100 ps/nm (center-line) and NDM link (bottom-line). NRZ- (left-column) aRZ- (center-column) and iRZ-PDM-QPSK (right-column).

We repeated the same analysis using a transmission link composed of 16x100km-long spans of LEAF fiber. The Q-factor vs. channel power is depicted in Fig. 4.19 for the DM30, DM100 and NDM link respectively. Please note that we limited our measurements up to a Q-factor of 14.2 dB to cap on simulation time. For the DM link the trends are quite similar to those obtained for the SMF transmission link, hence the XPolM again sets the PDM-QPSK performance. It is worth to note that the absolute Q-factor values are similar because the SMF link is longer than the LEAF link and this compensates for the stronger cross-channel nonlinearities associated with the LEAF fiber, because of the smaller fiber dispersion. For the NDM link the difference between XPolM and XPM effects is smaller than in the DM links

and the XPM is slightly stronger than XPolM. As in Fig. 4.17, comparing upper and middle figures, we note an increase of WDM-performance for increasing RDPS, for all the pulse formats.

4.3 PDM-BPSK homogeneous systems

In this section, we numerically analyze the nonlinearities impact on the performance of a 42.8 Gb/s PDM-BPSK probe channel (10Gbaud plus 7% of FEC overhead) in a homogeneous WDM transmission system. As for PDM-QPSK, we investigate the dependence of the central PDM-BPSK Q-factor both on the number of neighboring channels and on the channel power in a 19-channel WDM transmission system. Since in PDM-QPSK systems we observed similar performances for aRZ and NRZ, in this section we do not consider the aRZ pulse shape. We focus on the NRZ and iRZ (66%, 50% and 33% carved) pulse shapes, showing both the mean Q-factor and the Q-factor fluctuation through its Probability Density Function (PDF) and Cumulative Distribution Function (CDF).

4.3.1 Numerical Setup

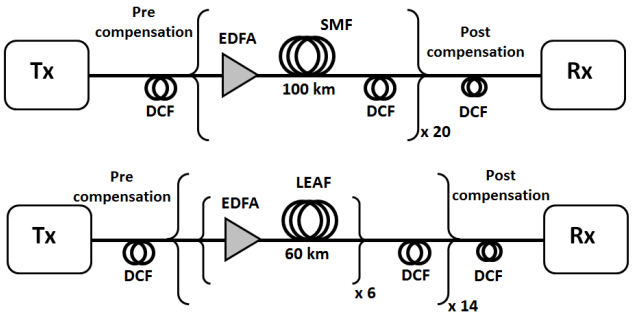


Figure 4.20: Simulation setup of 42.8 Gb/s PDM-BPSK homogeneous system: sub-marine link (top) and terrestrial link (bottom).

The system setup reported in Fig. 4.20, in which all channel were 42.8 Gb/s PDM-BPSK and were synchronous in time, was emulated with Optilux[3]. All lasers had first their SOP independently randomized over the Poincaré sphere, and were then modulated by nested Mach-Zehnder modulators with independent random sequences of 1024 symbols each. Dur-

ing multiplexing, each channel was filtered by a 3^{rd} order super-Gaussian optical filter of bandwidth 0.4 nm. We simulated two different transmission links, based on different fibers:

- *Sub-marine link*: DM-link formed by 14 trunks of 6×60 km spans of LEAF ($D = 4$ ps/nm/km @1550 nm, $A_{eff} = 72\ \mu m^2$, $n_2 = 2.7 \cdot 10^{-20}$ m^2/W, $\alpha = 0.2$ dB/km, $D^{'} = 0.085$ ps/nm^2/km) with RDPS=0 ps/nm

- *Terrestrial link*: DM-link formed by 20×100 km spans of SMF ($D = 17$ ps/nm/km @1550 nm, $A_{eff} = 80\ \mu m^2$, $n_2 = 2.5 \cdot 10^{-20}$ m^2/W, $\alpha = 0.2$ dB/km, $D^{'} = 0.057$ ps/nm^2/km) with RDPS=30 ps/nm

Before transmission, we inserted a pre-compensating fiber whose dispersion was chosen according to the straight line rule [48] in the terrestrial link, while no pre-compensating fiber was used in the sub-marine link. In each link we set to zero the overall cumulated dispersion through a post-compensating fiber. All compensating fibers were supposed to be purely linear in the terrestrial link, while also nonlinear effects were emulated in the sub-marine link. Fiber propagation was obtained by solving the Manakov-PMD equation through the Split Step Fourier Method (SSFM, Sec. 1.2.3), neglecting FWM [22]. We assumed flat gain amplifiers with different noise figure for the different pulse shapes, so as to have the same Q-factor in the linear regime. The entire link noise was loaded as a unique noise source before detection, thus neglecting nonlinear phase noise. Before detection, we perfectly compensated the optical linear impairments, i.e., GVD and birefringence (PMD was set to zero) by applying the inverse Jones matrix of the optical line.

The central channel was extracted by a 3^{rd} order super-Gaussian optical filter of bandwidth 0.4 nm and then detected with a DSP based coherent receiver including: mixing with an ideal local oscillator, low pass filtering over a bandwidth of 15 GHz, polarization recovery through a CMA algorithm proposed for BPSK signal[67] working with two samples per symbol, sampling, phase-recovery with the V&V algorithm with several number of taps, decision, and finally differential decoding [65]. We performed 100 learning loops, in absence of ASE, to bring the CMA to convergence. We estimated the bit error rate (BER) of the central PDM-BPSK through the Monte Carlo algorithm by counting on average 100 errors, and then converting the estimated BER to Q-factor. When not otherwise specified, each BER was averaged over 25 different runs with different random seeds, corresponding to the selection of different WDM random data patterns, launched signals SOPs, and fiber birefringence. For a fair comparison, we used the same random realizations when testing different pulse formats.

4.3.2 Simulation Results

In Sec. 4.2 we properly chose the optical receiver filter bandwidth to have same Q-factor in the linear regime for all the pulse shapes. In this section we use the same optical receiver filter and we adapt the noise figure to have same Q-factor in the linear regime for all the pulse shapes. Note that the two approaches lead to the same results. We start analyzing the different pulse formats back-to-back sensitivity. Fig. 4.21 depicts the PDM-BPSK Q-factor versus the bandwidth of the electrical filter. This figure shows a difference of almost 1 dB between the pulse shape with narrower spectrum (NRZ) and wider spectrum (iRZ 33%), for an electrical filter with a bandwidth of 15 GHz. The iRZ penalty w.r.t. NRZ at 22 GHz is due to the optical transmission filter (50 GHz) and is proportional to the spectrum bandwidth.

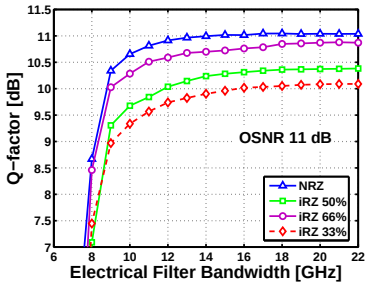


Figure 4.21: Back-to-Back PDM-BPSK Q-factor versus electrical filter bandwidth for: NRZ, iRZ 50%, iRZ 66% and iRZ 33% pulse shapes.

In order to evaluate the minimum number of WDM channels to correctly take into account the XPM- and XPolM-induced penalty, we evaluate the central channel Q-factor by increasing the number of interfering channels. To this aim we emulate the LEAF-based submarine link, in which the very long link and the small fiber dispersion increase the cross-nonlinearities.

In Fig. 4.22 the central channel Q-factor versus number of WDM channels for NRZ and iRZ 50% is reported. All channels were randomly oriented on the Poincarè sphere and their powers were set to -4 dBm. Several V&V number of taps were used: 3, 7, 21 and 41, respectively. All the nonlinear effects were emulated.

For both pulse shapes we observe a decreasing Q-factor for increasing number of WDM channels up to 33 channels. From Fig. 4.22 we also observe that the higher Q-factors are obtained using a large V&V number of taps (21 and 41) for NRZ pulse shape, suggesting us that the XPolM is the dominant cross-nonlinearity (Sec. 4.1.2). A smaller impact of V&V

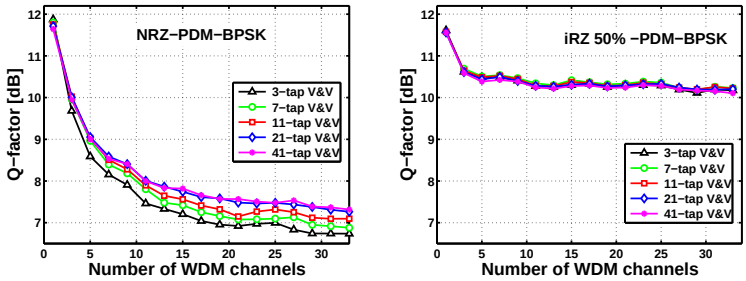


Figure 4.22: LEAF-based sub-marine link: PDM-BPSK Q-factor versus number of WDM channels. Pulse formats: NRZ (left), iRZ 50% (right).

number of taps is observed for the iRZ-PDM-BPSK system, probably due to the weaker nonlinear-induced penalties, as observed in Sec. 4.2 for PDM-QPSK systems.

To determine which is the dominant nonlinear effects we fix the number of channels to 19, considering such a number sufficient to correctly take into account the cross-nonlinear penalty. Fig. 4.23 shows the central PDM-BPSK Q-factor versus the channel power for NRZ, iRZ 66%, iRZ 50% and iRZ 33% using a 21-tap V&V algorithm. The errobars indicate the polarization-induced standard deviations.

Fig. 4.23 shows that XPolM is the dominant effect for NRZ and iRZ 66% shapes, while for iRZ 50% and iRZ 33% XPolM and XPM are comparable. Furthermore we observe that, as for PDM-QPSK (Sec. 4.2.4), the subcarrier time interleaving allows iRZ to perform better than NRZ. Indeed the iRZ format increases of almost 1.5 dB the Non-Linear Threshold (NLT) and induces weaker Q-factor standard deviations respect to NRZ.

To investigate in depth the Q-factor fluctuation induced by interfering SOPs and bit patterns, we visualize the Q-factor PDFs and CDFs. Due to the extremely time-consuming simulations for such analysis, we analyzed only the NRZ and iRZ 50% formats in the terrestrial link.

In Fig. 4.24 the Q-factor PDF and CDF for NRZ and iRZ 50% are reported. Each curve was obtained by collecting 1000 Q-factor values coming from simulations with different random seeds, corresponding to different WDM random data patterns and launched signals SOPs. We set the power 1 dB beyond the NLT to have the same nonlinear penalty and we varied the noise figure to have almost the same mean Q-factor. This figure confirms that the iRZ 50% slightly reduces the Q-factor fluctuations. Indeed, the figure on the right shows a slope of almost 1.2 decade/dB for iRZ 50% and 0.9 decade/dB for NRZ shape.

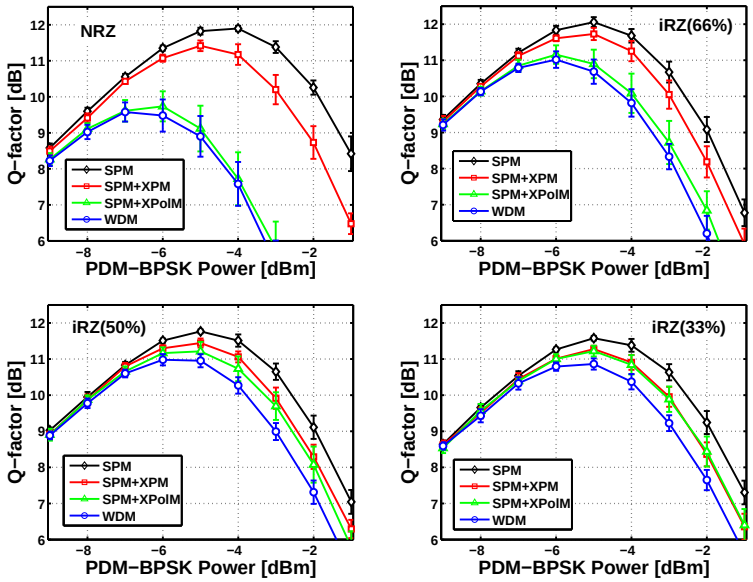


Figure 4.23: LEAF-based sub-marine link: PDM-BPSK Q-factor versus channel power. Pulse formats: NRZ, iRZ 66%, iRZ 50% and iRZ 33%. SPM case (black-diamond), SPM+XPM case (red-square), SPM+XPolM case (green-triangle) and WDM case (blue-circle).

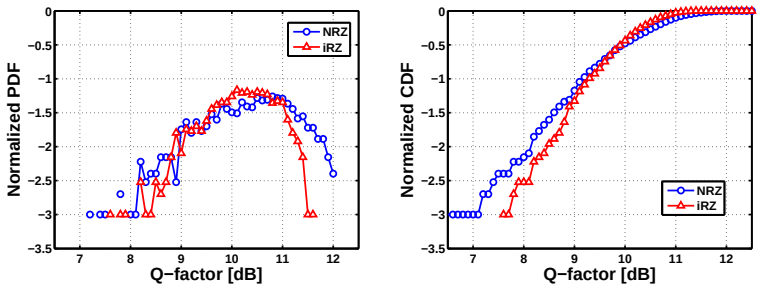


Figure 4.24: SMF-based terrestrial link: CDF and PDF for NRZ- (blue-circle) and iRZ 50% (red-triangle) shapes.

Chapter 5

Techniques to Mitigate Cross-channel Nonlinearities

In Chapter 4 we showed that both cross-phase modulation (XPM) and cross-polarization modulation (XPolM) are reduced by increasing channel walk-off. Hence cross-channel nonlinearities can be mitigated by the fiber group velocity dispersion (GVD), which, especially in non-dispersion managed (NDM) links, induces substantial channel walk-off [55, 65]. Another efficient way to increase channel walk-off is to use passive devices that introduce different delays on adjacent channels at specific points of the line. Such devices efficiently suppress XPM in on-off keying (OOK) systems [68] and PDM-QPSK systems [69, 70], and should prove to be effective also against XPolM.

Furthermore in Sec. 4.2 we showed the benefits of the interleaved return-to-zero (iRZ) pulse format in mitigating cross-channel effects in PDM-QPSK homogeneous systems transmitted on DM links. Such benefits are confirmed by experiments [57] and simulations [55]. However, polarization mode dispersion (PMD) should reduce the iRZ advantages by partially temporally re-aligning the polarization tributaries, thus destroying the basic idea of interleaving. On the other hand PMD makes the states of polarization (SOPs) of different channels to follow different paths over the Poincaré sphere, thus reducing their cross-interaction, provided that linear PMD is fully compensated at the receiver [60]. Hence a quantitative analysis of the PMD impact on iRZ transmission is of great interest.

In this chapter we compare by simulation different optical methods to improve the resilience of coherent 112-Gb/s PDM-QPSK WDM transmissions against cross-channel nonlinearities. Such methods consist of 1) increasing the line group velocity dispersion (GVD),

or 2) the line polarization- mode dispersion (PMD), or 3) inserting in-line cross-phase modulation (XPM) suppressors. Such methods are tested using non-return-to-zero (NRZ), aligned return-to-zero (aRZ), and iRZ pulse formats. We show that the nonlinearity- mitigating effect underlying all three methods is an increase of the interchannel decorrelation, obtained by either increasing walk-off (methods 1 and 3) or by depolarizing the WDM channels (method 2). This chapter extends the work of [71], where the effectiveness of the previous solutions was analyzed for the first time in a coherent scenario employing PDM-QPSK channels.

5.1 Optimal suppressor delay in mitigating XPM and XPolM

In [68] the authors showed the existence of an optimal XPM-suppressor delay to minimize the XPM variance in single polarization OOK systems. In this section we search such optimal XPM-suppressor delay D in mitigating cross-channel effects (both XPM and XPolM), in a pump-probe transmission system. To this investigation we emulated the resonant DM link (RDPS=0) composed of 20x100km-long spans of NZDSF fiber, reported in Fig. 5.1, whose cross-nonlinearities are dominant.

