**UNIVERSITÀ DEGLI STUDI DI PARMA**

*Dottorato di Ricerca in Tecnologie dell'Informazione*

*XXVII Ciclo*

# A STUDY OF NONLINEAR THRESHOLD AND OUTAGE PROBABILITY IN COHERENT OPTICAL SYSTEMS

Coordinatore:

*Chiar.mo Prof. Marco Locatelli*

Tutor:

*Chiar.mo Prof. Alberto Bononi*

Dottorando: *Nicola Rossi*

Gennaio 2015

*A Claudia*

*Alla mia Famiglia*

# Contents

# Introduction

The evolution in the last decades of long haul optical transmission systems followed the need to cope with an ever increasing demand for data capacity. The advent of coherent detection, firstly proposed at the end of the 1980s [1], allowed to match this demand by increasing the system spectral efficiency and consequently its capacity. Besides, the access to the full information of the optical field combined with the development of high speed digital electronics enabled digital signal processing (DSP) at the receiver. Among the advantages coming from DSP we find the use of polarization division multiplexing (PDM) to double system capacity, the use of multilevel modulation formats such as N-phase shift keying (PSK) and N-quadrature amplitude modulation (QAM), and the possibility to compensate in the electrical domain for linear impairments as chromatic dispersion (CD), polarization mode dispersion (PMD) and optical filtering [2]. In particular, CD and PMD compensation at the receiver side opened the possibility to deploy dispersion uncompensated (DU) links as opposed to legacy dispersion managed (DM) ones. One major question for such new coherent systems is about their reach when using both wavelength division multiplexing (WDM) and PDM. With PDM, cross-nonlinearities make neighboring channels interact depending on both their power and their state of polarization, through the well known scalar cross-phase modulation (XPM) and the recently addressed cross-polarization modulation (XPolM). These nonlinear distortions are sensitive to channel spacing in the WDM comb, hence are enhanced in the emerging flexible grid transmissions that suppress channel guardbands (Nyquist-WDM systems) [3].

With the investigations presented in this thesis we aim at shedding light on multiple aspects of modern coherent optical transmissions. The first one is how XPM and XPolM relate to single channel nonlinear distortion, i.e., self-phase modulation (SPM) effect, as the granularity, i.e., the signaling rate, of the system changes. The second one is to determine the best granularity, i.e., which per-channel symbol rate minimizes the bit error rate (BER) at fixed distance, or alternatively maximizes the largest transparent transmission distance (i.e., the reach) at fixed BER. The third one is how the presence of polarization dependent losses (PDL) affects the BER of coherent systems, especially in presence of fiber nonlinearity whose interplay with PDL was highlighted in recent studies. To address these issues we provide a systematic simulation study of system performance, by exploiting the possibility in the Optilux software [4] to selectively activate SPM, XPM or XPolM in the nonlinear step of the Split Step Fourier algorithm (SSFA) that simulates the field propagation along the fiber.

The thesis is organized as follows.

In the first chapter we review the main impairments in long haul optical transmissions, by splitting them into two categories, linear and nonlinear. Among linear impairments we find fiber attenuation, chromatic dispersion, polarization mode dispersion and polarization dependent loss. Among nonlinear impairments we find SPM, XPM and XPolM. We do not include in our discussion other nonlinear impairments such as four wave mixing, nonlinear phase noise and non elastic scattering (stimulated Raman and Brillouin scattering).

In the second chapter we investigate the optimal granularity for WDM long haul optical systems in both DM and DU links. We analyze, for the most popular modulation formats such as PDM-binary PSK, PDM-quaternary PSK (QPSK), polarization switched (PS)-QPSK and PDM-16QAM, the power margin at target BER and fixed distance, when varying the system symbol rate at fixed bandwidth efficiency. We also provide a new simple formula to predict the system reach according to side data obtained in the process of estimating the nonlinear threshold. In the last part of the chapter we provide reach simulations to verify the accuracy of the new reach formula.

In Ch. 3 we show that, differently from what expected by intuition, the nonlinear interference (NLI) noise can, in some particular system arrangements, take a non-circular distribution in DU networks. This study was performed during my 6-month Internship at Alcatel-Lucent Bell-Labs France. Since each channel travels over different lightpaths, the large amount of group velocity dispersion that can be cumulated by neighboring channels in their upstream propagation can induce a non-circular NLI during downstream propagation.

In Ch. 4 we address the impact of PDL on coherently detected systems, by concentrating on the PDM-QPSK modulation format. We study the interaction of PDL with the main fiber nonlinearity manifestations such as SPM, XPM, XPolM by investigating the average Q-factor performance for pathological cases of PDL, and by selectively activating SPM, XPM or XPolM in the SSFA nonlinear step. We analyze i) a homogeneous (PDM-QPSK) setup on a DM link; ii) a hybrid setup (PDM-QPSK – On Off Keying) on a DM link; iii) a homogeneous (PDM-QPSK) setup on a DU link. We furthermore show, for each setup i)–iii), the effect of PDL-nonlinearity interplay on the Q-factor distribution, whose variance is increased, thus resulting in an increased outage probability.

In Ch. 5, we present a Stratified Sampling based algorithm to efficiently assess the outage probability of coherent systems in presence of PDL. The proposed method allows significant savings in computational time with respect to the standard Monte Carlo algorithm and does not lose in efficiency in presence of polarization mode dispersion and above all fiber nonlinearity, where available semi-analytical models miss the PDL-nonlinearity interaction contribution to the outage probability.

Finally, in the Appendix we discuss about the correct procedure to estimate the Q-factor distribution. While the estimation of the average performance can be obtained in feasible time, the estimation of the complete Q-factor distribution is generally unfeasible by simulation. This reason usually leads to expedients that, though speeding up the estimation procedure, can introduce significant errors in the estimated outage probability. Moreover, we provide some warnings on the setting of some key numerical parameters in a SSFA based simulation to correctly reproduce fiber nonlinear effects.

# Chapter 1

# Fiber-Optic Coherent Transmissions

This first chapter recaps the main characteristics of a fiber-optic coherent transmission system. Concerning the propagation of an electric field along an optical fiber, the field evolution can be described in absence of polarization effects through the scalar nonlinear Schrödinger equation (NLSE) and in their presence through the coupled nonlinear Schrödinger equation (CNLSE) or its simplified form known as the Manakov equation [5]. According to these equations we can identify the main impairments in long haul optical transmissions. Usually, impairments are classified into linear and nonlinear effects. Among linear effects we find fiber attenuation, chromatic dispersion (CD), polarization mode dispersion (PMD) and polarization dependent loss (PDL), while among nonlinear effects we find self phase modulation (SPM), cross phase modulation (XPM), cross polarization modulation (XPolM) and four wave mixing (FWM). In this chapter we will briefly review the theory behind these major impairments, except for FWM which is normally well suppressed in most links of practical interest.

The distinguishing sign of second-generation coherent transmission is given by the presence of a digital signal processing (DSP) unit in the coherent receiver. DSP provides remarkable resilience against linear impairments as PMD and CD, and also enables the use of polarization division multiplexing (PDM) technique at the transmitter side. In this chapter we will briefly describe the composition of coherent receivers and the PDM transmission technique.

## 1.1 NLSE

The nonlinear Schrödinger equation is the reference model to describe the evolution of an electric field $A(z,t)\,[\sqrt{\mathrm{W}}]$ propagating through an optical fiber. If the signal is launched in single polarization we have the following partial differential equation (PDE) [6]:

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

where $z$ is the distance [m], $t$ is the time [t], $j$ is the imaginary unit, $\alpha$ is the fiber attenuation and $\gamma$ is 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 $f_0$, $\lambda_0$ the central frequency/wavelength of $A(z,t)$, respectively, and $c$ the speed of light.

## 1.2 Linear Impairments

### 1.2.1 Fiber attenuation

An important parameter of the fiber is the loss during propagation of the injected power. The fiber **attenuation** $\alpha\,[\mathrm{km}^{-1}]$ of the traveling light is mainly due material absorption and Rayleigh scattering. Absorption comes from impurities such as the OH ion, while Rayleigh scattering arises from density fluctuations frozen into the fused silica during manufacture.

Assuming all parameters zero except $\alpha$, the NLSE (1.1) becomes:

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

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

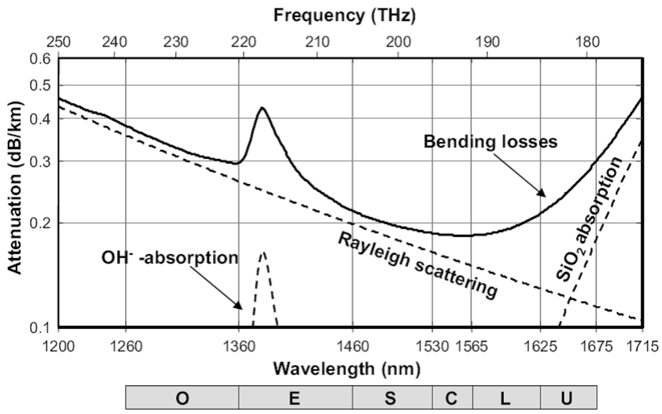


Figure 1.1: Fiber attenuation of a single-mode silica fiber, indicating lowest-loss regions associated with C- and L-transmissions bands. Dashed curves show the contributions resulting from Rayleigh scattering, $OH^-$ and $SiO_2$ absorptions.

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:

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

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

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

Fiber attenuation depends on the wavelength of the light, showing a minimum around 1550 nm, as shown in Fig. 1.1. Most of today transmissions use the C band, with fibers exhibiting typical attenuation of 0.2 [dB/km] around $\lambda = 1550$ nm. To the attenuation one can associate the ***attenuation length*** $L_\mathrm{A} = 1/\alpha$ as a measure of the distance over which nonlinear effects are significant. For a typical fiber having $\alpha = 0.2\,[\text{dB/km}]$ it is $L_\mathrm{A} = 21.715\,[\text{km}]$.

### 1.2.2 Chromatic Dispersion and Group Velocity Dispersion

**Chromatic dispersion** (CD) manifests through the frequency dependence of the refractive index $n(\omega)$ and refers to the medium response to the interaction of an electromagnetic wave with the bound electrons of a dielectric. 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 the fiber.

Mathematically, the effects of fiber dispersion are accounted for by expanding the wave propagation constant $\beta(\omega)$ in a Taylor series about the signal carrier frequency $\omega_0$:

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

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

An important feature of chromatic dispersion is that pulses at different wavelengths propagate at different speeds inside a fiber because of a mismatch in their group velocities. The **group velocity** $v_g$ of the signal along the fiber is accounted for by $\beta_1 = \left(\frac{d\beta}{d\omega}\right)_{\omega=\omega_0} = \frac{1}{v_g(\omega_0)}$ [s/m], hence is a delay per unit length. Two signals centered at wavelengths $\lambda_1$ and $\lambda_2$ generally have different group velocities, hence travel at different speed. The delay per unit length between the two signals is the walk-off parameter $d_{12}$ equal to:

$$d_{12} = \beta_1(\lambda_1) - \beta_1(\lambda_2) = \frac{1}{v_g}(\lambda_1) - \frac{1}{v_g}(\lambda_2)\,.$$

Positive values of $d_{12}$ mean that the channel at wavelength $\lambda_1$ travels slower than the other. The walk-off parameter weights the impact of cross-channel nonlinear effects, such as XPM, XPolM and FWM.

In presence of only $\beta_1$ the NLSE writes as:

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

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

Assuming equal to zero all parameters except $\beta_{1,2,3}$ the NLSE becomes:

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

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

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

whose solution is:

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

Physically, the envelope of an optical pulse moves at the group velocity $v_{\mathrm{g}}$ while the $\beta_2$ parameter represents the dispersion of the group velocity and is responsible for pulse broadening. This phenomenon is known as **group velocity dispersion** (GVD), and $\beta_2$ is the GVD parameter. Typically, $\beta_2$ and $\beta_3$ are expressed as functions of the wavelength through the fiber **dispersion** parameter $D$ and the fiber **dispersion slope** $D^{'}$ defined as:

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

where $\lambda_0 = c/f_0$.

The chromatic dispersion $D$ is usually expressed in [ps/(nm·km)] while the dispersion slope $D^{'}$ in [$\mathrm{ps}^2$/(nm·km)]. By chromatic dispersion the various spectral components of the signal travel at different speeds, hence causing a pulse broadening in the time domain which leads to bit-to-bit overlaps, called Inter-Symbol-Interference (ISI).

To the $\beta_2$ and $\beta_3$ parameters one can associate the ***dispersion length*** $L_{\mathrm{D}} = \beta_2/T_0^2$ and the ***dispersion slope length*** $L_{\mathrm{D}}^{'} = \beta_3/T_0^3$, being $T_0$ the symbol time.

### 1.2.3 Polarization Mode Dispersion

**Polarization mode dispersion** (PMD) [7] arises from radial asymmetries in the fiber. An ideal optical fiber has a perfectly circular core which implies that two orthogonal polarizations of the fundamental mode travel at the same speed. However, realistic conditions such as imperfections in manufacturing or induced thermal and mechanical stresses, break the circular symmetry of the fiber, causing the two polarizations to propagate at different speeds. This property is referred to as modal birefringence, which is stochastic and fluctuates both in time and along fiber length, depending on the characteristics of the fiber and on local condition, such as temperature. As a result of birefringence, there is an exchange of energy between orthogonal polarizations. 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 $\Delta\beta = \beta_s(\omega) - \beta_f(\omega)$ between these propagation constants is the strength of the birefringence, while $\hat{l}(z)$ represents the birefringence orientation. Usually, it is assumed that fibers are characterized by linear birefringence. If the frequency dependent term $\Delta\beta_1$ is zero, there is no pulse distortion, and the overall result of birefringence is just a rotation of the signal state of polarization (SOP) on the Poincaré sphere (App. A). If an input pulse excites both polarizations, the two components travel along the fiber at different speeds because of their different group velocities. The pulse becomes broader during propagation because group velocities change according to random changes in fiber birefringence. This broadening causes ISI and is referred to as polarization mode dispersion. The extent of pulse broadening is related to the differential group delay (DGD) $\Delta\tau$ [ps], i.e., the time delay at the receiver between the two principal modes of the transmission link. For a fiber of length $L$, $\Delta\tau$ is given by:

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

where $v_{gs}$ and $v_{gf}$ are the group velocities of the slow and fast modes, respectively. Because of random changes in birefringence orientation occurring along the fiber, the PMD is characterized by the root-mean-square (rms) value of $\Delta\tau$.

When $\hat{l}$ 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, is known as first-order PMD while in the general case, mode coupling produces PMD at all orders, hence a pulse is not only split in two, but each of the shadow pulses suffers a different amount of linear distortion, including GVD, that differently affects the polarized component of the signal.

### DGD statistics

The statistics of DGD depend on the fiber length. If fact, two regimes can be identified according to fiber length: a short-length regime where birefringence adds linearly with distance and a long-length regime where the birefringence is no longer additive due to random variations in its orientation $\hat{l}(z)$. In today's terrestrial and submarine transmission systems the long-length regime applies and it has been shown that the DGD accumulates as a three-dimensional random-walk, therefore a statistical approach must be adopted. To calculate the probability density of the DGD, the problem can be reduced to that of the probability density of the magnitude of the sum of three-dimensional Gaussian vectors having random orientation and length. The result is the well-known Maxwellian distribution. By this way, the DGD increases on average with the square root of distance. Fiber PMD is often specified by using the PMD coefficient $D_{\text{pmd}}\,[\text{ps}/\sqrt{\text{km}}]$, and the rms value of DGD can be obtained as $\sqrt{E\left[\Delta\tau^2\right]} = D_{\text{pmd}}\sqrt{L_{\text{tot}}}$ where $L_{\text{tot}}\,[\text{km}]$ is the total distance. For Maxwellian distribution the average DGD $E\left[\Delta\tau\right]$ can be obtained by the rms DGD as $E\left[\Delta\tau\right] = \sqrt{\frac{8}{3\pi}E\left[\Delta\tau^2\right]}$.

To emulate the stochastic behavior of real fibers PMD in computer software, many implementations are possible. The choice implemented in the Optilux Toolbox, i.e., the open source software used for all results shown in this thesis, is the following. Once the rms DGD $\sqrt{E\left[\Delta\tau^2\right]}$ we want to emulate is specified, either by knowledge of the PMD coefficient $D_{\text{pmd}}$ or of the average DGD $E\left[\Delta\tau\right]$, we concatenate $N$ PMF sections of length $L/N$, with $L$ the fiber length, all having the same deterministic DGD $\xi = \sqrt{E\left[\Delta\tau^2\right]/N}$. If the eigenvectors of the concatenated sections are uniformly distributed on the Poincaré sphere, the probability density function of the random variable $\Delta\tau$ has maximum value $N\xi$ (all sections aligned) and average value $E[\Delta\tau]$. If the number of sections $N$ is sufficiently large, it can be shown that $\Delta\tau$ has a Maxwellian distribution:

$$p(\Delta\tau, N) = \frac{3(\Delta\tau)^2}{\xi^3 N}\sqrt{\frac{6}{\pi N}}\exp\left(-\frac{3(\Delta\tau)^2}{2N\xi^2}\right). \tag{1.2}$$

When concatenating few sections $N$ of deterministic DGD $\xi$, with randomly oriented eigenvectors, the expression of $p(\Delta\tau, N)$ becomes instead [8]:

$$p(\Delta\tau, N) = \frac{\Delta\tau}{2\xi(N-2)!}\sum_{k=0}^{\lfloor Nm\rfloor}(-1)^k \binom{N}{k}(Nm-k)^{(N-2)} \tag{1.3}$$

where $m = \frac{1}{2}\left(1 - \frac{\Delta\tau}{N\xi}\right)$.