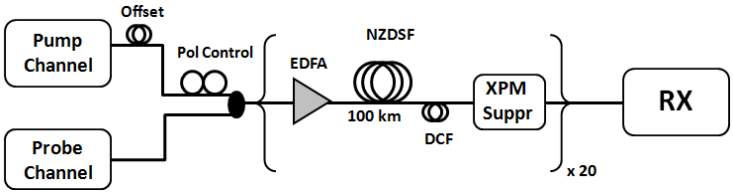


Figure 5.1: Pump-Probe numerical setup: XPM suppressor [68] is applied at the end of each span.

The pump and probe sequences are not synchronous in time, since before transmission we applied a random offset to pump channel sequence. We used 9 different offsets, from zero to one symbol time. The pump channel is randomly oriented on the Poincarè sphere. For each suppressor delay we measured the Q-factor of the probe channel, extrapolated from the BER, averaged 10 random realizations of pump SOP, transmission pattern and time offsets. BER was estimated through Monte Carlo algorithm stopped after counting 400 errors.

We started by reproducing a 10 Gb/s OOK probe – 10 Gb/s OOK pump transmission system, exactly as in [68]. The probe power was 2 dBm, while the pump channel had 8 dBm

of power. Fig. 5.2 depicts the symbol variance obtained in [68] (left-side) and our simulated Q-factor (right-side). Both approaches confirm the existence of an optimal delay value D of almost 100 ps, i.e one symbol time.

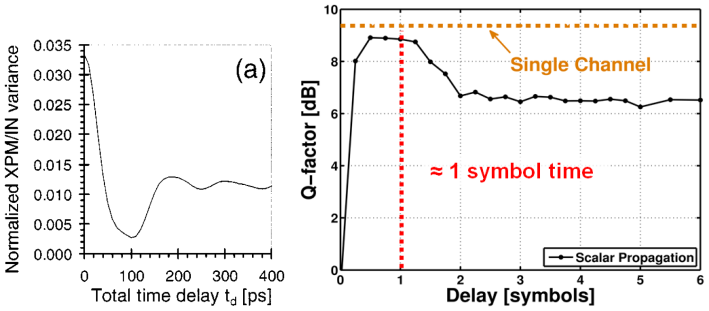


Figure 5.2: Optimal delay in a two channel OOK system, from [68] (left-side) and after simulation.

We next substituted the OOK probe with a 112 Gbit/s PDM-QPSK signal with power -3 dBm to check the effectiveness of the XPM suppressor in presence of a time-varying polarization, obtaining the Q-factor shown in Fig. 5.3. The neighboring channels were 112 Gbit/s PDM-QPSK with power 2 dBm (Fig. 5.3 (left)) or 10 Gbit/s OOK with power 1 dBm (Fig. 5.3 (right)). In the simulations we made use of our technique for separating the nonlinear effects to verify the existence of an optimal delay for each nonlinear cross effect, namely XPM and XPolM.

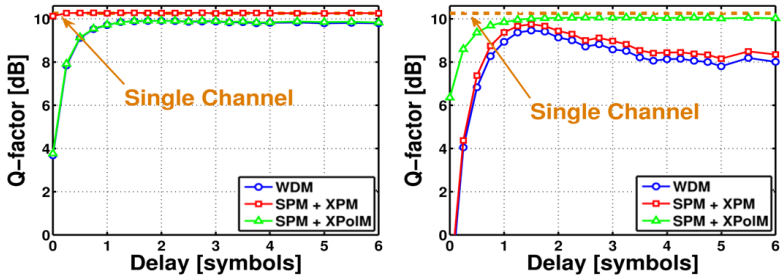


Figure 5.3: 112 Gb/s PDM-QPSK Q-factor in function of XPM suppressor delay, propagating with another 112 Gb/s PDM-QPSK (left) or 10 Gb/s OOK (right).

Fig. 5.3(right) shows that when XPolM is neglected (square) the optimal XPM suppressor

delay is almost one symbol time, i.e., 35.7 ps. On the contrary, when the XPM is turned off (triangle), an optimal delay does not exist and the best solution is to maximize the decorrelation among the channels. Fig. 5.3 (left), which depicts the PDM-QPSK Q-factor when the pump channel is PDM-QPSK, confirms that the optimal delay does not exist when XPolM is the dominant effect.

This different behavior of XPM and XPolM is an indicator that while XPM grow in a resonant way along propagation, XPolM does not. Therefore, any attempt to make XPolM destructively interfere with itself fails. The best way to counteract XPolM is thus to induce a complete decorrelation span-by-span.

5.2 Numerical Setup

To compare the different optical methods to reduce the cross-channel nonlinearities, we simulated with the open-source software Optilux[3] the transmission of a 19-channel 112Gb/s PDM-QPSK homogeneous WDM system with 50 GHz channel spacing. All lasers had first their SOP independently randomized over the Poincaré sphere, and were then modulated by nested Mach-Zehnder modulators with independent random sequences of 1024 symbols each. During multiplexing, each channel was filtered by a 2^{nd} order super-Gaussian optical filter of bandwidth 0.4 nm. The simulated link was composed of 20×100 km spans of single mode fiber (SMF), with zero overall cumulated dispersion obtained with an ideal linear post-compensating fiber. Two different setups were considered: 1) a DM link with pre-compensation whose dispersion was chosen according to the straight line rule [48] and a residual dispersion per span (RDPS) of 30 [ps/nm], and 2) a NDM link without pre- and in-line compensation. PMD was emulated only in the DM link, since in NDM links the interaction between PMD and Kerr nonlinearity is known to be negligible [60].

The XPM suppressor [68], optionally used only in DM links, was implemented by a demultiplexer followed by a bank of delay lines and a multiplexer, as sketched in Fig. 5.4. Each channel in the suppressor was delayed by D [ps] with respect to its smaller-wavelength neighbor.

Fiber propagation was obtained by solving the Manakov-PMD equation through the split step Fourier algorithm. Fiber birefringence and PMD were emulated by using 50 random waveplates per span. We assumed flat gain amplifiers with 6 dB noise figure at each span end, although the entire link noise was loaded as a unique noise source before detection. Before detection, we perfectly compensated optical linear impairments (GVD and PMD) by applying the inverse Jones matrix of the optical line, which allows us to focus entirely on

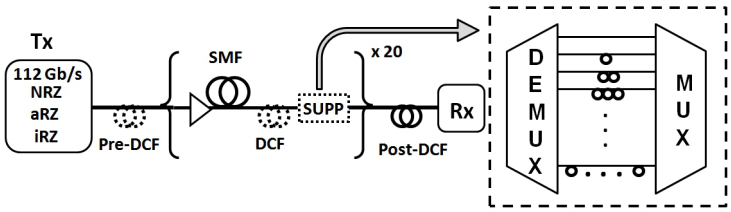


Figure 5.4: Numerical setup. Dotted box: XPM suppressor proposed in [68].

the extra penalty coming from the interplay of linear and nonlinear distortions along the link [60].

The central channel was extracted by a 2nd order super-Gaussian optical filter (bandwidth 0.25 nm (NRZ) and 0.22 nm (aRZ and iRZ) to have same Q-factor in the linear regime) and then detected with a DSP based coherent receiver including: mixing with an ideal local oscillator, low pass filtering over a bandwidth of 17 GHz, polarization recovery through the DARMAVG least-mean-square algorithm [62], sampling, phase-recovery with the Viterbi algorithm using 7 taps, decision, and finally differential decoding [65]. We estimated the bit error rate (BER) through the Monte Carlo algorithm by counting on average 100 errors, and then converting the estimated BER to Q-factor [25]. To take into account the stochastic nature of PMD, each BER was averaged over 40 different runs with different random seeds. Each seed corresponded to selection of different WDM random data patterns, SOPs, and fiber waveplates realizations. For a fair comparison, all numerical setups used the same random seeds.

5.3 Impact of variable delay and DGD

In this section we investigate the impact of the XPM suppressor on the performance of the NRZ, aRZ and iRZ-based systems in a 19-channel PDM-QPSK DM link without PMD. Unless otherwise noted, a suppressor was inserted at every span.

Fig. 5.5(a) shows the Q-factor versus suppressor delay D. The error bars indicate the Q-factor standard deviation. In the NRZ and aRZ cases we set the power to 0 dBm while for iRZ we used 1 dBm. With this choice, all formats work 1 dB beyond the power of maximum Q-factor at $D = 0\,ps$ and $DGD = 0\,ps$ (cfr. Fig. 5.6). We note that the XPM suppressor is effective for all formats, with an increasing Q-factor for increasing values of D, thus confirming that the best option is to maximize the decorrelation among channels. Q-factor is seen to saturate

after a delay of roughly 4 symbols (142.8 ps).

We next studied the impact of PMD on the same three pulse formats in the DM link without suppressor. In Fig. 5.5(b) we show the Q-factor versus average Differential Group Delay (DGD) obtained at the same powers as in Fig. 5.5(a). This figure shows that the DGD improves the Q-factor for all pulse formats, and that Q-factor saturates for an average DGD larger than 20 ps, in agreement with [60].

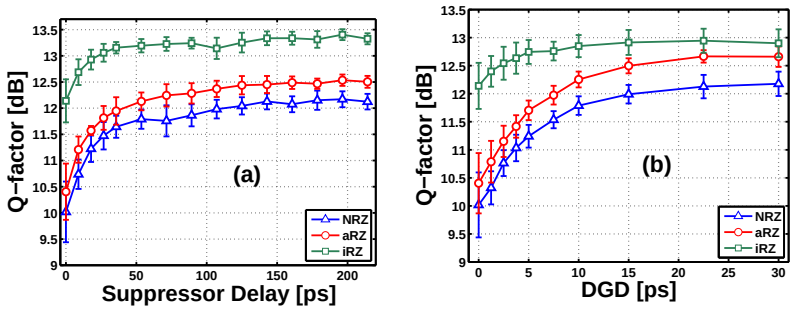


Figure 5.5: Q-factor versus suppressor delay D at zero DGD (a). Q-factor versus average DGD in absence of XPM suppressor (b). 19-channels PDM-QPSK system DM 20 x 100 km SMF link with pulse formats NRZ, aRZ and iRZ (symbol time = 35.7 ps).

The stochastic fluctuations of the Q-factor are mostly due to XPolM and are related to the random, symbol-dependent SOP orientation of the PDM-QPSK signals. In fact, the standard deviation is larger at small DGD, where XPolM is expected to be larger [60]. It is worth noting that iRZ has a smaller standard deviation than aRZ and NRZ at DGD=0, since iRZ is more tolerant to XPolM in absence of PMD (cfr. Sec. 4.2). Note that the iRZ Q-factor increases for increasing DGD, even if PMD degrades the iRZ pulse-interleaving, because PMD-induced depolarization is more effective in reducing XPolM.

5.4 XPM suppressor in absence and in presence of PMD

For the same WDM SMF-based DM link, we report in Fig. 5.6 the Q-factor versus power in absence/presence of either PMD (average DGD = 0 or 22.5 ps) or XPM-suppressor placed at the end of each span (delay D equal to 0 or 10 symbols, both at DGD = 0 ps). As a reference, in the same graphs we also report the single channel DM-case and the WDM NDM case (both in absence of DGD).

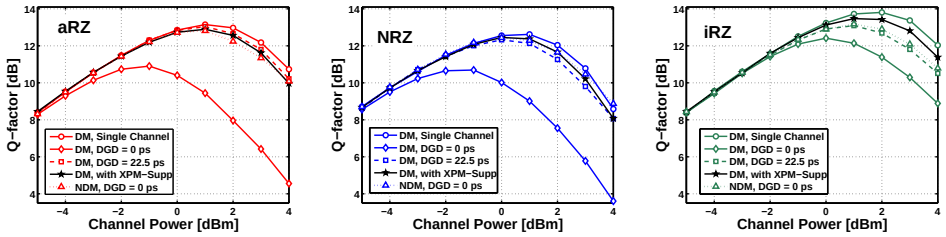


Figure 5.6: Q-factor versus channel power for a 19-channel PDM-QPSK 20 x 100 SMF link for different pulse formats and average DGD values. Black curves (stars) for XPM-suppressor calculated with D = 10 symbols, DGD = 0 ps.

The figure confirms that in absence of DGD the WDM NDM link largely outperforms the DM one. However, PMD improves the DM performance yielding Q-factors very close to the NDM case, also for iRZ, confirming that the PMD-induced depolarization compensates for the degraded time-interleaving. From the figures, we also note that for aRZ and NRZ the DM link with XPM suppressor has similar performance as the NDM link, while for iRZ the DM link with XPM suppressor is slightly superior to the NDM link. Reason is that the XPM suppressor reduces cross-channel interactions, but does not degrade pulse time-interleaving. It is thus the best option for a PDM-QPSK link with iRZ pulses.

In a final test we investigated more in detail the performance of the DM link with the XPM suppressor. Fig. 5.7 shows Q-factor versus power for all pulse formats either in absence of PMD, or with an average DGD = 22.5 ps. In any case a suppressor with D = 10 symbols was present at each span. From the figure, we note that PMD improves performance except for iRZ, where we observe a small decrease of the Q-factor in the nonlinear regime (descending region) making iRZ performance similar to aRZ. We ascribe such a worsening to the PMD-induced deterioration of the pulses' time-interleaving.

Fig. 5.8 shows the Q-factor versus power at DGD = 0 ps, where suppressors having D = 10 symbols were inserted every ten spans (circles), every five spans (squares), every two spans (triangles) and at all spans (stars) for the three pulse formats. By comparison with Fig. 5.6 we note that a DM link with suppressors is roughly as effective as an NDM link only when suppressors are inserted at every span, with a quick performance deterioration as the number of suppressors is reduced.

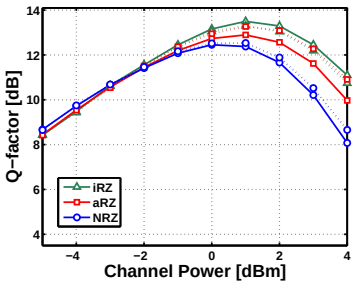


Figure 5.7: Q-factor versus power for a 19-channel PDM-QPSK DM link with XPM suppressor at all spans (ten-symbol delay) with average DGD = 0 ps (solid line) and 22.5 ps (dotted line).

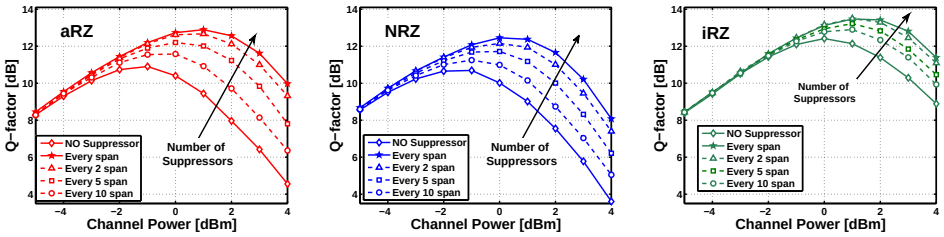


Figure 5.8: Q-factor versus power for a 19-channel PDM-QPSK DM link with zero DGD and XPM suppressor at all spans (stars), every ten spans (circles), every four spans (squares) and every other span (triangles), for aRZ, NRZ and iRZ pulse format.

5.5 Impact of DGD and RDPS

Given the benefits of the NDM link and the cumulated DGD along the link, in this section we provide a deeper investigation about the system performance improvement induced by both cumulated DGD and RDPS. To this aim we emulated a 19 channel WDM system transmitted over a 20x100km SMF-based DM link, with variable RDPS and variable amount of DGD. In Fig 5.9 we report the Q-factor of the central PDM-QPSK as a funtion of DGD and RDPS for NRZ, aRZ and iRZ pulses, respectively.

As a general result we observe that for all the pulse formats both DGD and RDPS improve system performance. For NRZ and aRZ pulses, DGD and RDPS seem have the same effectiveness; indeed doubling the RDPS from 100 $[ps/nm]$ to 200 $[ps/nm]$ or doubling DGD from 2.5 $[ps]$ to 5 $[ps]$induces the same Q-factor improvement. Furthermore for high RDPS (larger

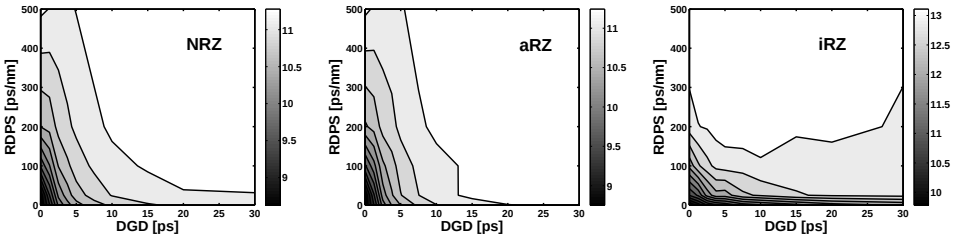


Figure 5.9: Q-factor for 19-channel NRZ, aRZ and iRZ-PDM-QPSK systems relying on SMF-based DM link with variable RDPS and variable cumulated DGD.

than 300 $[ps/nm]$) the DGD effect is masked by the large dispersion cumulated along the link; on the other hand, high cumulated DGD (larger than 10 $[ps]$) mask the impact of RDPS. From Fig. 5.9 (right) we infer that increasing the RDPS is more effective than improving the DGD to mitigate nonlinearities in the iRZ-PDM-QPSK system. Indeed the DGD improvement is remarkable only for links in which the RDPS is smaller than 100 $[ps/nm]$. We still ascribe this behavior to the sensitivity of iRZ pulse interleaving to PMD, as discussed in Sec. 5.4.

5.6 XPM suppressor impact on DOP and nonlinear-induced SOP rotation

In this section we provided an intuitive explanation of the reasons behind the effectiveness of the XPM suppressor against XPolM. To this aim we analyzed the XPolM-induced SOP rotation and channel depolarization *along* propagation, getting a picture of the SOP and Degree Of Polarization (DOP) of a reference signal at specific coordinates along the optical line. We simulated a WDM system based on a central Constant Wave (CW) signal of power -10 dBm surrounded by 18 NRZ-PDM-QPSK signals with power of 5 dBm each. The CW channel was oriented along the $\widehat{S_1}$ axis, while the PDM-QPSK channel were randomly oriented. The CW channel allows us to easily measure the DOP and to better visualize its SOP rotation during the propagation.

The transmission link was based on 20×100 km spans of SMF. Three different setups were considered: 1) a DM link with zero RDPS (labeled DM0), 2) the same DM link in which a XPM suppressor is applied at the end of each span (labeled DM0+XPM supp) and 3) a NDM link. In all these links, pre-compensation was not applied and the overall cumulated dispersion was set to zero.

We start showing in Fig. 5.10 the CW DOP at different points along the link, discretized with steps of 1 km. The measurement was repeated for five different random seeds, which correspond to random interfering channel pattern realizations and random SOP orientations. Concerning the DM0 link, in Fig. 5.10(left) we note a monotonically decreasing DOP for increasing distance, with a regular repeating pattern. Within the effective length of each transmission SMF fiber the DOP experiences an abrupt change, and then slows down in the following low-power km. Such a pattern induces an almost parabolic decrease of the DOP from span to span, with different curvatures depending on the random seed. The parabolic behavior ceases towards the end of the link because of the decrease of the power levels.

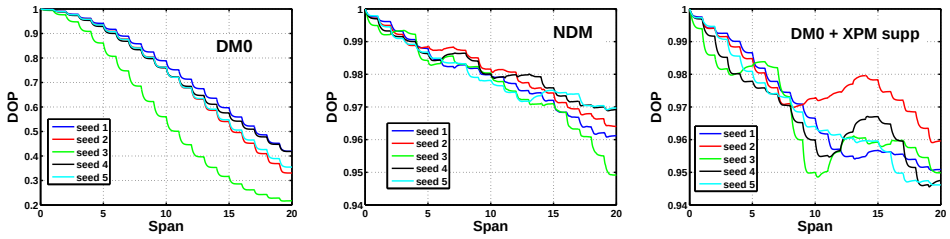


Figure 5.10: CW DOP decreasing along the link, when it propagates with 18 NRZ-PDM-QPSK channels in three links: DM link with zero RDSP (DM0), DM0 link in which XPM suppressor at the end of each span (DM0 + XPM supp) and NDM link.

For the NDM link (center figure) and the DM0 link with XPM suppressor (right figure), we observe that the DOP evolution along the link is much more irregular. Furthermore the final DOP (after 20 spans) is by far larger than for the DM0 link, indicating a smaller XPolM-induced depolarization (please note the different scales). The larger depolarization induced by XPolM in the DM0 link without XPM suppressor is an indicator that the performance is more sensitive to the interfering channel SOPs. On the contrary in the NDM and the DM0 link with XPM suppressor we expect a chaotic behavior along propagation, which does not allow specific bad orientations of interfering SOPs to add constructively (i.e., resonantly) along propagation [2].

Now we move to analyze the CW SOPs rotation induced by XPolM in the three transmission links. We fixed the time at the beginning and at the end of the first 35.7 ps (symbol duration at 28 Gbaud) of the CW channel, reporting in Fig. 5.11 their SOP trajectory during propagation, with steps of 1 km.

During the propagation along the DM0 link (Fig. 5.11 left), we note that the two CW samples evolved on the Poincarè sphere with an almost periodic pattern similar to a “spring”.

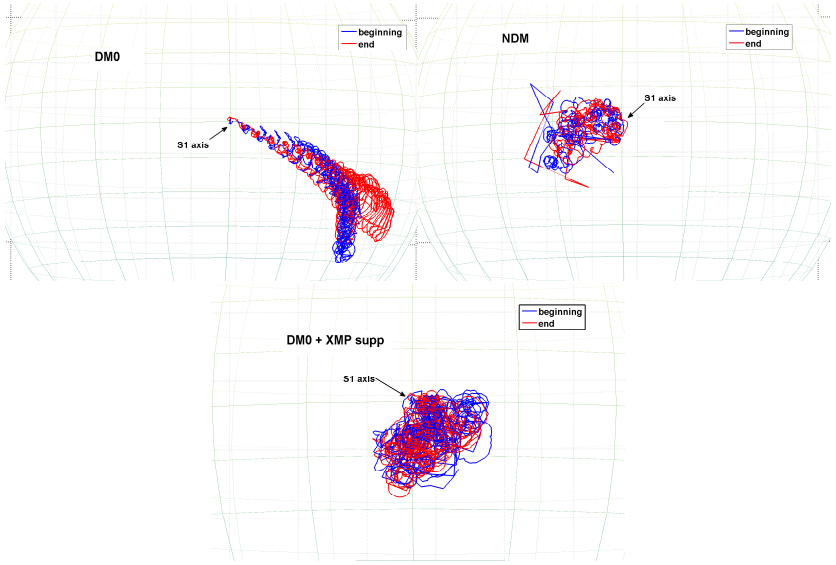


Figure 5.11: CW SOP evolution along the link due to the XPolM-induced rotation coming form the propagation with 18 NRZ-PDM-QPSK interfering channels in a DM0 with and without XPM suppressor and in a NDM link. Two different samples are considered: beginning and end of 35.7 ps time slot. Random seed 1.

These trajectories have a periodic component (a "drifts") and an uncorrelated component that leads to a random-like motion. The difference between the two final SOPs provides a direct estimation of the depolarization within a symbol period.

In Fig. 5.12 we zoomed over such a SOP trajectory, visualizing the SOP evolution of a single sample along five spans. In the figure we indicated the starting/ending coordinates of the optical fibers encountered along propagation. For each transmission fiber we also indicated the coordinate of the corresponding effective length.

We note that the CW follows similar trajectories during the propagation along each span. The reason is strictly related to the chosen zero RDPS that perfectly re-aligns in time all channels at the beginning of each span, thus creating a periodic interference with period equal to the span length. This means that, under a first approximation, the nonlinear rotation induced by the interfering patterns is the same in each span except for a different starting point, whatever the initial SOP of the interfering channels. Hence, this perfect re-alignment creates a constructive addition of the SOP trajectories that forces the CW SOP to follow a

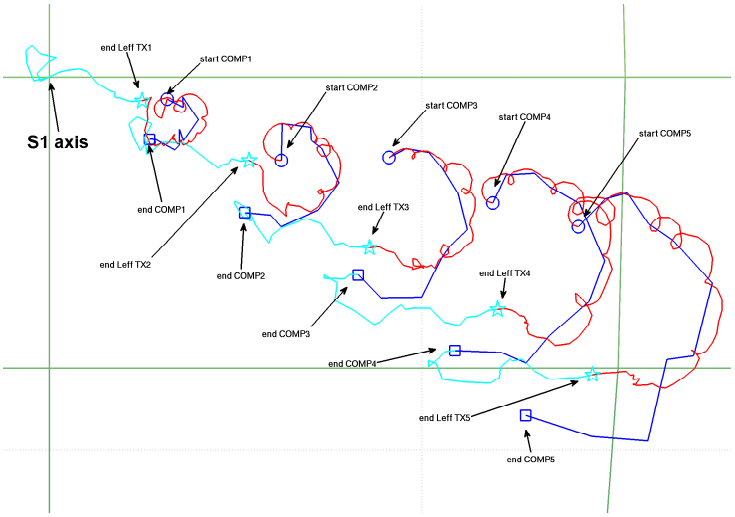


Figure 5.12: DM0 link: SOP evolution along the first five span of a single CW sample.

long trajectory on the Poincarè sphere.

We also note that the DCF fibers (purely linear in our simulations) induce a SOP change through their GVD[1]. The DCF-induced trajectory is of opposite direction w.r.t the one induced by the transmission fiber. Hence, XPolM may be partially resonant in dispersion compensated links, since its contribution slowly changes direction along propagation. Without the DCF each ring of the apparent "spring" motion is not recovered, thus the SOP trajectory changes *quickly* and chaotic along propagation, as shown in Fig. 5.11 (center and left).

Indeed, Fig. 5.11 (center) shows for the NDM the predicted chaotic behavior (random walk) that averages itself without letting the SOP travel long distances over the sphere. Also the distance between the final SOP of the two samples and consequently the depolarization are smaller than after the propagation along the DM0 link. The application of the XPM suppressor to the DM0 link seems to delete the periodicity of the SOP pattern, leading to an overall SOP evolution like a random walk, as in the NDM link.

Based on these observations, we can conclude that in a DM map the XPolM-induced rotations at each span add constructively, inducing a larger signal depolarization and consequently a stronger Q-penalty. The application of the XPM suppressor at the end of each span breaks such a resonance reducing both the depolarization and the Q-factor penalties, as in the NDM link.

[1]GVD does not depolarize in the frequency domain, but here we are in the time domain.

Chapter 6

Mode Division Multiplexing using an LCOS-based Spatial Modulator

As discussed in Chapter 2 mode division multiplexing (MDM) over few-mode fibers (FMF) has recently appeared as a promising alternative to keep up with transmission capacity growth [41, 42, 43, 44]. In this chapter we summarize the experimental work made at Alcatel-Lucent Bell-Labs France on this topic, focusing mainly on the modal cross-talk induced by a mode converter based on a liquid-crystal on silicon (LCOS) spatial modulator. For our transmission experiments, we employed a LCOS-based mode converter spatial modulator and a prototype FMF [45]. This LCOS-based approach to mode conversion is attractive because of the possibility to reconfigure the phase plates to any desired mode conversion. The FMF has the advantage of exhibiting very large effective-index differences and very large group delays between different modes and thus low linear crosstalk between modes, with only 0.22dB/km loss.

In Sec. 6.1 we describe the mode converter and the scheme of mode multiplexer and demultiplexer. In Sec. 6.2 we evaluate the impact of the modal cross-talk induced by imperfect mode multiplexing/demultiplexing in LCOS device on a 100 Gb/s PDM-QPSK performance. In Sec. 6.3, we demonstrate the transmission of two 100Gb/s PDM-QPSK signals on two modes of the FMF, which we shall call LP11a and LP11b, using an additional coherent detector and a more sophisticated DSP with respect to the standard single mode transmission. More

details on the receiver algorithms and other transmission experiments obtained using the combination of LCOS-based mode converter and FMF are discussed in [42, 72, 73, 74, 75].

6.1 Mode Conversion and (De-)Multiplexing

We realized in a single device two distinct functions: mode conversion and mode multiplexing/demultiplexing. For what concerns mode conversion, an LCOS-based spatial modulator is used to change the phase of the transverse distribution of the optical field in a so-called 4f-correlator configuration, where the spatial light modulator is placed in the center, as depicted in Fig. 6.1. Input fiber and output fiber are collimated by a lens, whose distance from the fiber output corresponds to its focal length 'f'. As a result, the two-dimensional Fourier transform of the spatial light distribution in the fiber is obtained in the central plane, spaced by another 'f' between the lenses. This Fourier transform is modulated in phase using a multiplicative mask, which is programmed on the LCOS. The resulting field distribution in the output fiber can be easily calculated by applying the inverse Fourier transform to the product of mask and incoming Fourier transform. In Fig. 6.1 the signal coming from a single-mode fiber (SMF) is converted into one of the modes supported by the output FMF: LP11a.

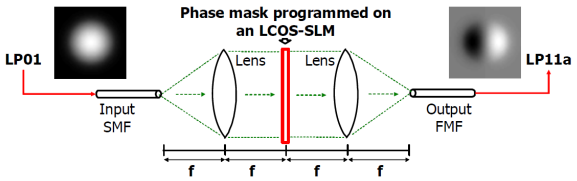


Figure 6.1: Scheme of mode conversion using spatial light modulator (SLM) phase mask on the Fourier plane of a 4f configuration.

By mode multiplexing, we mean the combination of more than one converted mode into the output FMF. In a simple but inefficient way this can be achieved by using free-space half-mirrors after mode conversion in a scheme like the one of Fig. 6.2. Each half mirror adds a 3dB loss. In our case we have built two devices that can be used as both multiplexers or demultiplexers. One of them has two SMF inputs/outputs and the other has four SMF inputs/outputs. We choose to use as a multiplexer the one with two SMF inputs and as a demultiplexer the one with four SMF outputs, where only 3 ports can be used for mode conversion and the fourth is directly coupled into SMF from FMF.

These two devices share a single LCOS of 1920x1080 pixels, operated in reflective mode.

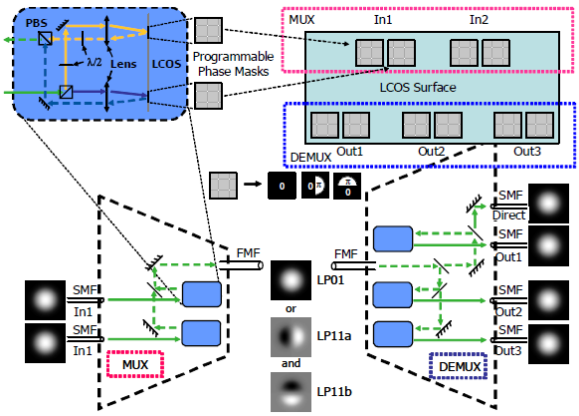


Figure 6.2: Scheme of mode multiplexer (2x1) and demultiplexer (1x4). The mode converters are realized with a polarization diversity scheme depicted in the inset and each beam uses a spot of the LCOS surface that can be programmed with on of the three possible masks.

The 2x1 converter uses the lower size of the LCOS screen. At each one of its two inputs, the light beam from the collimated SMF pigtail is split along two optical paths, one for each polarization, by a polarization splitter, and hits the LCOS onto two of four possible spots. Note that we use only phase modulation, whereas ideal mode conversion would require phase and amplitude masks, at the expense of an excessively complex design. Each spot on the LCOS device counts approximately 80x80 pixels and is programmed with the phase mask corresponding to the desired mode, i.e., according to the profiles in Fig. 6.2. Light from all four paths is sent back to two polarization beam combiners, then into a 2x1 free-space combiner and collimated into the FMF. At the receiver end, the 1x4 demultiplexer is designed similarly, but with four SMF-fiber pigtails as output ports (hence 6 spots on the LCOS phase modulator for polarization diversity plus a directly connected output port without mode conversion), the input being the FMF. It uses the upper part of the LCOS device.