### 1.2.4 Polarization Dependent Loss

**Polarization dependent loss** (PDL) refers to an energy loss that is preferential to one polarization state. PDL mainly comes from optical components such as couplers, isolators, circulators, add-drops and amplifiers, while the contribution of the fibers is typically negligible. Hence we can consider the transmission link as a sequence of birefringent elements, i.e., the fibers, interleaved with lumped PDL elements. Differently from birefringence, which causes an exchange of energy between orthogonal polarizations, PDL imposes an energy loss to a specific polarization state, while the energy of the orthogonal polarization is undiminished. This is the behavior of a partial polarizer. PDL is defined [9] in decibels as

$$\rho_{\mathrm{dB}} = 10\log_{10}\left(\frac{T_{\max}}{T_{\min}}\right) \tag{1.4}$$

where $T_{\max}$ and $T_{\min}$ are the maximum and minimum transmission intensity through the system. The relation between the optical field $\overrightarrow{Z}$ at the output of a generic PDL element and the field $\overrightarrow{A}$ at its input is expressed as:

$$\begin{bmatrix} Z_x \\ Z_y \end{bmatrix} = \underline{\underline{B}}^{\dagger}\begin{bmatrix} 1 & 0 \\ 0 & \exp(-\alpha) \end{bmatrix}\underline{\underline{B}}\begin{bmatrix} A_x \\ A_y \end{bmatrix} \tag{1.5}$$

where $\overrightarrow{A} = [A_x, A_y]^{\mathrm{T}}$ is the modulated input signal, subscripts $x$ and $y$ indicate the two polarizations, $j$ is the imaginary unit and $\alpha = \rho_{\mathrm{dB}}/(20\log_{10}\mathrm{e})$, with

$$\underline{\underline{B}} = \underline{\underline{U_3}}(-\theta)\underline{\underline{U_2}}(\varepsilon) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\varepsilon & j\sin\varepsilon \\ j\sin\varepsilon & \cos\varepsilon \end{bmatrix} \tag{1.6}$$

where $(\theta, \varepsilon)$ are azimuth and ellipticity of the PDL element eigenvector, which can point anywhere on the Poincaré sphere. In eq. (1.5) the average total power after PDL is decreased and can be recovered by an amplifier with gain $G = 2/(1+\exp(-2\alpha))$. The cascade of PDL and afore-mentioned amplifier is represented by a single PDL element giving output optical field

$$\begin{bmatrix} Z_x \\ Z_y \end{bmatrix} = \underline{\underline{B}}^{\dagger} \begin{bmatrix} \sqrt{1+\Gamma} & 0 \\ 0 & \sqrt{1-\Gamma} \end{bmatrix} \underline{\underline{B}} \begin{bmatrix} A_x \\ A_y \end{bmatrix} \tag{1.7}$$

where $\Gamma = \tanh(\alpha)$ is the normalized PDL coefficient such that $\rho_{\mathrm{dB}} = 10\log_{10}\left(\frac{1+\Gamma}{1-\Gamma}\right)$.

Since PDL has a stochastic nature as the DGD, in an optical link with $N$ elements of PDL $\rho_i$, $i = 1,\ldots,N$, the effective PDL $\rho$ becomes a random variable, with the same statistics shown in subsec. 1.2.3 for DGD when $\rho_i = \xi$, $i = 1,\ldots,N$.

### Effective PDL in presence of PMD

When both PDL and PMD are present along the link, their interaction generates several peculiar effects [10]. However, when considering a concatenation of separated PMD and PDL elements, the transmission matrix $\underline{\underline{T}}(\omega)$ of this concatenation can be decomposed as

$$\underline{\underline{T}}(\omega) = \underline{\underline{L}}(\omega)\underline{\underline{P}}(\omega)\underline{\underline{D}}(\omega)$$

where $\underline{\underline{P}}$ represents a partial polarizer and $\underline{\underline{D}}$ and $\underline{\underline{L}}$ are unitary matrices. The partial polarizer is a Hermitian matrix, hence with real eigenvalues and orthogonal eigenvectors, and can be decomposed as $\underline{\underline{S}}\,\underline{\underline{\Lambda}}\,\underline{\underline{S}}^{\dagger}$ where $\underline{\underline{S}}$ is a matrix whose columns are the eigenvectors of $\underline{\underline{P}}$ and $\underline{\underline{\Lambda}}$ is a diagonal matrix whose entries are the corresponding eigenvalues. By this way the effective PDL can be obtained by singular value decomposition of the transmission matrix $\underline{\underline{T}}$.

From previous considerations, a concatenation of PMD and PDL elements can be decomposed in general as $\underline{\underline{T}}(\omega) = \underline{\underline{H}}(\omega)\underline{\underline{V}}(\omega)$ where $\underline{\underline{V}}$ is a unitary matrix and $\underline{\underline{H}}$ an Hermitian positive one. Hence, $\underline{\underline{V}}$ represents an effective PMD while $\underline{\underline{H}}$ represents the effective PDL. However, $\underline{\underline{V}}$ is not equal to the concatenation of all the PMD elements nor is $\underline{\underline{H}}$ equal to the concatenation of the PDL elements. The effective PDL component can therefore be obtained from $\underline{\underline{T}}$ by taking advantage of the Hermitian properties of $\underline{\underline{H}}$:

$$\underline{\underline{T}}\,\underline{\underline{T}}^{\dagger} = \underline{\underline{H}}^2 = \underline{\underline{B}}\,\underline{\underline{\Lambda}}^2\underline{\underline{B}}^{\dagger}$$

while the effective DGD $\Delta\tau_{\text{eff}}$ can be obtained from the eigenvalues of the matrix $j\partial_{\omega}(\underline{\underline{V}})\underline{\underline{V}}^{-1}$.

## 1.3 Coupled-NLSE and Manakov-PMD equations

The coupled nonlinear Schrödinger equation (CNLSE) is the vectorial version of the scalar NLSE that takes into account of polarization effects, either linear as PMD (Sec. 1.2.3) or nonlinear, and it used to study the propagation of polarization division multiplexed (PDM) signals (Sec. 1.6).

If we represent the electric field by a complex vector with two elements $\overrightarrow{A} = [A_x, A_y]^{\mathrm{T}}$, the CNLSE can be expressed in a general form as [11, 12]:

$$\begin{aligned}\frac{\partial\overrightarrow{A}}{\partial z} + \frac{\alpha}{2}\overrightarrow{A} - j\beta_2\frac{\partial^2\overrightarrow{A}}{\partial\tau^2} + j\frac{\Delta\beta_0(z)}{2}\left(\hat{l}(z)\cdot\overrightarrow{\sigma}\right)\overrightarrow{A} + \frac{\Delta\beta_1(z)}{2}\left(\hat{l}(z)\cdot\overrightarrow{\sigma}\right)\frac{\partial\overrightarrow{A}}{\partial\tau} =\\ = -j\gamma\left[\left|\overrightarrow{A}\right|^2\overrightarrow{A} - \frac{1}{3}\left(\overrightarrow{A}^{\dagger}\sigma_3\overrightarrow{A}\right)\sigma_3\overrightarrow{A}\right]\end{aligned} \tag{1.8}$$

where $\tau = t - \frac{z}{v_g}$ is the retarded time frame, discussed in Sec. 1.2.2, while $^{\dagger}$ means transponse-conjugate. Comparing (1.8) with the *scalar* NLSE (1.1), we note two new terms appearing in the linear part, i.e., in the left-hand side (l.h.s.) of (1.8), that relate to the fiber birefringence $\Delta\beta_0$ and to the DGD $\Delta\beta_1$, while the nonlinear part, i.e., the right-hand side (r.h.s) of (1.8) is modified by the extra term with $\sigma_3$.

The symbol $\overrightarrow{\sigma}$ in the l.h.s. of (1.8) is the so-called spin-vector, whose elements are the three Pauli matrices shown in App. A, hence $\overrightarrow{\sigma}$ is actually a tensor. The scalar product $\left(\hat{l}(z)\cdot\overrightarrow{\sigma}\right)$ yields a unitary Jones matrix, i.e., a $2\times 2$ complex matrix with unit determinant, that produces mode coupling. Of the three Pauli matrices, only the third, i.e., $\sigma_3$, appears in the nonlinear term in (1.8), because 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 usually implemented for its numerical solution, namely the Manakov-PMD 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, in the numerical integration of (1.8) it is necessary to choose a step (in $z$) sufficiently small to take into account of $\Delta\beta_0$. One must choose an integration step $\Delta z$ such that the phase rotation $\Delta\beta_0\Delta z$ is small compared with $2\pi$. An important parameter of transmission fibers associated to $\Delta\beta_0$ is the ***beat length*** $L_{\mathrm{B}} = \frac{2\pi}{\Delta\beta_0}$. Hence the integration step $\Delta z$ should be smaller than $L_{\mathrm{B}}$. However, since $L_{\mathrm{B}}$ is typically of the order of meters (or tens of meters), for standard fibers the integration of (1.8) becomes extremely time-consuming.