6.2 LCOS-induced cross-talk

In this section we characterize our LCOS-based mode multiplexer fed with a 100Gb/s PDM-QPSK signal in a transmission experiment over 40km of few-mode fiber. The tolerance of the Q-factor performance to the crosstalk-induced phase mismatch is quantified.

6.2.1 Experimental set-up

In our set-up, depicted in Fig. 6.3(a), the light from a tunable laser is passed into an integrated transmitter. The transmitter uses a serializer to produce four 28Gb/s electrical pseudo-random signals of length $2^{15}-1$, each shifted by 8192 bits, which feed a quad-driver and a polarization-multiplexed nested Mach-Zehnder modulator. It generates a data stream at 112Gb/s PDM-QPSK, including 12% protocol and forward error correction (FEC) overhead. This stream is sent to an optical amplifier, connected to an input of the mode multiplexer, for mode conversion and coupling into the FMF. The LCOS surface hosts ten spots programmed with the phase mask corresponding to the desired mode, as depicted in Fig. 6.3(b). Each mode needs two spots for its horizontal (H) and vertical (V) polarization components. In our configuration the only used multiplexer input "In1" is set to convert the fundamental mode into the LP11b mode. The converted signals are then collimated into the FMF at the multiplexer output. This 40-km long FMF supports up to 4 spatial modes (LP01, LP11, LP21, and LP02), with low attenuation (0.22 dB/km) and large effective areas (>120 μm^2) for all guided modes. The differential mode group delay (DMGD) per unit length between LP01 and LP11 is quite large, at 4.35 ps/m (4872 symbols after 40km at 28GBaud) and their effective-index difference is $> 10^{-3}$, which suggests negligible mode coupling as in [76].

At the receiver end, the FMF goes into the 1x3 mode-demultiplexer, which performs mode conversion using the remaining 6 spots on the LCOS phase modulator and collimates the light to the three SMF-fibers at the outputs. The first two outputs are used for the back-conversion of the LP11b and LP11a modes of the FMF and are detected by a noise-loaded coherent receiver with joint mode/polarization diversity. The third output is not converted and it is sent to an SMF which captures mainly the LP01 mode of the FMF, which is measured with a power meter. Outputs 1 and 2 go through variable attenuators, optical preamplifiers and filters for setting the desired received OSNR before the coherent receivers. Each coherent receiver uses a free-space coherent mixer, a free-running local oscillator and four balanced photodiodes in a polarization diversity configuration. Sampling is performed with two real-time oscilloscopes (16GHz analog bandwidth) synchronously triggered. They provide 4 complex signals containing the optical field. To discriminate between the degenerated modes LP11 along the two polarization axes, a 4x4 MIMO equalizer is needed, as opposed to the conventional 2x2 MIMO equalizer used for polarization demultiplexing over SMF [42].

One important characteristic of the LCOS device is its possibility to independently program its pixels on a gray scale with many discrete values. This scale corresponds through a given monotonic function to a phase shift φ in the range $[0:2\pi]$. The masks required for LP11 conversion are programmed using a simple phase shift $\varphi = \pi$ on half of the surface.

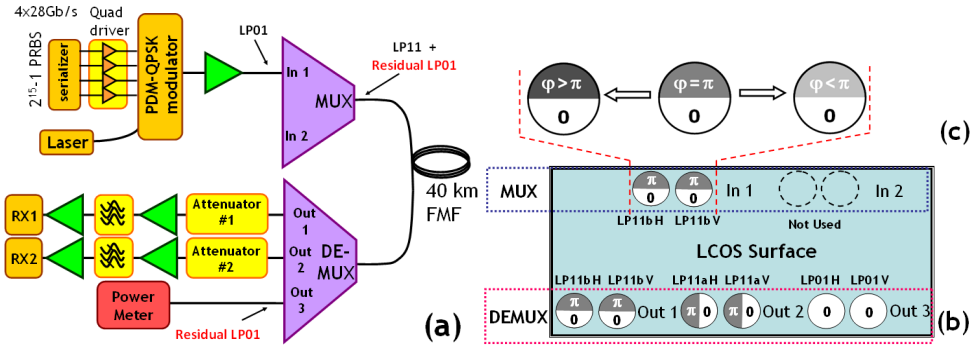


Figure 6.3: System setup (a); LCOS surface with phase mask (b) and φ represented as a gray value (c).

The relationship between the gray value and the phase shift is dependent on the position of the mask on the LCOS surface, which we take into account using a mask-specific additive phase term φ_M. To compensate for φ_M we apply a gray correction parameter ε to the phase shift $\varphi = \pi + \varepsilon\pi + \varphi_M$, which is represented as a different gray value in Fig 6.3(c). Ideally, a phase shift $\varphi = \pi$ should be applied by the spatial light modulator, which means that an optimum value of the gray correction parameter $\varepsilon' = -\varphi_M/\pi$ should be found. A first method to quantify this optimum value is to minimize the residual LP01 power that propagates in the FMF whenever a "sub-optimal" phase mask of the multiplexer does not convert all the power of the LP01 into LP11. A second method for its optimization is to measure the phase shift impact on the Q-factor. Both these techniques will be applied for the optimization of the LP11 phase mask used at the input "In 1" of the multiplexer. The masks of the demultiplexer have already been optimized with similar procedure.

6.2.2 Experimental results

In a first test we transmitted a signal al $\lambda = 1533.47nm$ through the "In 1" port of the multiplexer (LP11b) and we varied the gray correction parameter of both horizontal (ε_H) and vertical (ε_V) polarization phase masks. This variation generates a residual LP01 mode at the transmitter side, which is coupled into the FMF together with the LP11b mode. The power of this residual LP01 mode is demultiplexed through the "Out 3" port and measured. In Fig. 6.4(a) the LP01 power at the "Out 3" port is reported as a function of ε_H and ε_V. For each curve, only the ε of one polarization's mask is varied, while the ε of the other mask is set to

its optimum value. We found that these two phase masks have different optimum values for ε, i.e., $\varepsilon'_H = 0.03$ and $\varepsilon'_V = 0.09$ give the minimum LP01 power corresponding to the lowest cross-talk contribution of the "In 1" port mode converter. Because of these different optimum ε' values, we will consider only variations $\Delta\varepsilon = (\varepsilon'_H - \varepsilon_H) = (\varepsilon'_V - \varepsilon_V)$ with respect to the optimum couple of values, hence $\Delta\varepsilon = 0$ corresponds to optimum $\varepsilon = \varepsilon'$ for each polarization ($\varphi = \pi$). Fig. 6.4(b) depicts the residual LP01 power as a function of $\Delta\varepsilon$. We observe a rapid residual LP01 power increase for increasing $|\Delta\varepsilon|$, which becomes larger than 5 dB for $|\Delta\varepsilon| > 0.2$. From our definition of phase shift this value corresponds to a phase mismatch of ±36°. In a second test we measured the Q-factor of LP11 by varying the $\Delta\varepsilon$ in the range of [-0.4:0.4], which corresponds to a phase mismatch of ±72°. In Fig. 6.5(a), LP11 Q-factors of both polarization tributaries are reported. Despite a constant offset in their performance, we measured a Q-factor penalty below 1dB in the range of ±27° for both polarizations. This corresponds to an interfering power increase of roughly 3 dB, as depicted in Fig. 6.4(b).

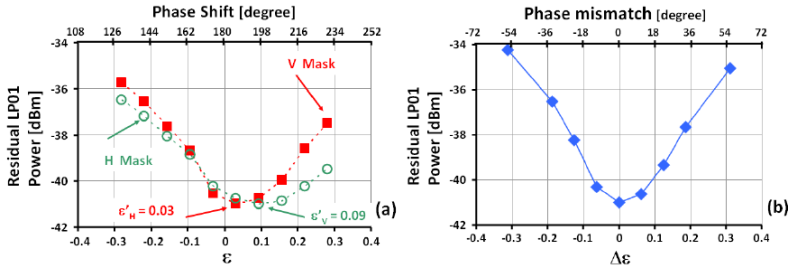


Figure 6.4: Residual LP01 power ("Out 3"), for different ε of the horizontal (H) and vertical (V) polarization phase mask (a), and for various $\Delta\varepsilon = (\varepsilon'_H - \varepsilon_H) = (\varepsilon'_V - \varepsilon_V)$, where $\Delta\varepsilon = 0$ corresponds to optimum ε' for each polarization (b). The corresponding phase shift and phase mismatch are reported as reference.

We then repeated the ε optimization using two other wavelengths, namely 1532.29nm and 1534.64nm. Interestingly, here, the optimum values differed from the values obtained at 1533.47nm, being $\varepsilon'_H = 0.09$ and $\varepsilon'_V = -0.09$ at 1532.29 and $\varepsilon'_H = 0.09$ and $\varepsilon'_V = 0.09$ at 1534.64nm respectively. Starting from these ε' values, we measured the LP11 Q-factor penalty by varying phase mismatch as in the previous case. The results are depicted in Fig. 6.5(b) using empty and full symbols for the two polarization tributaries. We observe a similar phase mismatch impact for the three wavelengths although ε'_H and ε'_V are different for the three wavelengths. This shows that the gray correction parameter ε, as defined above, shows not only the spatial dependency but also the wavelength dependency of the LCOS. However,

the varying ε' values could cause severe Q-factor penalties in potential WDM transmissions, where all wavelengths of one mode could be converted by the same phase mask.

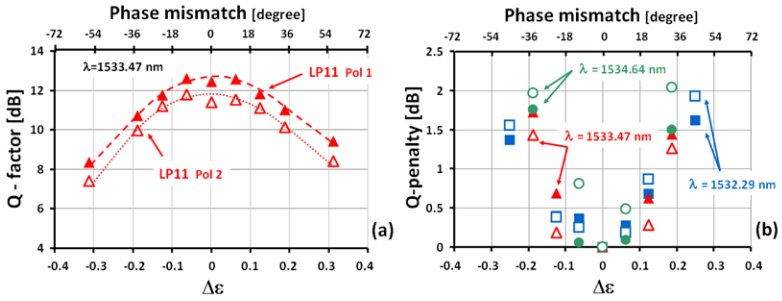


Figure 6.5: LP11 Q-factor (a) and LP11 Q-factor penalty for three wavelengths, versus $\Delta\varepsilon$ and phase mismatch (b).

6.3 Transmission Experiment

6.3.1 Experimental set-up

In the transmission set-up, depicted in Fig 6.6, the light from a laser at 1533.47nm is passed into an integrated transmitter, which generates a data stream at 112Gb/s PDM-QPSK. This stream is replicated along two fiber paths decorrelated by several thousands of symbols using a SMF patch-cord, and fed into two optical amplifiers, connected to the two inputs of the mode multiplexer. For a transmission experiment using the fundamental mode, only one single input is used and the black mask is programmed on the LCOS for not performing any mode conversion. In the case of a transmission of mode division multiplexed signals, the two inputs are used and the masks corresponding to LP11a and LP11b conversion are programmed on the two converters used by the multiplexer. When the device is operated with LP01 mode (no mode conversion) the losses of the multiplexer and demultiplexer are about 16dB. However, when the masks for LP11 conversion are used, the losses increase up to 25dB because of the mode conversion extra losses.

The mode demultiplexer receives the signals from the FMF and performs mode conversion before sending the modes out on different SMFs connected on its output ports. Each port is connected on a series of components that form a receiver with adjustable received power. We use an attenuator to change the power that enters into a first Erbium doped fiber amplifier

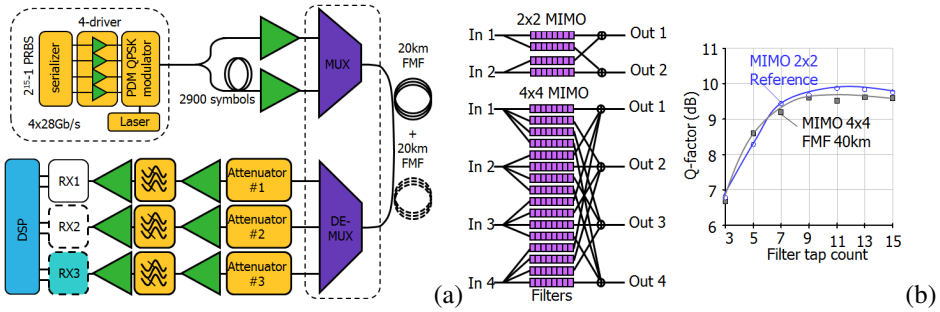


Figure 6.6: Experimental setup (a). Representation of the FIR filters used for mode/polarization demultiplexing in standard single mode case and 4x4-MIMO dual-mode case. Q-factor vs. tap number (T/2-spacing) on the right graph (b).

(EDFA), a filter for noise suppression and a second EDFA for insuring the required powers on the photodiodes of the coherent receiver. The power at the input of the first EDFA determines the signal to noise ratio and thus the performance. In all configurations, the power into each receiver is varied using dedicated attenuators and the performance is measured. LP01 is detected with a constant phase mask and using a single receiver. When the phase masks are set to select LP11a and LP11b, two receivers are used to perform joint detection and signal processing. Furthermore, when the residual information on the LP01 mode is employed, a third receiver can be added which detects the residual LP01 optical field. In the next section we will describe only the transmission results obtained using two receivers. The benefits obtained adding the third receiver are discussed in [72].

The two receivers provide 4 complex signals representing the optical field. The structure of the 4x4 MIMO equalizer, needed to discriminate between the degenerate modes LP11 along the two polarization axes is depicted in Fig. 6.6(b). The 4 complex signals are split into 4, each sent into an FIR filter having up to 15 taps. Each of the 4 outputs is the combination of the 4 inputs, filtered by a dedicated FIR. We vary the FIR tap count in order to estimate the required length of the equalizer. We use 100Gb/s PDM-QPSK performance in back-to-back as a reference (MIMO 2x2 over SMF). An almost flat Q-factor is measured from 9 taps to 15 taps, both for 100Gb/s PDM-QPSK and 2x100Gb/s MDM after 40km, as shown in Fig. 6.6(b). The 4x4 MIMO receiver operates in blind mode, using traditional CMA for FIR update. The complexity of such an architecture is only doubled (per transmitted bit) compared to standard single mode operation, as 16 adaptive filters are required to process 2x100Gb/s (4x4 MIMO) compared to 4 filters for 1x100Gb/s (2x2 MIMO). It should be noted that CMA

tends to converge to the more powerful signal tributaries. In particular, in presence of multiple inputs it is necessary to correctly initialize the filter taps for allowing the detection of all the polarization/mode tributaries. Failing to do so would cause different outputs converging to the same input. We performed identification after detection by comparing the delays of the different tributaries.

6.3.2 Experimental Results

The first experimental results have been recorded after 20km of FMF fiber. To ease comparisons, we choose to focus on the same received power in all configurations namely -31dBm. In Fig. 6.7(a), the fundamental mode LP01 shows 1.1dB sensitivity penalty after 20km compared to the reference back-to-back. After the same distance, the average Q-factor of the LP11a + LP11b modes detected with the 4x4 MIMO receiver is reduced by another 0.6dB with respect to the LP01. After 40km of FMF, the modes LP11a + LP11b are further degraded, but by no more than 1.1dB. The values represented in Fig. 6.7(a) are averaged over all the received mode/polarization tributaries (2 for the LP01 and 4 for the LP11). Fig. 6.7(b) depicts the actual performance of all mode/polarization tributaries of LP11.

The Q-factors of mode LP11a-polarization X appear significantly lower (by typically 2.5dB) than the other Q-factors. This is attributed to degraded characteristics of our mode mux/demux along one polarization. By correcting the imperfections that limit the performance of this first prototype of LCOS based mode multiplexer/demultiplexer we can expect to reach longer distances.