An alternative approach, with the advantage of requiring much smaller computation time, is that of averaging the impact of signal polarization over the nonlinear term in (1.8). When $L_{\mathrm{B}}$ is small enough compared to $\Delta z$, the nonlinear effects are averaged if the rapid variations of the state of polarization of $\overrightarrow{A}$ are such as to cover uniformly the Poincaré sphere. In this case we say that the term $\frac{1}{3}\left(\overrightarrow{A}^{\dagger}\sigma_3\overrightarrow{A}\right)\sigma_3\overrightarrow{A}$ undergoes a *complete mixing*, and on average is reduced to $\frac{1}{9}\left|\overrightarrow{A}\right|^2\overrightarrow{A}$.

Hence, over sufficiently long distances to ensure complete mixing, the CNLSE can be simplified to

$$\frac{\partial\overrightarrow{A}}{\partial z}+\frac{\alpha}{2}\overrightarrow{A}-j\beta_2\frac{\partial^2\overrightarrow{A}}{\partial\tau^2}+j\frac{\Delta\beta_0(z)}{2}\left(\hat{l}(z)\cdot\overrightarrow{\sigma}\right)\overrightarrow{A}+\frac{\Delta\beta_1(z)}{2}\left(\hat{l}(z)\cdot\overrightarrow{\sigma}\right)\frac{\partial\overrightarrow{A}}{\partial\tau}=-j\gamma\frac{8}{9}\left|\overrightarrow{A}\right|^2\overrightarrow{A} \tag{1.9}$$

which is known as the ***Manakov-PMD equation*** [12]. The Manakov-PMD equation is generally regarded as a simple and reliable way to model optical fibers affected by both fiber nonlinearity and PMD. Note that although the birefringence term is still there, the great advantage is that there is no need for integration steps smaller than $L_B$. In other terms, the Manakov equation amounts to a 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 $\frac{8}{9}$. The factor of $\frac{8}{9}$ in the nonlinear coefficient has been verified experimentally [13, 14].

## 1.4 Nonlinear Impairments

The response to light of any dielectric medium 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 intensity of the field $|A|^2$ [6, 15]:

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

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

The nonlinear coefficient $\gamma[1/(\mathrm{W}\cdot\mathrm{km})]$ is due to the Kerr effect and its relation to the fiber nonlinear index $n_2$ is the following:

$$\gamma = \frac{2\pi n_2}{\lambda_0 A_{\text{eff}}} \tag{1.10}$$

where $A_{\text{eff}}\,[\mu\mathrm{m}^2]$ 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 peak power. A direct comparison between the nonlinear length and the dispersion length allows to deduce the propagation regime inside the optical fiber. $L_{\text{NL}} \gg L_{\text{D}}$ implies propagation in the dispersion-limited, or purely linear, regime; on the opposite, with $L_{\text{NL}} \ll L_{\text{D}}$ the propagation is in the nonlinear regime. The intensity dependence of the refractive index leads to a large number of nonlinear effects.

Nonlinear 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 briefly discuss SPM, XPM and XPolM. See [6] for an introduction to FWM and [16, 17] for advanced studies of FWM.

### 1.4.1 Self Phase Modulation

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

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

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

Neglecting $\beta_{1,2,3}$, the NLSE in a single channel transmission is:

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

whose solution is:

$$A(z,t) = A(0,t) \mathrm{e}^{-\frac{\alpha}{2} z - j\gamma |A(0,t)|^2 L_{\mathrm{eff}}(z)} = A(0,t) \mathrm{e}^{-\frac{\alpha}{2} z - j\Phi_{\mathrm{SPM}}(z,t)} \tag{1.11}$$

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

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

which is a measure of the distance over which the nonlinear effect is significant.

Note from (1.11) 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 energy conservation principle the energy carried by at 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 rotation in the time domain.

### 1.4.2 Cross Phase Modulation

XPM refers to the nonlinear phase shift of an optical field induced by another field having a different wavelength. Considering a WDM simplified scenario in which two scalar optical fields (single polarization) $A_a$ and $A_b$ at frequencies $\omega_a$ and $\omega_b$ respectively, propagate simultaneously inside the fiber, the nonlinear phase shift for the field at $\omega_a$ is given by

$$\phi_{\mathrm{NL}} = n_2 k_0 L \left( |A_a|^2 + 2\,|A_b|^2 \right) \tag{1.12}$$

where we have neglected all terms that generate power at frequencies other than $\omega_a$ and $\omega_b$. The two intensity-dependent terms on the r.h.s. of (1.12) are due to SPM and XPM, respectively. An important feature of XPM is that its contribution to the nonlinear phase shift is twice that of SPM for equally intense optical fields of different wavelengths. 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:

$$A(z,t) = \sum_{k=1}^{M} A_k(z,t) \mathrm{e}^{j\Delta\omega_k t} \tag{1.13}$$

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 $\omega_0$ of the WDM comb. The nonlinear part of (1.1) for $A_k(z,t)$ for $k = 1,\ldots,M$ is:

$$\sum_k \frac{\partial A_k}{\partial z} \mathrm{e}^{j\Delta\omega_k t} = -j\gamma \sum_{n,l,m} A_n A_l A_m^* \mathrm{e}^{j(\Delta\omega_n + \Delta\omega_l - \Delta\omega_m)t} \,. \tag{1.14}$$

We can rewrite (1.14) as:

$$\frac{\partial A_k}{\partial z} = -j\gamma \sum_{\substack{n,l,m \\ \omega_n + \omega_l - \omega_m = \omega_k}}^{M} A_n A_l A_m^* \qquad k = 1,\ldots,M \tag{1.15}$$

where $n$, $l$, $m$ can range from 1 to $M$ but must satisfy $\omega_n + \omega_l - \omega_m = \omega_k$. Hence in (1.15) we discarded all the 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 $\omega_k$, $k = 1, \ldots, M$ 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 (1.15) yields:

$$\frac{\partial A_k}{\partial z} = -j\gamma \left( |A_k|^2 A_k + 2 \sum_{m \neq k} |A_m|^2 A_k \right) . \tag{1.16}$$

### 1.4.3 Cross Polarization Modulation

Eq. (1.16) highlights the contribution to the nonlinear part of (1.1) coming from both single-channel and cross-channel effects on condition that the WDM transmitted comb includes scalar optical fields. When PDM is employed, nonlinear effects does not limit to a nonlinear phase rotation but alter also the state of polarization of the field. XPolM refers to a nonlinear change in the state of polarization of an optical field induced by another field having a different wavelength.

To derive the equivalent expression of (1.16) in case of PDM transmission, we consider the complex optical field $\overrightarrow{A}(z,t)$ expressed as:

$$\overrightarrow{A}(z,t) = \sum_{k=1}^{M} \overrightarrow{A_k}(z,t) \mathrm{e}^{j\Delta\omega_k t} \tag{1.17}$$

with $\overrightarrow{A_k}(z,t) = \left[A_{kx}(z,t), A_{ky}(z,t)\right]^{\mathrm{T}}$. Removing all term related to linear impairments in (1.9) we obtain

$$\frac{\partial \overrightarrow{A}}{\partial z} = -j\gamma \frac{8}{9} \left| \overrightarrow{A} \right|^2 \overrightarrow{A} . \tag{1.18}$$

Following the same procedure of Sec. 1.4.2 in discarding all the terms falling outside the bandwidth of $\overrightarrow{A_k}(z,t)$ we can write (1.18) as [18]

$$\frac{\partial \overrightarrow{A_k}}{\partial z} = -j\gamma\frac{8}{9}\left\{\underbrace{|A_k|^2\overrightarrow{A_k}}_{\text{SPM}} + \underbrace{\frac{3}{2}\sum_{m\neq k}|A_m|^2\overrightarrow{A_k}}_{\text{XPM}} + \underbrace{\frac{1}{2}\sum_{m\neq k}\left(\overrightarrow{a_m}\cdot\overrightarrow{\sigma}\right)\overrightarrow{A_k}}_{\text{XPolM}}\right\} \tag{1.19}$$

in which we now distinguish among cross-channel effects according on their impact on the propagating channel. We identify XPM term as the responsible of the induced nonlinear phase rotation, while XPolM term is the responsible of a nonlinear change of the field state of polarization. The factor $3/2$ in XPM is a polarization-averaged value, as XPM runs between a value 1 when $\overrightarrow{A_m}$ is orthogonal to $\overrightarrow{A_k}$, with $m \neq k$, and the classic value 2 of scalar propagation.

## 1.5 SSFA for Numerical Solution of NLSE

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

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

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:

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

are two differential equations easy to solve. This is the case of the NLSE where closed form solutions exist considering the only presence of fiber dispersion or with the only presence of fiber nonlinearity. Hence with the SSFA the field propagation through the fiber is modeled as a concatenation of small steps within which the two operators are applied separately and alternately. 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 reduced as well, and $A(L,t)$ is closer to the exact solution. Regarding SSFA implementation, the linear operator $\mathscr{L}$ is efficiently evaluated in the frequency domain while the nonlinear operator $\mathscr{N}$ in the time domain. Such approach calls for Fast Fourier Transform (FFT) and Inverse FFT (IFFT) to switch efficiently between the two domains.

The choice of the SSFA 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 SSFA step size is to run many simulations for decreasing step sizes until some convergence is observed. Focusing on the nonlinear step, we must ensure a correct description of the relative motion of the channels to not incur in overestimation of cross-channel nonlinearities. To this aim, within a single step the edge channels in a WDM comb must walk past each other by a time smaller than the symbol time, with smaller the time better the accuracy. In literature this method is called *walk-off method* [20].

### 1.5.1 Nonlinearity-decoupling method

In this section we describe the nonlinearity-decoupling method, which will be used in chapters 2 and 4 to selectively activate the desired nonlinear effect, namely SPM, XPM or XPolM, in the nonlinear step of the SSFA. The Manakov nonlinear step for the $k$-th channel of a PDM WDM transmission is expressed in (1.19) where are identified SPM, XPM and XPolM contributions to the overall nonlinearity. At each time instant, the exact solution of (1.19) when SPM, XPM and XPolM are simultaneously active is (we omit the time dependence, since the nonlinear step is memoryless):

$$\overrightarrow{A_k}(z) = \mathrm{e}^{\frac{-j\bar{\gamma}z\left(A_k^2(0)+3\sum_{m\neq k}A_m^2(0)\right)}{2}}\,\underline{e}^{-\frac{j\bar{\gamma}z}{2}(\overrightarrow{s_t}(0)\cdot\overrightarrow{\sigma})}\overrightarrow{A_k}(0) \tag{1.20}$$

where $\bar{\gamma} = (8/9)\gamma$ is the Manakov-averaged nonlinear coefficient, $\overrightarrow{s_t}(z) = \overrightarrow{s_t}(0) \triangleq \sum_m \overrightarrow{a_m}(0)$ is the real pivot vector (a constant in $z$), and $\underline{e}^{(\cdot)}$ denotes a matrix exponential, which can be computed for the generic real vector $\overrightarrow{v}$ of norm $|\overrightarrow{v}|$ as [21]:

$$\underline{e}^{-j(\overrightarrow{v}\cdot\overrightarrow{\sigma})} = \cos(|\overrightarrow{v}|)\sigma_0 - j\sin(|\overrightarrow{v}|)\left(\frac{\overrightarrow{v}}{|\overrightarrow{v}|}\cdot\overrightarrow{\sigma}\right) \tag{1.21}$$

where $\sigma_0$ is the 2x2 identity matrix.

When, instead, we desire to simulate the nonlinear step of SSFA with only one nonlinear effect in (1.19) active, (1.19) solution becomes respectively:

$$\overrightarrow{A_k}_{,\mathrm{SPM}}(z) \triangleq \mathrm{e}^{-j\bar{\gamma}A_k^2(0)z}\overrightarrow{A_k}(0) \tag{1.22}$$

$$\overrightarrow{A_k}_{,\mathrm{XPM}}(z) \triangleq \mathrm{e}^{-j\bar{\gamma}\frac{3}{2}\sum_{m\neq k}A_m^2(0)z}\overrightarrow{A_k}(0) \tag{1.23}$$

$$\overrightarrow{A_k}_{,\mathrm{XPolM}}(z) \triangleq \mathrm{e}^{\frac{j\bar{\gamma}A_k^2(0)z}{2}}\underline{e}^{-\frac{j\bar{\gamma}}{2}(\overrightarrow{s_t}(0)\cdot\overrightarrow{\sigma})z}\overrightarrow{A_k}(0) \tag{1.24}$$

In the case, labeled as XCI, where both cross-channel nonlinearities, i.e., XPM and XPolM, are active, the solution of (1.19) is:

$$\overrightarrow{A_k}_{,\mathrm{XCI}}(z) \triangleq \mathrm{e}^{-j\bar{\gamma}\frac{3}{2}\sum_{m\neq k}A_m^2(0)z}\overrightarrow{A_k}_{,\mathrm{XPolM}}(z)\,. \tag{1.25}$$

Note that eq. (1.19) is indeed a set of coupled differential equations, with as many equations as the number of channels $M$ in the WDM comb. We will refer to it as the *separate-field* propagation (SFP) [4]. Hence a SSFA simulation with SFP consist of forming a matrix whose columns are the fields $\overrightarrow{A_k}$ discretized over the FFT grid corresponding to a frequency window equal to the channel spacing $\Delta f$. In our implementation we allocate a window larger than $3\Delta f$ to avoid aliasing due to nonlinear broadening. Then the matrix is propagated along $z$ such that at each step:

i) independent Gaussian ASE noise samples are added at each grid point if the $z$ coordinate corresponds to the position of an EDFA;

ii) the $M$ signal fields are impaired only by linear effects;

iii) the $M$ fields are impaired only by nonlinear effects, according to (1.22)-(1.25).

Note that point i) is applied only in the case of distributed noise, and amounts to adding to each channel a white ASE over a bandwidth equal to the channel spacing $\Delta f$.

We refer to the *unique-field* propagation (UFP), instead, when the WDM signal (1.17) is treated as a unique channel and the Manakov nonlinear step is just $\frac{\partial\overrightarrow{A}}{\partial z} = -i\bar{\gamma}A^2\overrightarrow{A}$. In this case a much shorter sampling time is required, typically of the order of the inverse of 3 times the WDM occupied bandwidth, to account for at least first-order FWM.

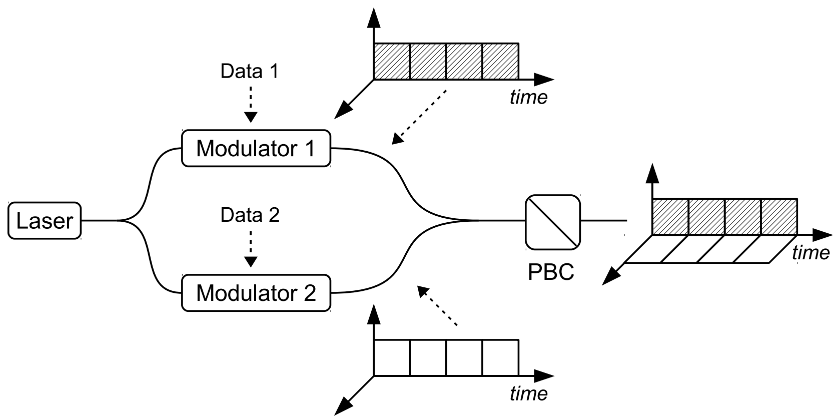


Figure 1.2: Scheme of PDM transmitter.

## 1.6 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 to doubling the spectral efficiency. For example, PDM-quaternary phase shift keying (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. In Fig. 1.2 is depicted the scheme of a PDM transmitter which operates as following. The light from the continuous wave (CW) laser is split into two copies through a 3 dB coupler and each copy is sent into a distinct modulator driven with different electrical data. The output of the two modulators is then mapped on the two polarizations of the PDM channel with a polarization beam combiner (PBC). To ensure a correct procedure, all the components used in the PDM generation process must be polarization maintaining. Usually, we refer to the two polarization tributaries as either parallel/orthogonal or X/Y tributaries.

PDM transmission can be combined with different modulation formats such as binary-PSK (BPSK) [22], QPSK [23] or 16 quadrature amplitude modulation (16QAM) [24]. In this thesis we will resort to all these options, namely PDM-BPSK, PDM-QPSK and PDM-16QAM.

To ensure a given total bit rate, the symbol rate of each PDM tributary is halved with respect to an equivalent transmission with single polarization signals, because the the total bit rate of the PDM signal is the sum of the bit rate of both polarization tributaries. Hence, at the same total bit rate, the optical bandwidth of PDM signals is halved with respect to signals employing only one polarization. Consequently, the requirements on bandwidth of opto-electronic devices as well as the drive frequency of modulators are relaxed, being halved with respect to single polarization transmissions. As a drawback, PDM transmission requires almost twice as many components.

The narrower spectral width 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 [25]. Furthermore, narrow signal spectra allow the deployment 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 [26]. On the other hand, tolerance against nonlinearities decreases, in general, with the application of PDM due to nonlinear interactions between both polarizations [27] and to the lower symbol rate which makes signals more sensitive against cross nonlinearities [28].