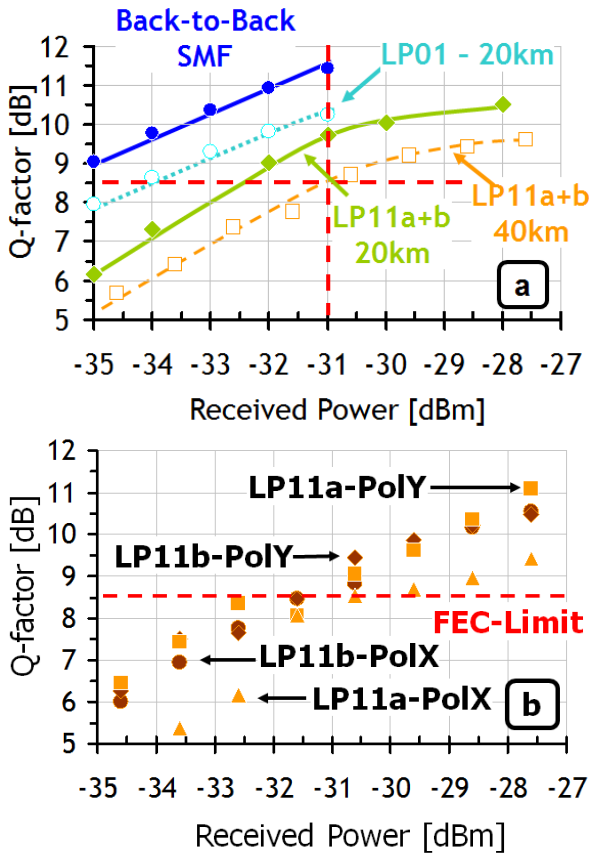


Figure 6.7: (a) Q-factor versus received power for LP01 (back-to-back and 20km FMF transmission) and LP11a+b (20km and 40km FMF transmission); (b) Details of Q-factor of the 4 mode/polarization tributaries of LP11 after 40km.

Chapter 7

Conclusions

In this thesis we analyzed a new kind of cross-nonlinear effect that depends on neighboring channels' power and state of polarization (SOP), called XPolM to distinguish it from the XPM, which depends only on the neighboring channels' power. Thanks to the possibility in Optilux software [3] to take into account separately the nonlinear terms of the propagation equation, the impact of such a nonlinearity on the PDM-QPSK performance was investigated as a function of some transmission parameters such as the channel power, spacing and state of polarization (SOP). We also identified when the bit error rate (BER) is dominated by XPolM in a wide range of transmission setups.

In Ch. 3 we showed how the XPolM-induced SOP rotation and PDM-QPSK constellation distortion depend on the rotation axis orientation. We showed that the XPolM can act as an additional XPM term, inducing a differential phase noise or can induce both phase and amplitude noise, i.e., a cross-talk between the two PDM-QPSK subcarriers.

In Ch. 4 we quantified the nonlinear penalty in hybrid systems (PDM-QPSK–OOK) and homogeneous systems (PDM-QPSK and PDM-BPSK).

By analyzing the cross-channel nonlinear effects induced by 10 Gb/s OOK channels on an 112 Gb/s PDM-QPSK channel, we showed that both XPM and XPolM decrease with increasing walk-off and that the penalties coming form spectrally distant channels are mainly due to XPolM while the penalties coming from the neighboring channels are mainly due to XPM. We also identified the channel spacing beyond which the XPolM induces stronger impairments than XPM. The role of the carrier phase estimation based on the Viterbi and Viterbi algorithm against both XPM and XPolM was clarified, showing the number of taps that maximize the performance in presence of XPM and XPolM. There is no general rule

about which is the dominant cross-nonlinear impairment in hybrid systems, since it depends on the length of the phase recovery algorithm. Finally we showed that the XPolM-induced penalty is minimized when the interfering OOK channels have the same polarization as one of the two PDM sub-channels, i.e., when the XPolM acts as an additional XPM term and there is no cross-talk between the two subcarriers.

Moving to the homogeneous 112 Gb/s PDM-QPSK transmission systems, we confirmed that both XPM and XPolM decrease for increasing walk-off but at different rates. Furthermore we observed that in dispersion managed (DM) links almost the entire penalty is due to XPolM, while XPM adds a negligible contribution to the overall penalty. In non-dispersion managed (NDM) links the difference between XPolM and XPM is almost null. It turns out that the large walk-off of a NDM link strongly reduces both XPM and XPolM, leading to a distortion induced by power fluctuations (XPM) similar to the distortion induced by the SOP fluctuations (XPolM). We also showed the benefits of the time-interleaving return-to-zero (iRZ) pulse format in mitigating the cross-nonlinear effects.

In the last part of the chapter we numerically analyzed the nonlinearities impact on the performance of a 42.8 Gb/s PDM-BPSK in homogeneous WDM transmission systems. Different pulse shapes were analyzed: NRZ and iRZ (66%, 50% and 33% carved), showing both the mean Q-factor and the Q-factor fluctuations. We observed that XPolM is the dominant effect for NRZ and iRZ 66% shapes, while for iRZ 50% and iRZ 33% XPolM and XPM are comparable. The benefits of iRZ are also confirmed not only in reducing the average cross-nonlinearities, but also in reducing the performance fluctuations induced by different SOP orientations and transmitted patterns.

In chapter **5** we compared different optical solutions to improve the resilience of coherent 112 Gb/s PDM-QPSK WDM transmissions against cross-channel nonlinearities. We showed that decorrelating the channels, through either PMD, or delay-line XPM suppressor, or by removing dispersion management, improves performance and reduces the gap among iRZ, NRZ, and aRZ pulse formats. In iRZ-PDM-QPSK the worsening of the pulses' time-interleaving due to PMD is more than offset by the positive PMD-induced depolarization effect that reduces XPolM. We find that the best option is to use iRZ-PDM-QPSK in a DM link with an XPM suppressor at each span to decorrelate channels without neither compromising the time-orthogonality of the PDM tributaries, as with PMD, nor inducing more nonlinear self-effects, as with NDM. Unlike [68], for PDM-QPSK systems we did not find an optimum delay value for the XPM suppressor, so the best option is to maximize the decorrelation among channels. Furthermore by comparing the degree of polarization (DOP) and the SOP trajectory along a DM link with and without the XPM suppressor and NDM link, we clarified

the reason why the XPM suppressor is also effective against the XPolM.

In Ch. 6 the transmission of two 100Gb/s PDM-QPSK signals at the same wavelength has been demonstrated over two modes of a 40km few-mode fiber. The key components are a reprogrammable free-space mode mux/demux, based on a Liquid Crystal on Silicon device (LCOS), a prototype few-mode fiber with low mode coupling and two polarization- and mode-diversity coherent detectors supported by MIMO processing. We characterized the LCOS-based mode multiplexer fed with a 100Gb/s PDM-QPSK signal in a transmission experiment over 40km of few-mode fiber. The tolerance of the Q-factor performance to the crosstalk-induced phase mismatch was quantified, by introducing defects in the phase mask of a programmable LCOS-based mode converter generated a residual LP01 mode. We have measured that a multiplexer phase mismatch of more than 27° induces a Q-factor penalty larger than 1 dB.

Finally in the Appendix we discussed three different rules to correctly simulate the cross-nonlinearities, showing also the effects of their wrong choice on the simulated performance. In particular we showed the importance of a correct choice of the nonlinear step of the Split-Step Fourier Method and the sequence length. We showed that an incorrect evaluation of all these parameters can induce an artificial overestimation of the cross-nonlinearities.

Appendix A

Stokes Representation

We first define four matrices:

$$\sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \sigma_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad \sigma_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \sigma_3 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$$

which are called the ***Pauli matrices***[1]. The first matrix is the 2x2 identity matrix, which we also indicate by I.

We now transform any 2x1 complex vector $\overrightarrow{A}$ into a 4x1 real vector $\underline{a} = [a_0, a_1, a_2, a_3]^T$, by defining $a_i \stackrel{\triangle}{=} \overrightarrow{A}^{\dagger} \sigma_i \overrightarrow{A}$, which is a real quantity because the Pauli matrices are Hermitian. We call this the *Stokes representation* of the field $\overrightarrow{A}$. If we factor out the common phase term and express the field as:

$$\overrightarrow{A} = \begin{bmatrix} A_x \\ A_y \end{bmatrix} = \begin{bmatrix} A_x e^{i\phi_1} \\ A_y e^{i\phi_2} \end{bmatrix} = e^{i(\phi_x+\phi_y)/2} A \underline{\widehat{A}}$$

where $\phi = \phi_y - \phi_x$ is the differential phase, and $\widehat{A} = \begin{bmatrix} \cos\chi \, e^{-i\phi/2} \\ \sin\chi \, e^{i\phi/2} \end{bmatrix}$ is the unit Jones SOP vector, with $A_x = A\cos\chi$, $A_y = A\sin\chi$, then it is easy to verify that

[1] The numbering of the original Pauli matrices, used in quantum mechanics, is different: the labels of the above matrices 1 2 3 are shifted cyclically by one, namely they are 3 1 2. Our labeling is conventional for the Poincaré representation.

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} A_x^2 + A_y^2 \\ A_x^2 - A_y^2 \\ 2\Re\left[\overrightarrow{A}_x^* \overrightarrow{A}_y\right] \\ 2\Im\left[\overrightarrow{A}_x^* \overrightarrow{A}_y\right] \end{bmatrix} = A^2 \begin{bmatrix} 1 \\ \cos 2\chi \\ \sin 2\chi \cos\phi \\ \sin 2\chi \sin\phi \end{bmatrix} \tag{A.1}$$

It is thus clear that the representation $\underline{a}$ of the field $\overrightarrow{A}$ is faithful up to a common, arbitrary phase factor, i.e., the family of fields $\overrightarrow{A}e^{i\delta}$, δ real, all map into the same vector $\underline{a}$. The *Poincaré representation* of the field $\overrightarrow{A}$ is the vector $\overrightarrow{a} \stackrel{\triangle}{=} [a_1, a_2, a_3]$, which can be considered as the projection on the 3D space of the 4D Stokes vector $\underline{a}$. The *Poincaré representation* of its SOP Jones vector $\widehat{A}$ is the unit vector $\widehat{a} \stackrel{\triangle}{=} [a_1, a_2, a_3]/A^2$, which can be plotted as a point on the 3-dimensional unit sphere, the so-called Poincaré sphere.

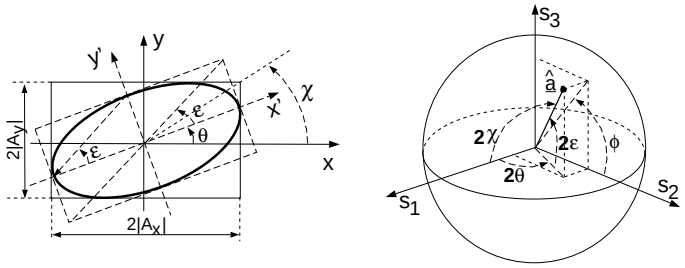


Figure A.1: (Left) Jones SOP parameters; (Right) Stokes SOP parameters on Poincaré sphere.

Fig. A.1 shows the relation between the polarization ellipse [77] giving the parameters of the Jones SOP vector $\underline{\widehat{A}}$ and the Poincarè unit vector $\underline{\widehat{a}}$, also called the *Stokes SOP vector*.

Note that one can write $\underline{\widehat{A}} = U_3(\theta) \begin{bmatrix} \cos\varepsilon \\ i\sin\varepsilon \end{bmatrix}$, where by $U_3(\alpha)$ we indicate the matrix that operates a *counterclockwise* rotation by an angle α:

$$U_3(\alpha) = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$$

From such a description, one gets an explicit expression of the Jones and Stokes SOP vectors as:

$$\widehat{\underline{A}} = \begin{bmatrix} \cos\theta\cos\varepsilon - i\sin\theta\sin\varepsilon \\ \sin\theta\cos\varepsilon + i\cos\theta\sin\varepsilon \end{bmatrix} \Leftrightarrow \widehat{\underline{a}} = \begin{bmatrix} \cos 2\theta\cos 2\varepsilon \\ \sin 2\theta\cos 2\varepsilon \\ \sin 2\varepsilon \end{bmatrix} \tag{A.2}$$

where the azimuth and elevation angles $2\theta, 2\varepsilon$ are also indicated in Fig A.1. Note the 2 multiplicative factor of θ and ε when moving between Jones and Stokes spaces.

In the figure $\widehat{S_1}$, $\widehat{S_2}$,, and $\widehat{S_3}$ are the Cartesian reference axes, and we note that their labeling (related to the labeling of the Pauli matrices) is such that linearly polarized SOPs ($\varepsilon = 0$) have $a_3 = 0$, i.e., they lie on the equator of the sphere.

Conversely, starting from an assigned Stokes SOP vector $\widehat{r} = [r_1, r_2, r_3]^T$, we find its associated Jones SOP vector $\widehat{R}$, as well as its orthogonal $\widehat{R_o}$ associated with $-\widehat{r}$ [2]as

$$\begin{cases} \widehat{R} = \begin{bmatrix} 1 + r_1 \\ r_2 + ir_3 \end{bmatrix} \frac{1}{\sqrt{2(1+r_1)}} \\ \widehat{R_o} = \begin{bmatrix} -r_2 + ir_3 \\ 1 + r_1 \end{bmatrix} \frac{1}{\sqrt{2(1+r_1)}} \end{cases} \tag{A.3}$$

Finally, given $\widehat{R}$ associated with Stokes vector $\overrightarrow{r} = [r_1, r_2, r_3]^T$, it is easy to check using the definitions that $\widehat{R}^*$ is associated with $[r_1, r_2, -r_3]^T$, i.e., conjugation in Jones corresponds to a change in sign in the third coordinate, the one relative to circular polarization.

[2]Both $\widehat{R}$ and $\widehat{R_o}$ are unique up to a common phase term. For instance, by adding a common phase rotation of π to the second eigenvector one gets the equivalent form $\widehat{R_o} = [r_3 - ir_2, -(1 + r_1)]^T / \sqrt{2(1 + r_1)}$.

Appendix B

SOP of a PDM-QPSK signal

A PDM-QPSK signal has two QPSK signals in the X and Y components of the polarized field. Here we want to evaluate its SOP in the Stokes space. As shown in App. A, calling $\overrightarrow{A}(t) = [A_x(t), A_y(t)]^T$ the electric field, the Stokes representation of the SOP is:

$$\overrightarrow{A} = \begin{bmatrix} A_x^2 + A_y^2 \\ A_x^2 - A_y^2 \\ 2\Re\left[\overrightarrow{A}_x^* \overrightarrow{A}_y\right] \\ 2\Im\left[\overrightarrow{A}_x^* \overrightarrow{A}_y\right] \end{bmatrix} = \left(|A_x|^2 + |A_y|^2\right) \begin{bmatrix} 1 \\ \widehat{a} \end{bmatrix} \tag{B.1}$$

where $\widehat{a} = [a_1, a_2, a_3]^T = \begin{bmatrix} \cos 2\theta \cos 2\varepsilon \\ \sin 2\theta \cos 2\varepsilon \\ \sin 2\varepsilon \end{bmatrix}$ is the Poincaré unit vector or instantaneous SOP vector and θ is the azimuth and ε the ellipticity of the SOP. The Poincaré sphere represents this vector $\widehat{a}$ as a point on the unit sphere[1]. In the main body of the thesis, to have a unique reference sphere at each temporal instant, we represent the SOP on a Poincaré with a radius equals the average signal power $|A|^2 \triangleq |A_x|^2 + |A_y|^2$. Consequently the magnitude of Stokes vectors are not 1 but equals the actual signal power.

The Poincaré sphere can be completely spanned by:

$$-\tfrac{\pi}{2} \leq \theta \leq \tfrac{\pi}{2}$$
$$-\tfrac{\pi}{4} \leq \varepsilon \leq \tfrac{\pi}{4}$$

[1] The main axes of the Poincarè sphere are called $\widehat{S_1}, \widehat{S_2}, \widehat{S_3}$.

θ is the angular coordinate along a parallel of the sphere, ε the angular coordinate along a meridian.