At same symbol rate, the OSNR required from a PDM signal is 3 dB higher than a single polarization signal, as a consequence of the total signal power split on the two orthogonal components that carry different information.

The major downside of PDM transmission systems is represented by receiver complexity. However, recent advances in electronics allow robust reception thanks to coherent detection (Sec. 1.7) and advanced signal processing. Hence, paired with coherent detection and digital signal processing, PDM allows to design optical transmission systems with increased spectral efficiency and higher robustness against linear effects.

## 1.7 Coherent Detection

In the coherent receiver, the signal is detected by the beating with a local oscillator (LO) laser. This solution was proposed at the end on 1980s to achieve better receiver sensitivity and longer unrepeated distance [29–31] but set aside at the beginning of the 1990s with the introduction of the Erbium Doped Fiber Amplifiers (EDFA). In fact, the beating with LO laser, much more powerful than the signal, serves as a signal amplifier and the receiver sensitivity can be improved by up to 20 dB compared to direct detection without pre-amplification [32]. With EDFA amplification longer unregenerated distance can be achieved by periodically amplifying the optical signal and better sensitivity can be achieved by pre-amplifying the received signal. Since direct detection in conjunction with EDFA pre-amplification gives very similar performance to coherent detection without most of the technical issues, further research in coherent optical communications have been dropped for twenty years.

In the last decade, coherent technologies have restarted to become interesting solutions to meet the ever-increasing bandwidth demand with multi-level modulation formats. Besides, today is possible to treat the electrical signal in a digital signal processing (DSP) core thanks to the development of high-speed digital electronics and analog-to-digital converters (ADC). DSP opened the possibility of using powerful algorithms to compensate for linear impairments incurring along the optical fiber link, such as chromatic dispersion and PMD [33, 34]. Although PMD is a time varying distortion, the digital coherent receiver can equalize it through adaptive algorithms. Also nonlinear impairments can be partially mitigated in the DSP unit with the recent development of digital back-propagation (DBP) techniques.

Hence, today's coherent receivers enjoy the high sensitivity of coherent detection with all the benefits of digital signal processing.

### 1.7.1 Coherent Receiver description

The scheme of the coherent receiver is depicted in Fig. 1.3. The optical signal beats with the light of a LO laser hence down-converting the signal from the optical carrier frequency to microwave carrier frequency in the range of GHz or tens of GHz. An optical beat signal is then generated at a frequency equal to an intermediate frequency $\omega_{\mathrm{IF}}$

$$\omega_{\mathrm{IF}} = \omega_{\mathrm{s}} - \omega_{\mathrm{ol}}$$

that is the difference between the frequency of the received signal $\omega_{\mathrm{s}}$ and that of the LO laser $\omega_{\mathrm{ol}}$. If the optical frequency of the signal is the same as that of the LO laser, the system is called **homodyne**, otherwise the system is called **heterodyne**.

Homodyne reception requires the LO frequency to be strictly locked in frequency and phase to the received signal and gives optimal receiver sensitivity [32]. Hence the frequency and phase of the LO must be controlled and adjusted continuously, a task that is usually performed with a phase-locked loop (PLL).

Heterodyne reception has the advantage to relax the constraints on the linewidth of the lasers, however as drawbacks it requires a receiver bandwidth at least twice the symbol rate [35], i.e., doubled if compared to what needed by homodyne detection, and the sensitivity is at least 3 dB worse compared to homodyne detection, due to the effective energy of a heterodyne-detected signal that is half of the effective energy with homodyne detection [36, 37]. Therefore, the implementation of heterodyne reception, easier in principle than homodyne detection, becomes challenging for high bit rate operations.

A halfway solution is provided by **intradyne** systems, where the frequency of the LO is *approximately* the same as the frequency of the signal. The frequency mismatch between the signal and the LO is recovered in the digital domain by processing the baseband signal. By this way there is no need to continuously adjust the LO frequency and phase, with the maximum tolerable frequency-difference determined by the signal processing. Furthermore, a phase-diversity receiver is essential as the LO is not phase-locked.

Intradyne detection aims to detect the in-phase and quadrature components of the signal and therefore a 90° hybrid is required. Both components of the (baseband) optical field can be transferred to the electrical domain by ADCs; thus, the same receiver can operate with any kind of optical modulation format. Moreover, any phase-drift can be compensated through DSP.

Polarization-diversity receivers can be considered as multiple-input multiple-output

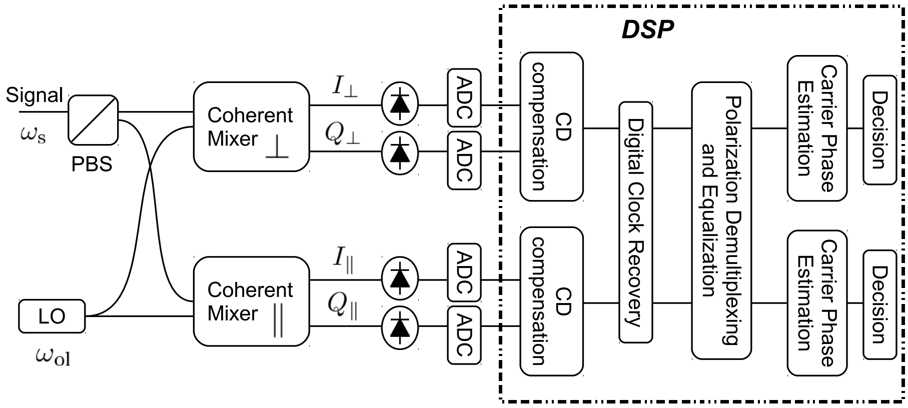


Figure 1.3: Coherent Receiver scheme.

(MIMO) receivers as they offer the potential of detecting PDM signals without any additional component. Okoshi et al. were the firsts, in 1987, to demonstrate a receiver combining intradyne detection with polarization diversity in which all the characteristics of the optical field were translated to the electrical domain [38].

Today's coherent receivers, use on one hand, intradyne polarization-diversity detection to convert all the characteristics of the optical field (i.e., amplitude, phase and polarization) to the electrical domain [31] and on the other hand advanced algorithms to compensate for transmission impairments [33, 34]. This requires the detection and digitalization of four signals, i.e., the in-phase (I) and quadrature (Q) components of two arbitrary, but orthogonal, polarization states (Fig. 1.3).

### 1.7.2 Coherent Receiver scheme

The scheme of the coherent receiver is depicted in Fig. 1.3. A polarization beam splitter (PBS) splits the signal in two and each one of the incoming polarization is sent into two coherent mixers. The four optical interference signals are then sampled, and digitalized by ADCs. 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 the DSP unit and symbols are then identified by a simple threshold method. The DSP is done in several steps as described in Fig. 1.3.

The first DSP operation is the compensation of the CD, either by using Finite Impulse Response (FIR) filters, or by using the Fast Fourier Transform (FFT) method as described in [39]. These filters may be very long, depending on the amount of CD to be compensated for, but are static and do not require 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 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 [40]. The filters used within this part, have to adapt themselves continuously to the incoming signal, to follow polarization fluctuation and PMD variations [33]. 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 LO and the carrier frequency of the signal as described in [40, 41].

More details about the coherent receiver and DSP algorithms can be found in [42, 43].

# Chapter 2

# Nonlinear Threshold in Long-haul Coherent Transmissions

The adoption of digital signal processing (DSP)-based coherent reception in long-haul optical transmissions led to several advantages. For instance, it provided strong resilience against linear impairments as polarization mode dispersion (PMD) and cumulated group velocity dispersion (GVD) [2]. The ability of DSP to compensate at the receiver for large amounts of cumulated GVD and to adaptively track the slow (if compared to symbol time) dynamics of PMD, opened the possibility to design long-haul optical links without dispersion management, i.e., dispersion uncompensated (DU) links [44]. The great amount of GVD cumulated in DU links before reception mitigates the effects of fiber nonlinearity, thus allowing an increased system reach compared to dispersion managed (DM) links, at equal bit error rate (BER) performance.

DSP also enabled the use of the polarization division multiplexing (PDM) technique, and provided higher system capacity without resorting to a wider spectrum. The further increase of demanded capacity also caused the increase of the bandwidth efficiency, i.e., a more tight spacing among wavelength division multiplexed (WDM) channels, and the use of modulation formats with higher spectral efficiency.

However, for a selected bandwidth efficiency, it is important to know for system design purposes how the granularity of the WDM spectrum, i.e., the number of carriers to fill the system aggregate bandwidth, affects the achievable distance when delivering a target BER [45].

The aim of this chapter is to provide this information, by a systematic simulation study. We study, at a fixed bandwidth efficiency, the symbol rate maximizing the achievable distance, i.e, the *reach*, while assuring a desired BER, for several single-carrier modulation formats of interest, in both DM and DU links. In each studied case, we also highlight the individual contribution of each and every Kerr nonlinearity, namely, self-phase modulation (SPM), cross-phase modulation (XPM) and cross-polarization modulation (XPolM). Our objective is to estimate the constrained nonlinear threshold (NLT) as the per-channel symbol rate varies. We also present a new formula yielding the system reach based on knowledge of NLT measurements. Later on, for selected system configurations, we perform reach simulations with the twofold aim of checking the accuracy of the new reach formula and of highlighting the accumulation rate of fiber nonlinearity with distance.

## 2.1 Nonlinear Threshold for System Design

### 2.1.1 Nonlinear Threshold

The performance of an optical transmission system is usually evaluated in terms of the achieved average BER or equivalently, after a one-to-one conversion, in terms of the average Q-factor $Q = \sqrt{2}\mathrm{erfc}^{-1}(2\mathrm{BER})$. $Q$ plays here the role of the electrical signal-to-noise ratio (SNR) in a fictitious equivalent on-off keying (OOK) transmission. At a fixed distance, the Q-factor vs transmitted power plot can be represented in two ways: at fixed noise figure (NF), i.e., optical SNR (OSNR) increases for increasing transmitted power, or at fixed OSNR, i.e., the NF increases for increasing transmitted power. In the first case, Q-factor vs power plot is usually referred to as *bell curve*. An example of bell curve is shown in Fig. 2.1.

On the bell curve we can identify two regions: i) the “linear region”, i.e., the ascending left part of the curve, where the performance is limited by the Amplified Spontaneous Emission (ASE) noise and $Q$ increases with increasing power; ii) the

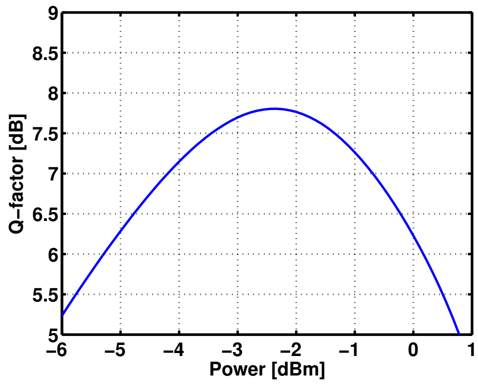


Figure 2.1: Q-factor versus Power at fixed distance. (Amplifiers NF fixed).

"nonlinear region", i.e., the descending right part of the curve, where the performance is limited by fiber nonlinearity and $Q$ decreases with increasing power. We refer to the power delivering the highest Q-factor, i.e., the lowest BER, as the *unconstrained* Nonlinear Threshold (NLT).

Instead of reasoning in terms of Q-factor, whenever the BER is due to additive gaussian noise, we can rely on the electrical SNR $S$ at the decision gate. This is the case with coherent reception including polarization demultiplexing (assuming PDM transmission), ideal linear electrical equalization, followed by matched filtering and ideal carrier estimation. The SNR at the decision gate is:

$$S = \frac{P}{N_{\text{ASE}} + N_{\text{NLI}}} \tag{2.1}$$

where $N_{\text{ASE}}\,[\text{mW}]$ is the ASE noise power, $N_{\text{NLI}}\,[\text{mW}]$ is the Nonlinear Interference (NLI) noise power, both measured over the optical signal's bandwidth, and $P\,[\text{mW}]$ is the per-channel average signal power. It has been proved [46, 47] that NLI can be modeled as a signal-independent circular Gaussian noise like the ASE noise. The BER can be thus expressed as $\text{BER} = f(S)$, where the functional relation depends on the modulation format [48]. Hence the optimization of the BER becomes equivalent to the optimization of the SNR $S$.

Recently, a Gaussian Noise (GN) model, that assumes the NLI as a zero-mean

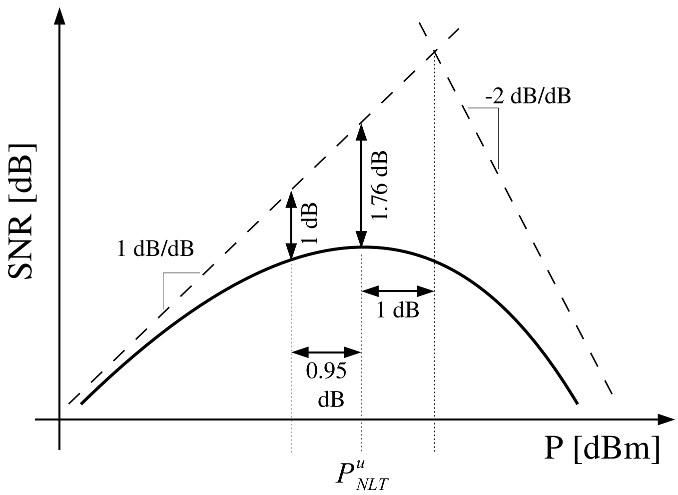


Figure 2.2: SNR versus Power at fixed distance.

signal-independent circular Gaussian additive noise, was proposed to justify and predict the performance of DU coherent communications with orthogonal frequency division multiplexing (OFDM) [49,50] and with single-carrier signals [51,52]. Alternative formulations express such NLI as a multiplicative noise according to the logarithmic perturbation method [53]. In this work we follow the additive noise approach. According to the GN model, the NLI noise power is approximated by $N_{\text{NLI}} = a_{\text{NL}} P^3$, where $a_{\text{NL}}\,[\text{mW}^{-2}]$ is a power-independent coefficient, hence (2.1) becomes

$$S = \frac{P}{N_{\text{ASE}} + a_{\text{NL}} P^3}\,. \tag{2.2}$$

In this case at the unconstrained NLT power, where the SNR $S$ is maximized, the ASE noise power $N_{\text{ASE}}$ is twice the NLI noise power $N_{\text{NLI}}$ [51]. It has been demonstrated [51] that at NLT power, both Q-factor $Q$ and SNR $S$ are maximized.

The unconstrained NLT $P^{\text{u}}_{\text{NLT}}$ it can be shown that the SNR penalty with respect to (w.r.t.) linear propagation is 1.76 dB [54]. $P^{\text{u}}_{\text{NLT}}$ is 1 dB lower that the break-point power $P_{\text{B}}$ where ASE power equals the NLI power, as shown by Fig. 2.2. The power $P^{\text{u}}_{\text{NLT},1}$ yielding 1 dB of SNR penalty w.r.t. linear propagation is 0.95 dB lower than $P^{\text{u}}_{\text{NLT}}$. The *constrained* NLT $P_{\text{NLT}}$, instead, is defined as the power that maximizes the

SNR when the maximum reaches a desired value $S_0$. The required ASE noise power to achieve the desired SNR $S_0$ is $N^*_{\mathrm{ASE}} = \frac{2}{\sqrt{a_{\mathrm{NL}}(3S_0)^3}}$, thus yielding the constrained NLT as

$$P_{\mathrm{NLT}} = \frac{3}{2} S_0 N^*_{\mathrm{ASE}} = \frac{1}{\sqrt{3S_0 a_{\mathrm{NL}}}}. \tag{2.3}$$

Hence, when $N_{\mathrm{ASE}} > N^*_{\mathrm{ASE}}$ the desired SNR $S_0$ is not achievable, while for $N_{\mathrm{ASE}} < N^*_{\mathrm{ASE}}$ there are two powers achieving $S_0$, the largest one $P_{\mathrm{H}}$ in the “nonlinear region”, the other one $P_{\mathrm{L}}$ in the “linear region”. In the constrained scenario, the power $P_{\mathrm{NLT},1}$ yielding 1 dB of SNR penalty w.r.t. linear propagation is 1.05 dB lower [55] than $P_{\mathrm{NLT}}$, where the SNR penalty is always 1.76 dB.

The procedure we adopt to measure the *constrained* 1-dB NLT $P_{\mathrm{NLT},1}$ at reference $\mathrm{BER}_0$, i.e., SNR $S_0$, is the following:

1. We measure the OSNR value $\mathrm{OSNR}_0\,[\mathrm{dB}]$, on a 0.1 nm bandwidth, that yields the reference SNR $S_0\,[\mathrm{dB}]$ in back-to-back (b2b) transmission. Their relation is $\mathrm{OSNR}_0 = S_0 - 10\log_{10}(B_{RX}/\Delta\upsilon)$ where $B_{RX} = R$ is the matched-filter electrical bandwidth and $\Delta\upsilon = 12.5\,\mathrm{GHz}$ is the conventional 0.1 nm measurement bandwidth of the optical spectrum analyzer. Since $\mathrm{OSNR}_0$ scales with the symbol rate $R\,[\mathrm{Gbaud}]$, we first measure it at one symbol rate, 10Gbaud in our case, and we scale it to any symbol rate $R$ as $\mathrm{OSNR}_0^R = \mathrm{OSNR}_0^{10} + 10\log_{10}(R/10)$.