Let's now turn back to the problem of identifying the SOP of a PDM-QPSK signal on the Poincarè sphere. Both A_x and A_y at the sampling times and in absence of distortions take on the complex plane the values $\exp\left(i\left(\frac{\pi}{4}+m\frac{\pi}{2}\right)\right)$, $m=0,1,2,3$. Set now the observation time so that A_x at that time takes a specific value. A_y at the same time, being independently modulated from A_x, can be described on the complex plane as a vector rotated with respect to A_x by 0 or $\pm\frac{\pi}{2}$ or π radians. From (B.1) we can observe one of the following:

- $A_y = A_x, \Rightarrow \widehat{a} = \left[0, 2|A_x|^2, 0\right]^T$. Such a vector is a point on the equator of the Poincarè sphere, with direction $\widehat{S_2}$.
- $A_y = A_x e^{\pm i\pi} = -A_x, \Rightarrow \widehat{a} = \left[0, -2|A_x|^2, 0\right]^T$. The opposite vector as before.
- $A_y = A_x e^{\pm i\pi/2}, \Rightarrow \widehat{a} = \left[0, 0, \pm 2|A_x|^2\right]^T$. A vector with direction $\pm\widehat{S_3}$.

Note that, in any case, with the previous definition of the (X,Y) reference system, the PDM signal has the $\widehat{S_1}$ component equal to zero, hence the SOP, in absence of distortions, lies in the plane $\left(\widehat{S_2}, \widehat{S_3}\right)$. Fig. B.1 shows an example of a PDM-QPSK signal on the Poincaré sphere. At a given time the SOP is one of the points shown in the figure, so that over a sufficiently long observation time the plot takes the characteristic form of a cross in 3D Stokes space.

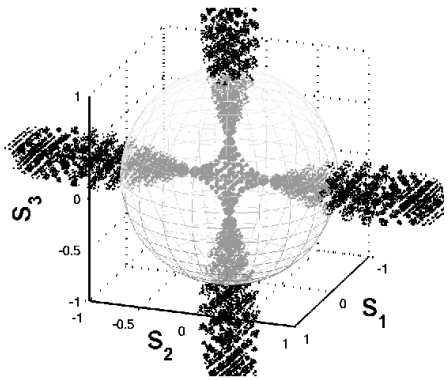


Figure B.1: SOP of a PDM-QPSK signal on the Poincarè sphere in absence of distortions. Magnitude of Stokes vectors are not 1 but equals the actual signal power. Radius of shown Poincaré sphere equals the average signal power.

Appendix C

Some Matrix Results

C.1 Exponential Matrix Expansion

We give here a simple method to explicitly evaluate the matrix exponential, based on the following [78]:

Theorem: Let the characteristic polynomial of (square) matrix A be

$$\Delta(\lambda) \stackrel{\triangle}{=} \det[\lambda I - A] = \prod_{i=1}^{m} (\lambda - \lambda_i)^{n_i}$$

where the m distinct eigenvalues $\{\lambda_i\}$ have multiplicities $\{n_i\}$. Let f and g be two arbitrary polynomials. If for every $i = 1, 2, ..., m$:

$$f^{(l)}(\lambda_i) = g^{(l)}(\lambda_i) \quad \text{for} \, l = 0, 1, ..., n_i - 1$$

(where $f^{(l)}$ indicates the $l-th$ derivative) then $f(A) = g(A)$.

When the above is true, we say that f and g agree on the spectrum of A. This theorem, which is an implication of the Cayley-Hamilton theorem, implies that *any* function f (including thus the matrix exponential) of an $n \times n$ matrix A can be made equal to a polynomial $g(\lambda) = \alpha_0 + \alpha_1 \lambda + ... + \alpha_{n-1} \lambda^{n-1}$ of degree at most $n-1$. As applications of the above, let's find the expression of the 2x2 matrix $f(B) = \exp^B = g(B) = \alpha_0 I + \alpha_1 B$. Let λ_1, λ_2 be the two distinct eigenvalues of B. Then the agreement on the spectrum of B gives

$$\begin{cases} \alpha_0 + \alpha_1 \lambda_1 = e^{\lambda_1} \\ \alpha_0 + \alpha_1 \lambda_2 = e^{\lambda_2} \end{cases}$$

whose solution is

$$\alpha_0 = \frac{\lambda_1 e^{\lambda_2} - \lambda_2 e^{\lambda_1}}{\lambda_1 - \lambda_2} \quad \alpha_1 = \frac{e^{\lambda_1} - e^{\lambda_2}}{\lambda_1 - \lambda_2}$$

C.2 Properties of Vector and Scalar Products

We list here some useful properties involving mixtures of vector and scalar products, given vectors a, b, $c \in \mathscr{R}^3$.

(i) $a \times b = -b \times a$

(ii) $a \times b \odot c = a \odot b \times c$ In words: we can exchange operators, for fixed position of vectors.

(iii) $a \times b \odot c = c \times a \odot b = b \times c \odot a$ In words: we can shift cyclically vectors, for fixed position of operators.

(iv) $a \times (b \times c) = (a \odot c)b - (a \odot b)c$

(v) $a \times (b \times c) = b \times (a \times c) + c \times (b \times a)$ In words: the rhs is the sum of two terms similar to the one on the lhs, but with swapping of vectors 1,2 for one, and of vectors 1,3 for the other.

(vi) The skew-symmetric matrix

$$[\underline{b}\times] \stackrel{\triangle}{=} \begin{bmatrix} 0 & -b_3 & b_2 \\ b_3 & 0 & -b_1 \\ -b_2 & b_1 & 0 \end{bmatrix}$$

corresponds to the vector product operation with vector $\underline{b} = [b_1, b_2, b_3]^T$. Hence for any vector c, we have: $b \times c = [b\times]c$. We also note that $[b_1\times] + [b_2\times] = [(b_1 + b_2)\times]$.

(vii) If $\widehat{\underline{b}}$ is a unit vector, we have the useful relation: $[\widehat{\underline{b}}\times][\widehat{\underline{b}}\times] = \underline{\widehat{b}\widehat{b}}^T - \underline{\underline{I}}$, where the matrix $\underline{\widehat{b}\widehat{b}}^T$ is called a *dyad*. We also have $[\widehat{\underline{b}}\times][\widehat{\underline{b}}\times][\widehat{\underline{b}}\times] = -[\widehat{\underline{b}}\times]$.

(viii) A generalization of the above is: $[\underline{d}\times][\underline{b}\times] = \underline{b}\,\underline{d}^T - \underline{d}^T\,\underline{b}\,\underline{\underline{I}}$

(ix) It may be useful to know the inverse of $A = \alpha I + \beta[\widehat{\underline{b}}\times]$. The determinant is $\det[A] = \alpha(\alpha^2 + \beta^2)$, which is nonzero if $\alpha \neq 0$. In such case the inverse is

$$A^{-1} = \frac{1}{\det[A]}\left[(\alpha^2+\beta^2)I - \alpha\beta[\widehat{\underline{b}}\times] + \beta^2[\widehat{\underline{b}}\times]^2\right] = \frac{1}{\det[A]}\left[\alpha^2 I - \alpha\beta[\widehat{\underline{b}}\times] + \beta^2[\underline{\widehat{b}\widehat{b}}^T]\right]$$

(x) For column vectors $\underline{b}$, $\underline{c}$ using (iv) we get: $[(\underline{b}\times\underline{c})\times] = (\underline{c}\underline{b}^T - \underline{b}\underline{c}^T)$. Hence using (viii) we finally get:

$$[(\underline{b}\times\underline{c})\times] = [\underline{b}\times][\underline{c}\times] - [\underline{c}\times][\underline{b}\times]$$

(can get it immediately immediately from (v))

Appendix D

How to correctly simulate cross-channel nonlinear effects

In this appendix we highlight three simple rules to correctly simulate cross-nonlinearities, showing also some artifacts that can arise with a non-correct setting of some simulation parameters.

D.1 Step length in Split Step Fourier Method

We start showing the correctly dimensioning of the nonlinear step in Split-Step-Fourier Method (SSFM). The SSFM relies on computing the propagation within the optical fibers as the concatenation of small steps, and treating the linear steps (made in the frequency domain) and the nonlinear steps (made in the time domain) separately (Sec. 1.2.3). Here we focused on the nonlinear steps, whose length has to be sufficiently long to take in account the walk-off between the two edge channels in a WDM comb to correctly take into account the cross-nonlinearities. As a general rule, we must ensure that within a single step the edge channels walk past each other by a time smaller than the symbol time, much smaller for a better accuracy. This way, we are correctly discretizing the relative motion of the channels. In the literature this method is called *walk-off method* [79] and is useful when the walk-off is large, i.e., at large chromatic dispersion or channel spacing.

Fig. D.1 sets the idea for a simple 2-channel transmission. In the upper line the transmitted interfering pattern is depicted; in the middle there is a “walking” probe symbol, which walks

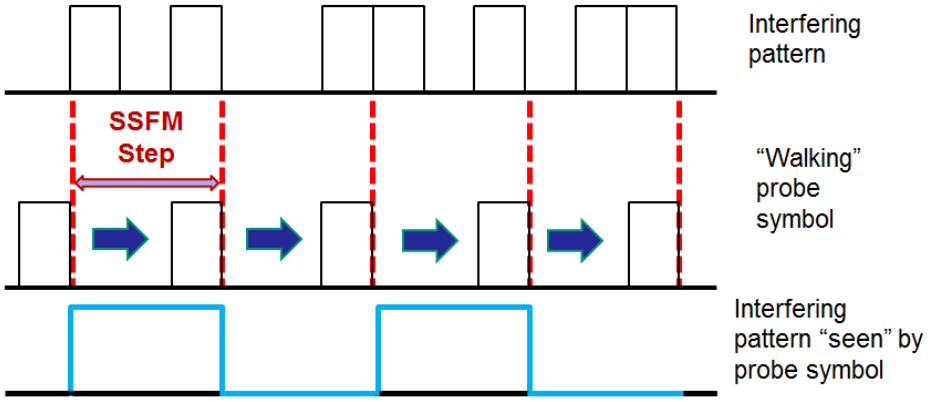


Figure D.1: SSFM step not short enough to correctly take into account the walk-off: interfering patter (top), "walking" probe symbol (center) and interfering patter seen by probe symbol (bottom).

step by step along the fiber. Propagating the probe symbol using a too long SSFM step, in the bottom line the interfering pattern actually felt by the probe symbol appears. The figure shows that with this nonlinear step length the probe symbol walks through 3 interfering symbols at each step. Hence, the impact of the interfering pattern (bottom) on the probe is as if its walk-off were 1/3 of its actual value and consequently, we over-estimate the cross-nonlinearities. We conclude that, to correctly take into account the walk-off, the nonlinear step must be smaller than the walk-off length. To ensure this condition, the nonlinear step length [km] has to be

$$L_{NL-step} \leq \frac{T_o}{D \cdot \Delta\lambda} \tag{D.1}$$

where $T_0\,[ps]$ is the interfering symbol period, $D\,\left[\frac{ps}{nm \cdot km}\right]$ is the transmission fiber dispersion and $\Delta\lambda\,[nm]$ is the channel spacing.

As said before, a wrong choice of the step length can induce a cross-nonlinearities over-estimation for far away channels. To quantify this claim, we simulated a system in which the WDM comb was composed of a 112 Gb/s PDM-QPSK probe channel surrounded by five 10 Gb/s OOK channels on each side, transmitted over a 20 spans of 100km-long SMF-based DM link. The OOK channels were 50 GHz spaced but there was a variable guard band between the nearest OOKs and the PDM-QPSK channel, as reported in Fig. D.2 (top). We increased the guard band by moving the two sets of five OOK channels away from the PDM-QPSK

channel and measured its Q-factor. The PDM-QPSK power was set to -2 dBm, while OOK power was set to -3 dBm.

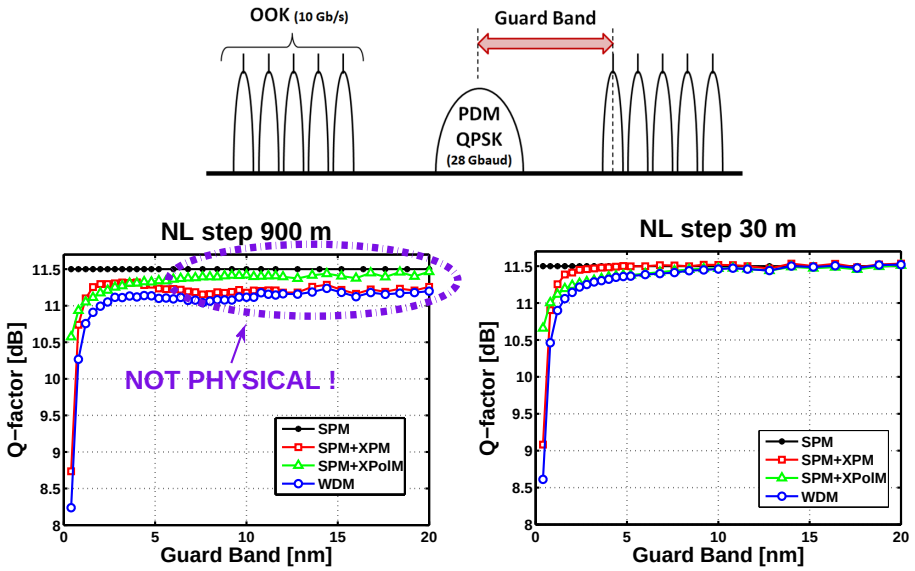


Figure D.2: Spectrum of a WDM comb formed by a PDM-QPSK probe channel surrounded by five OOK channels on each side. PDM-QPSK Q-factor vs Guard Band using 900 meters-long SSFM step (bottom left) or 30 meters-long SSFM step (bottom right) .

In Fig. D.2 (bottom-left) the Q-factor vs guard band for a step length of 900 meters is reported. We note that the 'WDM' curve (circle) and the 'SPM+XPM' one increase with increasing guard band, but they do not reach the single channel Q-factor value. This behavior is physically impossible because the XPM-induced penalty vanishes with increasing walk-off, in agreement with the walk-off window theory[52] and with results of Sec. 4.1 and Sec. 4.2. Looking at the results for a step length of 30 meters (Fig. D.2 bottom-right) we observe that, as logical, all three curves converge to single channel Q-factor value at increasing band guard.

D.2 Pattern length

Another critical parameter in evaluating cross-channel nonlinear impairments is transmission pattern length. In this section we show the two rules to correctly choose the pattern length in

NDM and DM links.

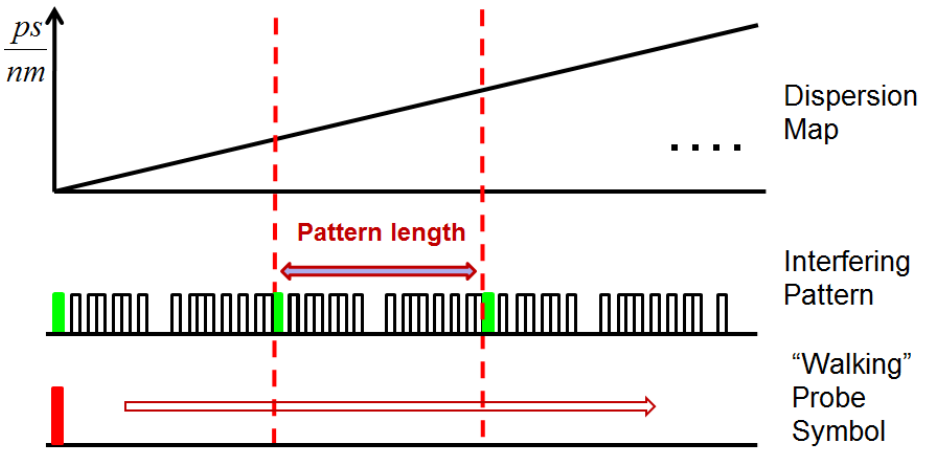


Figure D.3: NDM link: cumulated dispersion (top), interfering pattern (middle) and a "walking" probe symbol (bottom). The interfering pattern is periodically repeated due to the circular property of the FFT.

Starting with NDM link, Fig. D.3 reports a scenario in which the pattern length is not sufficiently long. The cumulated dispersion, the interfering pattern and "walking" probe symbol are reported in top-line, middle-line and bottom-line, respectively. Please note that the interfering pattern is periodically repeated due to the circular property of the Fast Fourier Transform (FFT).

The basic idea is to avoid that a probe symbol, along propagation in the fiber, "sees" the same interfering symbols, creating an artificial correlation that changes the results. As we can see in Fig. D.3, the probe symbol propagating along the fiber, walks w.r.t. the pump symbols but, since the pattern is not long enough, until the end of the link the reference symbol sees three times the same interfering symbol (the grey symbol in the middle line). In this way the nonlinear impact of the grey interfering symbol in artificially increased. Hence, for a correct simulation of cross-channel effects the pattern length for a NDM link has to be chosen such as

$$N_{symb} \geq \left\lceil \frac{D \cdot \Delta\lambda \cdot L_{span} \cdot N_{span}}{T_0} \right\rceil \tag{D.2}$$

where N_{symb} is the sequence length expressed in number of symbols, $D\left[\frac{ps}{nm \cdot km}\right]$ is the trans-

mission fiber dispersion, $\Delta\lambda\,[nm]$ is the channel spacing, $L_{span}\,[km]$ is the span length, N_{span} is the number of spans, $T_0\,[ps]$ is the interfering symbol period and $\lceil\cdot\rceil$ is the ceiling function. Note that for sake of simplicity dispersion slope is not taken into account and we assume the same number of symbols for all channels. Generally speaking this is true only when all channels have the same baud rate, since the FFT size must be the same for all channel. Hence, the pattern length is proportional to the channel baud rate.

To explain the pitfall coming from an insufficiently long pattern, we report the Q-factor vs channel spacing in a homogeneous 3-channel PDM-QPSK system transmitted over a 20x100km SMF-based NDM link, i.e., the same numerical setup as in Sec. 4.2.2. In Fig. D.4 the Q-factor vs channel spacing obtained by simulating a 4096-long sequence (circle) or a 1024-long sequence (triangle) is reported. To save on computation time, only one random seed is depicted.

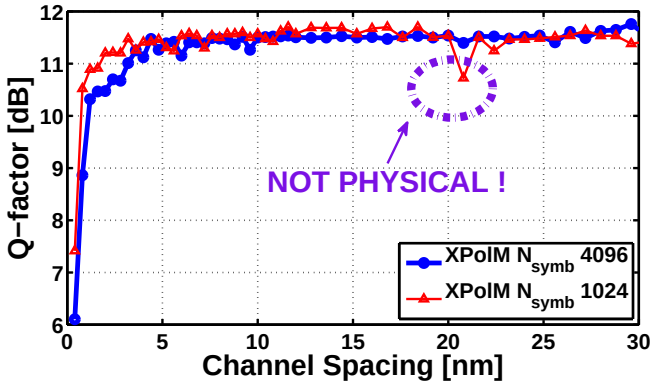


Figure D.4: PDM-QPSK Q-factor vs Channel Spacing in NDM link, simulating 4096-long sequence (circle) or 1024-long sequence (triangle).

The triangle curve (1024-long sequences) indicates a big «non-physical» Q-factor dip around 20.8 nm, demonstrating that the sequences are not sufficiently long. Indeed at 20.8 nm, 1024 symbols "walk off" after almost 100 km ($L = \frac{N_{symb}\cdot T_0}{D\cdot\Delta\lambda} = \frac{1024\cdot 35.7}{17\cdot 20.8} = 103.4\,[km]$) in a SMF fiber, thus artificially creating a resonance span-by-span as if the residual dispersion would be almost zero. In other words, the same interfering symbols impact several times (almost at each span) within the nonlinear length, thus artificially increasing its cross-nonlinearly induced-penalty.

Although to satisfy eq. (D.2) we would need at least 28571 symbols (at 30 nm), by using 4096-long sequences we obtained the circle curve in which the artificial dip disappeared.

Hence, in the cases in which the fulfillment of eq. (D.2) requires exceedingly long sequences, a trade-off between numerical accuracy and time consumption is necessary. It is worth to note that the Q-factor difference at low channel spacing is due to the different interfering patterns, coming form the difference in their length.

In NDM link we only take into account the «fast» walk-off, i.e., the walk-off induced by the fiber dispersion; on the contrary in DM links we have also to take into account the «slow» walk-off due to RDPS, in addition to the «fast» walk-off of a single span. For example in a DM link with zero RDPS at the end of each span the in-line compensating fiber realigns the probe and interfering sequences. Consequently only the number of symbols that walk off in a single span must be taken into account, namely ($\frac{\Delta\lambda \cdot D \cdot L_{span}}{T_0}$). In presence of RDPS this realignment is not perfect and the number of «mismatched» symbols ($\frac{\Delta\lambda \cdot RDPS}{T_0}$) must also be taken into account. Hence, in the DM case the sequence length N_{symb} expressed in number of symbols has to be

$$N_{symb} \geq \left\lceil \frac{\Delta\lambda \cdot (D \cdot L_{span} + RDPS \cdot N_{span})}{T_0} \right\rceil \tag{D.3}$$

where $D\left[\frac{ps}{nm \cdot km}\right]$ is the transmission fiber dispersion, $\Delta\lambda\,[nm]$ the channel spacing, $L_{span}\,[km]$ the span length, N_{span} the number of spans, $RDPS$ the residual dispersion per span, $T_0\,[ps]$ the interfering symbol period and $\lceil\cdot\rceil$ the ceiling function.

D.3 Type of sequences

In this last part we analyze a possible pitfall coming from a wrong choice of sequence type or an insufficient decorrelation among the interfering channel patterns. We focused our attention on two types of sequences: random and De Brujin. A k-ary De Bruijn sequence B(k, n) of order n is a cyclic sequence from a given alphabet A with size k for which every possible sub sequence of length n in A appears as a sequence of consecutive characters exactly once. Each B(k, n) has length k^n.

An example of De Brujin and Random sequences obtained for 7 different generating seeds is shown in Fig. D.5. This figure shows that the first part and the last part of the De Brujin sequences are the same for all seeds. This behavior is generally very rare in purely random sequences (i.e., those where each symbol is independently and uniformly randomly chosen from the alphabet) and can induce a numerical error in analysis in which the pattern randomness is important.

For instance, using the De Bruijn sequences in the setup utilized in Sec. 5.1 to measure

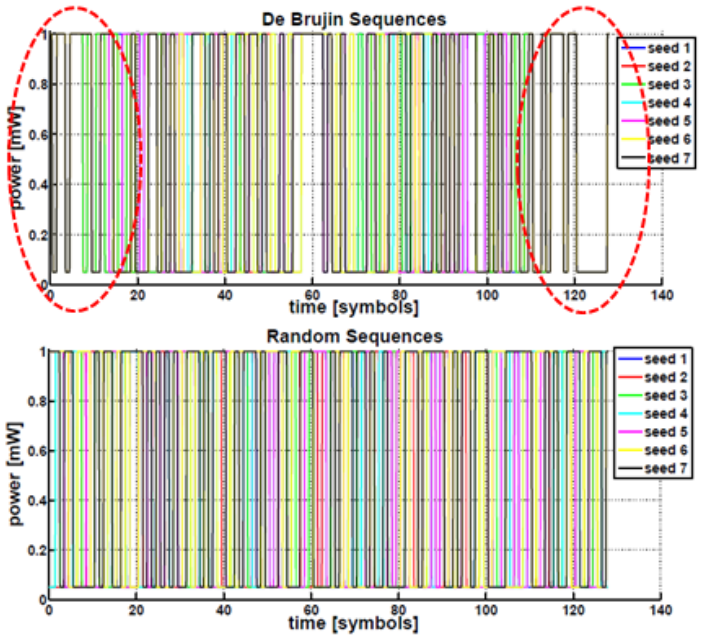


Figure D.5: Comparison between De Brujin (top) and Random (bottom) sequences for 7 random seeds generator.

the Q-factor vs. the suppressor delay, we note two artificial dips around a delay of 5 symbols, as shown in Fig. D.6 (triangle). Using random sequences for probe and pump signal we obtained the circle curve, in which the unphysical dips disappear. The dips could disappears also by properly shifting the two De Bruijn sequences. However, with short sequences and many channels, a good decorrelation cannot be ensured a priori among all the channels. For instance, in Fig. D.5 we note that the De Bruijn sequences coincide for a total of almost 20 symbols over a 128-long sequence, hence even by shifting them the probability of a realignment along the link due to the walk-off can be considerable. This artifacts can also appear with random sequences but with a much lower probability. For this reason, in such WDM setups it is best to use purely random sequences.

Since the De Brujin sequences are largely correlated the cross-nonlinearities can be largely over-estimated, as shown in Fig. D.7, which reports the PDM-QPSK Q-factor for a SMF-based DM link with RDPS = 30 ps/nm, analyzed in Sec. 4.2.4. This figure confirms that using De Brujin without time decorrelation can induce an over-estimation by of almost 2 dB

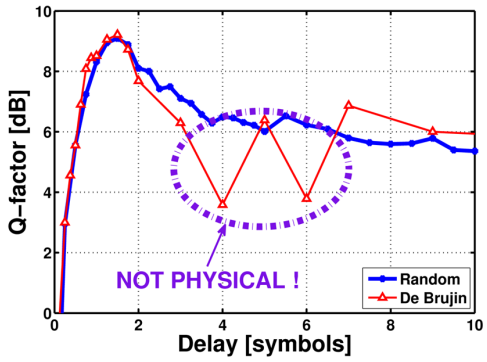


Figure D.6: PDM-QPSK Q-factor in function of the XPM suppressor delay (Fig. 5.2) using random (circle) or De Bruijn sequences (triangle).

for aRZ and NRZ pulse shapes and 0.7 dB for iRZ-PDM-QPSK (which unfortunately was published in [80]).

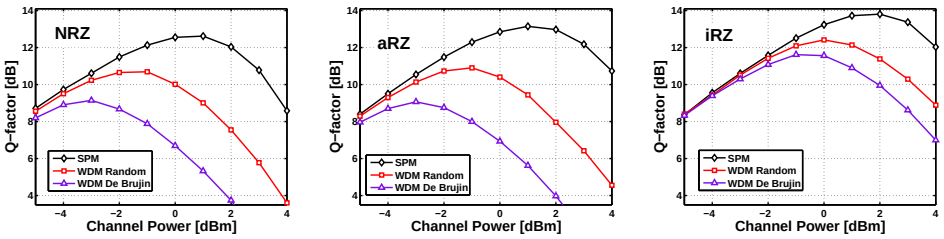


Figure D.7: PDM-QPSK Q-factor in function of the XPM suppressor delay (Ch. 5) using random (squares) or De Bruijn sequences (triangles).

Bibliography

[1] M. Karlsson and H. Sunnerud, "Effects of nonlinearities on PMD-induced system impairments," *J. Lightw. Technol.*, vol. 24, no. 11, pp. 4127 –4137, Nov. 2006.

[2] M. Winter, C.-A. Bunge, D. Setti, and K. Petermann, "A statistical treatment of cross-polarization modulation in DWDM systems," *J. Lightw. Technol.*, vol. 27, no. 17, pp. 3739 –3751, Sep. 2009.

[3] P. Serena, M. Bertolini, and A. Vannucci, "Optilux Toolbox," 2009. [Online]. Available: www.optilux.sourceforge.net

[4] G. P. Agrawal, *Nonlinear Fiber Optics*. 3rd ed. Academic Press, 2001.

[5] G. Keiser, *Optical Fiber Communications*. 3rd ed. McGraw-Hill, 2000.

[6] M. Eiselt, "Limits on WDM systems due to four-wave mixing: a statistical approach," *J. Lightw. Technol.*, vol. 17, no. 11, pp. 2261–2267, Nov. 1999.

[7] S. Kumar, "Analysis of degenerate four-wave-mixing noise in return-to-zero optical transmission systems including walk-off," *J. Lightw. Technol.*, vol. 23, no. 1, pp. 310–320, Jan. 2005.

[8] D. Marcuse, C. Manyuk, and P. Wai, "Application of the Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence," *J. Lightw. Technol.*, vol. 15, no. 9, pp. 1735–1746, Sep. 1997.

[9] C. Menyuk and B. Marks, "Interaction of polarization mode dispersion and nonlinearity in optical fiber transmission systems," *J. Lightw. Technol.*, vol. 24, no. 7, pp. 2806 – 2826, Jul. 2006.

[10] A. Galtarossa and M. C. R., *Polarization Mode Dispersion*. Springer Science, 2004.

[11] J. P. Gordon and H. Kogelnik, "PMD fundamentals: polarization mode dispersion in optical fibers," *Proc. of the National Academy of Science (PNAS)*, vol. 97, pp. 4541 –4550, Apr. 2000.

[12] I. Kaminow, "Polarization in optical fibers," *IEEE J. Quantum Electron.*, vol. 17, no. 1, pp. 15–22, Jan. 1981.

[13] M. Karlsson, J. Brentel, and P. Andrekson, "Long-term measurement of PMD and polarization drift in installed fibers," *J. Lightw. Technol.*, vol. 18, no. 7, pp. 941 –951, Jul. 2000.

[14] S. J. Savory, "Digital filters for coherent optical receivers," *Opt. Exp.*, vol. 16, no. 2, pp. 804–817, Jan. 2008.

[15] J. Renaudier, G. Charlet, M. Salsi, O. Bertran Pardo, H. Mardoyan, P. Tran, and S. Bigo, "Linear Fiber Impairments Mitigation of 40-Gbit/s Polarization-Multiplexed QPSK by Digital Processing in a Coherent Receiver," *J. Lightw. Technol.*, vol. 26, no. 1, pp. 36 –42, Jan. 2008.

[16] D. Van Den Borne, T. Duthel, C. R. S. Fludger, E. D. Schmidt, T. Wuth, C. Schulien, E. Gottwald, G. D. Khoe, and H. d. Waardt, "Electrical PMD Compensation in 43-Gb/s POLMUX-NRZ-DQPSK enabled by Coherent Detection and Equalization," Sep. 2007.

[17] E. L. Buckland and R. W. Boyd, "Electrostrictive contribution to the intensity-dependent refractive index of optical fibers," *Opt. Lett.*, vol. 21, no. 15, pp. 1117–1119, Aug. 1996.

[18] S. V. Chernikov and J. R. Taylor, "Measurement of normalization factor of n2 for random polarization in optical fibers," *Opt. Lett.*, vol. 21, no. 19, pp. 1559–1561, Oct. 1996.

[19] R. J. LeVeque and J. Oliger, "Numerical methods based on additive splittings for hyperbolic partial differential equations," *Mathematics of Computation*, vol. 40, no. 162, pp. 469–497, Apr. 1983.

[20] G. Charlet, M. Salsi, H. Mardoyan, P. Tran, J. Renaudier, S. Bigo, M. Astruc, P. Sillard, L. Provost, and F. Cerou, "Transmission of 81 channels at 40Gbit/s over a transpacific-distance erbium-only link, using PDM-BPSK modulation, coherent detection, and a new large effective area fibre." in *Proc. ECOC 2008,* Bruxelles, Belgium, Sep. 2008, paper Th.3.E.3.

[21] P. Winzer and A. Gnauck, "112-Gb/s polarization-multiplexed 16-QAM on a 25-GHz WDM grid," in *Proc. ECOC 2008,* Bruxelles, Belgium, Sep. 2008, paper We7.2.2.

[22] P. J. Winzer and R.-J. Essiambre, "Advanced modulation formats for high-capacity optical transport networks," *J. Lightw. Technol.*, vol. 24, no. 12, pp. 4711 – 4728, Dec. 2006.

[23] G. Charlet, M. Salsi, J. Renaudier, O. Bertran Pardo, H. Mardoyan, and S. Bigo, "Performance comparison of singly-polarised and polarisation-multiplexed coherent transmission at 10Gbauds under linear impairments," *Electronics Letters*, vol. 43, no. 20, pp. 1109 –1111, Sep. 2007.

[24] G. Charlet, J. Renaudier, O. Bertran Pardo, P. Tran, H. Mardoyan, and S. Bigo, "Performance comparison of singly-polarized and polarization-multiplexed at 10Gbaud under nonlinear impairments," in *Proc. OFC-NFOEC 2008,* San Diego, CA, Feb. 2008, paper OthU8.

[25] A. Bononi, P. Serena, and N. Rossi, "Nonlinear signal-noise interactions in dispersion-managed links with various modulation format," *Opt. Fiber Technol.*, vol. 16, no. 2, pp. 73–85, Mar. 2010.

[26] R. Linke and A. Gnauck, "High-capacity coherent lightwave systems," *J. Lightw. Technol.*, vol. 6, no. 11, pp. 1750 –1769, Nov. 1988.

[27] S. Ryu, S. Yamamoto, H. Taga, N. Edagawa, Y. Yoshida, and H. Wakabayashi, "Long-haul coherent optical fiber communication systems using optical amplifiers," *J. Lightw. Technol.*, vol. 9, no. 2, pp. 251 –260, Feb. 1991.

[28] L. Kazovsky, "Phase- and polarization-diversity coherent optical techniques," *J. Lightw. Technol.*, vol. 7, no. 2, pp. 279 –292, Feb. 1989.

[29] D. Malyon, T. Hodgkinson, D. Smith, R. Booth, and B. Daymond-John, "PSK homodyne receiver sensitivity measurements at 1.5 ÎŒm," *Electronics Letters*, vol. 19, no. 4, pp. 144 –146, Feb. 1983.

[30] M. Taylor, "Coherent detection method using DSP for demodulation of signal and subsequent equalization of propagation impairments," *IEEE Photon. Technol. Lett.*, vol. 16, no. 2, pp. 674 –676, Feb. 2004.

[31] ——, "Algorithms for coherent detection," in *Proc. OFC-NFOEC 2010,* San Diego, CA, Mar. 2010, paper OThL4.

[32] R. Linke, B. Kasper, N. Olsson, and R. Alferness, "Coherent lightwave transmission over 150 km fibre lengths at 400 Mbit/s and 1 Gbit/s data rates using phase modulation," *Electronics Letters*, vol. 22, no. 1, pp. 30 –31, Jan. 1986.

[33] E. Basch and T. Brown, "Introduction to coherent optical fiber transmission," *Communications Magazine, IEEE*, vol. 23, no. 5, pp. 23 –30, May 1985.

[34] G. Nicholson, "Probability of error for optical heterodyne DPSK system with quantum phase noise," *Electronics Letters*, vol. 20, no. 24, pp. 1005 –1007, Nov. 1984.

[35] T. Okoshi and Y. Cheng, "Four-port homodyne receiver for optical fibre communications comprising phase and polarisation diversities," *Electronics Letters*, vol. 23, no. 8, pp. 377 –378, Apr. 1987.

[36] S. Savory, "Digital Signal Processing Options in Long Haul Transmission," in *Proc. OFC-NFOEC 2008,* San Diego, CA, Feb. 2008, paper OtuO3.

[37] S. Savory, A. Stewart, S. Wood, G. Gavioli, M. Taylor, R. Killey, and P. Bayvel., "Digital equalization of 40 Gbit/s per wavelength transmission over 2.480 km of standard fiber without optical dispersion compensation," in *Proc. ECOC 2006,* Cannes, France, Sep. 2008, paper Th2.5.5.

[38] A. J. Viterbi and A. M. Viterbi, "Nonlinear estimation of PSK-modulated carrier phase with application to burst digital transmission," *IEEE Trans. Inf. Theory*, vol. 29, no. 4, pp. 543 – 551, Jul. 1983.

[39] K. Kikuchi, "Coherent transmission systems," in *Proc. ECOC 2008,* Bruxelles, Belgium, Sep. 2008, paper Th.2.A.1.

[40] G. Bosco, V. Curri, A. Carena, P. Poggiolini, and F. Forghieri, "On the performance of nyquist-WDM terabit superchannels based on PM-BPSK, PM-QPSK, PM-8QAM or PM-16QAM subcarriers," *J. Lightw. Technol.*, vol. 29, no. 1, pp. 53 –61, Jan. 2011.

[41] A. Li, A. Al Amin, X. Chen, and W. Shieh, "Reception of mode and polarization multiplexed 107-Gb/s CO-OFDM signal over a two-mode fiber," in *Proc. OFC-NFOEC 2011,* Los Angeles, CA, Mar. 2011, paper PDPB8.

[42] M. Salsi, C. Koebele, D. Sperti, P. Tran, P. Brindel, H. Mardoyan, S. Bigo, A. Boutin, F. Verluise, P. Sillard, M. Astruc, L. Provost, F. Cerou, and G. Charlet, "Transmission

at 2x100Gb/s, over two modes of 40km-long prototype few-mode fiber, using LCOS-based mode multiplexer and demultiplexer," in *Proc. OFC-NFOEC 2011,* Los Angeles, CA, Mar. 2011, paper PDPB9.

[43] R. Ryf, S. Randel, A. Gnauck, C. Bolle, R. Essiambre, P. Winzer, D. Peckham, A. McCurdy, and R. Lingle, "Space-division multiplexing over 10 km of three-mode fiber using coherent 6x6 MIMO processing," in *Proc. OFC-NFOEC 2011,* Los Angeles, CA, Mar. 2011, paper PDPB10.

[44] N. Hanzawa, K. Saitoh, T. Sakamoto, T. Matsui, S. Tomita, and M. Koshiba, "Demonstration of mode-division multiplexing transmission over 10 km two-mode fiber with mode coupler," in *Proc. OFC-NFOEC 2011,* Los Angeles, CA, Mar. 2011, paper OWA4.

[45] J. Carpenter and T. Wilkinson, "Holographic mode-group division multiplexing," in *Proc. OFC-NFOEC 2011,* Los Angeles, CA, Mar. 2011, paper OthN3.

[46] A. Bononi, A. Vannucci, A. Orlandini, E. Corbel, S. Lanne, and S. Bigo, "Degree of polarization degradation due to cross-phase modulation and its impact on polarization-mode dispersion compensators," *J. Lightw. Technol.*, vol. 21, no. 9, pp. 1903 – 1913, Sep. 2003.

[47] A. Bononi, P. Serena, N. Rossi, and D. Sperti, "Which is the dominant nonlinearity in long-haul PDM-QPSK coherent transmissions?" in *Proc. ECOC 2010,* Torino, Italy, Sep. 2010, paper Th.10.E.1.

[48] A. Bononi, P. Serena, and A. Orlandini, "A unified design framework for single-channel dispersion-managed terrestrial systems," *J. Lightw. Technol.*, vol. 26, no. 22, pp. 3617 –3631, Nov. 2008.

[49] G. Charlet, "Coherent detection associated with digital signal processing for fiber optics communication," *C. R. Physique 9 (2008)*, pp. 1012 – 1030, Dec. 2008.

[50] L. Li, Z. Tao, S. Oda, T. Tanimura, M. Yuki, T. Hoshida, and J. Rasmussen, "Adaptive optimization for digital carrier phase estimation in optical coherent receivers," in *IEEE/LEOS Photonics Society Summer Topical Meeting, Newport Beach, CA, USA*, Jul. 2008, paper TuC3.3.

[51] Z. Tao, L. Li, L. Liu, W. Yan, H. Nakashima, T. Tanimura, S. Oda, T. Hoshida, and J. Rasmussen, "Improvements to digital carrier phase recovery algorithm for high-performance optical coherent receivers," *IEEE J. Sel. Topics Quantum Electron.*, vol. 16, no. 5, pp. 1201 –1209, Sep.-Oct. 2010.

[52] A. Bononi, M. Bertolini, P. Serena, and G. Bellotti, "Cross-phase modulation induced by OOK channels on higher-rate DQPSK and coherent QPSK channels," *J. Lightw. Technol.*, vol. 27, no. 18, pp. 3974 –3983, Sep. 2009.

[53] S. Bigo, O. Bertran-Pardo, J. Renaudier, and G. Charlet, "Nonlinear impairments in coherent communication systems," in *IEEE/LEOS Photonics Society Summer Topical Meeting , Playa del Carmen, Mexico*, Jul. 2010, paper WC3.1.

[54] O. Bertran-Pardo, J. Renaudier, G. Charlet, H. Mardoyan, P. Tran, and S. Bigo, "Nonlinearity limitations when mixing 40-Gb/s coherent PDM-QPSK channels with preexisting 10-Gb/s NRZ channels," *IEEE Photon. Technol. Lett.*, vol. 20, no. 15, pp. 1314 – 1316, Aug. 2008.

[55] C. Xie, "Interchannel nonlinearities in coherent polarization-division-multiplexed quadrature-phase-shift-keying systems," *IEEE Photon. Technol. Lett.*, vol. 21, no. 5, pp. 274 –276, Mar. 2009.

[56] J. Renaudier, O. Bertran-Pardo, G. Charlet, M. Salsi, M. Bertolini, P. Tran, H. Mardoyan, and S. Bigo, "On the required number of WDM channels when assessing performance of 100Gb/s coherent PDM-QPSK overlaying legacy systems," in *Proc. ECOC 2009,* Vienna, Austria, Sep. 2009, paper Tu3.4.5.

[57] O. Bertran-Pardo, J. Renaudier, M. Charlet, Salsi, M. Bertolini, H. Mardoyan, P. Tran, C. Koebele, and S. Bigo, "PDM-QPSK: On the system benefits arising from temporally interleaving polarization tributaries at 100 Gb/s," *Opt. Exp.*, vol. 17, no. 22, pp. 19 902 – 19 907, Mar. 2009.

[58] A. Bononi, N. Rossi, and P. Serena, "Transmission limitations due to fiber nonlinearity," in *Proc. OFC-NFOEC 2011,* Los Angeles, CA, Mar. 2011, paper OWO7.

[59] C. Xie, "Inter-channel nonlinearities in hybrid OOK and coherent PDM-QPSK transmission systems with dispersion management," in *IEEE/LEOS Photonics Society Summer Topical Meeting , Playa del Carmen, Mexico*, Jul. 2010, paper TuA.3.2.

[60] P. Serena, N. Rossi, and A. Bononi, "Nonlinear penalty reduction induced by PMD in 112 Gbit/s WDM PDM-QPSK coherent systems," in *Proc. ECOC 2009,* Vienna, Austria, Sep. 2009, paper 10.4.3.

[61] D. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Commun.*, vol. 28, no. 11, pp. 1867–1875, Nov. 1980.

[62] J. Challis, "A procedure for determining rigid body transformation parameters," *J. Biomechanics*, vol. 28, no. 6, pp. 733 –737, Jun. 1995.

[63] S. Savory, "Digital coherent optical receivers: Algorithms and subsystems," *IEEE J. Sel. Topics Quantum Electron.*, vol. 16, no. 5, pp. 1164 –1179, Sep.-Oct. 2010.

[64] E. Tipsuwannakul, M. N. Chughtai, M. Forzati, J. Mårtensson, P. Andrekson, and M. Karlsson, "Influence of self- and cross-phase modulation on 40 Gbaud dual polarization DQPSK/D8PSK signals in 10 Gbit/s OOK WDM systems," *Opt. Exp.*, vol. 18, no. 23, pp. 24 178–24 188, Nov. 2010.

[65] G. Charlet, J. Renaudier, O. Bertran Pardo, P. Tran, H. Mardoyan, and S. Bigo, "Performance comparison of singly-polarized and polarization-multiplexed at 10Gbaud under nonlinear impairments," in *Proc. OFC-NFOEC 2008,* San Diego, CA, Feb. 2008, paper OThU8.

[66] C. Xia and D. Van Den Borne, "Impact of the channel count on the nonlinear tolerance in coherently-detected POLMUX-QPSK modulation," in *Proc. OFC-NFOEC 2011,* Los Angeles, CA, Mar. 2011, paper OWO1.

[67] M. Yan, Z. Tao, H. Zhang, W. Yan, T. Hoshida, and J. Rasmussen, "Adaptive blind equalization for coherent optical BPSK system," in *Proc. ECOC 2010,* Torino, Italy, Sep. 2010, paper Th.9.A.4.

[68] G. Bellotti and S. Bigo, "Cross-phase modulation suppressor for multispan dispersion-managed WDM transmissions," *IEEE Photon. Technol. Lett.*, vol. 12, no. 6, pp. 726 –728, Jun. 2000.

[69] X. Wei, X. Liu, C. Xie, and L. F. Mollenauer, "Reduction of collision induced timing jitter in dense wavelength-division multiplexing by the use of periodic-group-delay dispersion compensators," *Opt. Lett.*, vol. 12, no. 28, pp. 983 – 985, Mar. 2003.

[70] C. Xie, "Suppression of inter-channel nonlinearities in WDM coherent PDM-QPSK systems using periodic-group-delay dispersion compensators," in *Proc. ECOC 2009,* Vienna, Austria, Sep. 2009, paper P4.08.

[71] D. Sperti, P. Serena, and A. Bononi, "Optical Solutions to Improve PDM-QPSK Resilience Against Cross-Channel Nonlinearities: A Comparison," *IEEE Photon. Technol. Lett.*, vol. 23, no. 11, pp. 667 –669, Jun. 2011.

[72] M. Salsi, C. Koebele, D. Sperti, P. Tran, H. Mardoyan, P. Brindel, S. Bigo, A. Boutin, P. Sillard, M. Astruc, L. Provost, and G. Charlet, "Mode Division Multiplexing of 2 x 100Gb/s Channels using an LCOS based Spatial Modulator," *J. Lightw. Technol.*, p. Submitted, 2011.

[73] D. Sperti, M. Salsi, C. Koebele, P. Tran, H. Mardoyan, S. Bigo, A. Boutin, P. Sillard, and G. Charlet, "Experimental investigation of modal crosstalk using LCOS-based spatial light modulator for mode conversion," in *Proc. ECOC 2011,* Genève, Switzerland, Sep. 2011, paper Th.12.B.2.

[74] C. Koebele, M. Salsi, D. Sperti, P. Tran, P. Brindel, H. Mardoyan, S. Bigo, A. Boutin, P. Sillard, F. Cerou, and G. Charlet, "Two-mode transmission with digital inter-modal cross-talk mitigation," in *Proc. ECOC 2011,* Genève, Switzerland, Sep. 2011, paper Tu.5.B.4.

[75] C. Koebele, M. Salsi, D. Sperti, P. Tran, P. Brindel, H. Mardoyan, S. Bigo, A. Boutin, F. Verluise, P. Sillard, M. Astruc, L. Provost, F. Cerou, and G. Charlet, "Two mode transmission at 2x100Gb/s, over 40km-long prototype few-mode fiber, using LCOS-based programmable mode multiplexer and demultiplexer," *Opt. Express*, vol. 19, no. 17, pp. 16 593–16 600, Aug. 2011.

[76] F. Yaman, N. Bai, Y. K. Huang, M. F. Huang, B. Zhu, T. Wang, and L. G., "10x112Gb/s PDM-QPSK transmission over 5032 km in few-mode fibers," *Opt. Exp.*, vol. 18, no. 23, pp. 24 178–24 188, Nov. 2010.

[77] S. Huard, *Polarisation de la Lumiere*. Masson, Paris, 1994.

[78] C.-T. Chen, *Linear System Theory and Design*. Holt-Saunders Int. Edition, 1984.

[79] O. Sinkin, R. Holzlohner, J. Zweck, and C. Menyuk, "Optimization of the split-step Fourier method in modeling optical-fiber communications systems," *J. Lightw. Technol.*, vol. 21, no. 1, pp. 61 –68, Jan. 2003.

[80] D. Sperti, P. Serena, and A. Bononi, "A comparison of different options to improve PDM-QPSK resilience against cross-channel nonlinearities," in *Proc. ECOC 2010,* Torino, Italy, Sep. 2010, paper Th.9.A.1.