2. Since we search the 1-dB NLT, we set the OSNR as $\mathrm{OSNR}_{\mathrm{ref}} = \mathrm{OSNR}_0 + 1\,\mathrm{dB}$.

3. We fix a tentative transmitted power $P_{\mathrm{tx}}$. The amplifiers NF is set in order to achieve $\mathrm{OSNR}_{\mathrm{ref}}$. We simulate the wavelength division multiplexed (WDM) comb propagation along the link through the Split Step Fourier algorithm (SSFA) and the BER is evaluated with a standard Monte Carlo algorithm.

4. If $\mathrm{BER} > \mathrm{BER}_0$ ($\mathrm{BER} < \mathrm{BER}_0$) we decrease (increase) $P_{\mathrm{tx}}$ and repeat the point 3. Note that the amplifiers NF is tuned together with $P_{\mathrm{tx}}$ in order to impose $\mathrm{OSNR}_{\mathrm{ref}}$ value.

5. $P_{\mathrm{NLT},1}$ is found by interpolation of the BER versus $P_{\mathrm{tx}}$ plot for the few tested points around $\mathrm{BER}_0$.

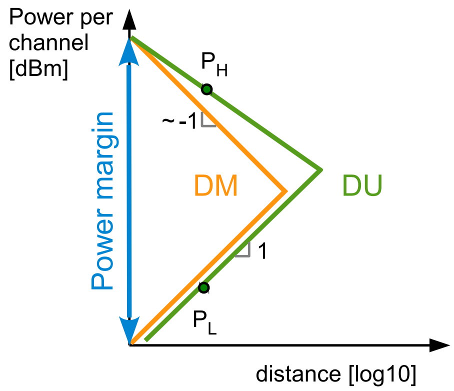


Figure 2.3: Simplified sketch of received BER contours versus both launched power and distance. $P_{\mathrm{H}}$ is in the nonlinear region. $P_{\mathrm{L}}$ is in the linear region. Their difference is the power margin.

In sec. 2.1.2 we will show that the knowledge of the constrained NLT is a valuable information for system design purposes if used together with the system parameter $\varepsilon$ accounting for the accumulation rate of the NLI with distance.

### 2.1.2 System Design

The design of long-haul optical links is traditionally based on the maximization of the system reach, defined as the maximum distance bridged by a channel without regeneration. To this goal, it is customary to evaluate the reach from BER, or equivalently Q-factor, contours versus both launched power and distance [56, 57]. A simplified straight-line sketch of such contours is given in Fig. 2.3 for both a DM and a DU system. The analysis can be performed equivalently on SNR contours in which the constrained NLT provides partial information on the upper branch, accounting for the "nonlinear region".

In the Gaussian noise model (2.2), SNR dependence on transmission distance appears implicitly in ASE noise power that scales linearly with the number of spans $N$ as

$$N_{\mathrm{ASE}} = \beta N \tag{2.4}$$

with $\beta = h\nu FGR$, where $h\nu$ is the photon energy at frequency $\nu$, $F$ is the amplifiers noise figure, $G$ is the amplifiers gain and $R$ is the matched-filter electrical bandwidth, equal to symbol rate. In NLI noise power, the dependence on distance is hidden in the power-independent coefficient $a_{\mathrm{NL}}$ that can be pragmatically approximated as

$$a_{\mathrm{NL}} = \alpha_{\mathrm{NL}} N^{1+\varepsilon} \tag{2.5}$$

where $\alpha_{\mathrm{NL}}$ is a constant and $\varepsilon$ is a system factor whose value is between 0 and 1, since the variance $N_{\mathrm{NLI}}$ of the NLI field scales as $N^1$ if the span contributions are uncorrelated equal-variance random variables (RVs), or as $N^2$ if the span contributions are the same RV.

In this picture, the $S_0$-constrained NLT $P_{\mathrm{NLT}}(N)$ at distance $N$ is a valuable information for the following reasons. Once we evaluate $P_{\mathrm{NLT}}(N)$, we also know the NF value $F_{\mathrm{NLT}}$ yielding $N^*_{\mathrm{ASE}}$, as stated in point 4 of the procedure in Sec. 2.1.1. $F_{\mathrm{NLT}}$ may take unpractical values during the constrained NLT measurement, in order to impose the target performance at distance $N$. However the constrained NLT $P_{\mathrm{NLT}}$ can be useful also in the unconstrained scenario. In fact, at distance $N$, when $N_{\mathrm{ASE}} > N^*_{\mathrm{ASE}}$, i.e., $F > F_{\mathrm{NLT}}$, the desired SNR $S_0$ is not achievable, while for $N_{\mathrm{ASE}} < N^*_{\mathrm{ASE}}$ there is a range of powers allowing $S > S_0$. The extremes of this range are $P_{\mathrm{H}}$ in the "nonlinear region" and $P_{\mathrm{L}}$ in the "linear region" and their difference is the *power budget* at distance $N$ (see Fig. 2.3). The expressions of $P_{\mathrm{H}}$ and $P_{\mathrm{L}}$ are [54, 55]:

$$\begin{aligned} P_{\mathrm{H}} &= 3S_0 N^*_{\mathrm{ASE}} \cos\left(\frac{\arccos(-N_{\mathrm{ASE}}/N^*_{\mathrm{ASE}})}{3}\right) \\ P_{\mathrm{L}} &= 3S_0 N^*_{\mathrm{ASE}} \cos\left(\frac{2\pi - \arccos(-N_{\mathrm{ASE}}/N^*_{\mathrm{ASE}})}{3}\right) . \end{aligned} \tag{2.6}$$

As shown in Fig. 2.3, $P_{\mathrm{H}}$ and $P_{\mathrm{L}}$ are the upper and lower branches, respectively, of the SNR contour versus both launched power and distance. Each branch has an asymptote, obtained by suppressing either ASE or NLI, respectively. The linear asymptote can be expressed as

$$P^{\mathrm{A}}_{\mathrm{L}} = P_{\mathrm{LT}}(1) N \tag{2.7}$$

where $P_{\mathrm{LT}}(1) = \beta S_0$ is the 1-span linear threshold achieving $S_0$ in the linear regime, hence it is a straight line with positive slope 1 dB/dB versus distance. On the upper branch, the nonlinear asymptote is

$$P_{\mathrm{H}}^{\mathrm{A}} = \frac{1}{\sqrt{S_0 a_{\mathrm{NL}}}} = \frac{1}{\sqrt{S_0 \alpha_{\mathrm{NL}} N^{1+\varepsilon}}} \tag{2.8}$$

hence a straight line with a negative slope $-(1+\varepsilon)/2$ dB/dB.

The two asymptotes cross at the point $N_0^{\mathrm{A}}$ which can be found from the values of the asymptotes at distance $N$ as

$$\frac{N_0^{\mathrm{A}}}{N} = \left(\frac{P_{\mathrm{H}}^{\mathrm{A}}(N)}{P_{\mathrm{L}}^{\mathrm{A}}(N)}\right)^{\frac{2}{3+\varepsilon}}$$

and $N_0^{\mathrm{A}}$ overestimates the true reach $N_0$ by $2.76/(1+\varepsilon/3)$.

The nonlinear equivalent of the 1-span linear threshold $P_{\mathrm{LT}}(1)$ is the 1-span constrained NLT power, which can be expressed from (2.3) and (2.5) as

$$P_{\mathrm{NLT}}(1) = \frac{1}{\sqrt{3 S_0 \alpha_{\mathrm{NL}}}} = \frac{P_{\mathrm{NLT}}(N)}{N^{-\frac{1+\varepsilon}{2}}}\,. \tag{2.9}$$

Since the true reach $N_0$ can be expressed as [55]:

$$N_0 = \left(\frac{P_{\mathrm{NLT}}(1)}{\frac{3}{2} P_{\mathrm{LT}}(1)}\right)^{\frac{2}{3+\varepsilon}} \tag{2.10}$$

by re-writing (2.9) as $P_{\mathrm{NLT}}(N)N = P_{\mathrm{NLT}}(1)N^{-\frac{1-\varepsilon}{2}}$ and from (2.10) we obtain

$$\frac{N_0}{N} = \left(\frac{P_{\mathrm{NLT}}(N)}{\frac{3}{2} P_{\mathrm{L}}^{\mathrm{A}}(N)}\right)^{\frac{2}{3+\varepsilon}} = \left(\frac{P_{\mathrm{NLT}}(N)/R}{\frac{3}{2} N h \nu F G S_0}\right)^{\frac{2}{3+\varepsilon}}\,. \tag{2.11}$$

The question is now how these SNR contours, and thus the reach, change as the symbol rate $R$ is changed. In the next section we will address this question by investigating a multiplicity of scenarios including both DM and DU links, and several single-carrier modulation formats of interest. It is clear that if $\varepsilon$ is independent of $R$, by maximizing the power margin $P_{\mathrm{NLT}}/P_{\mathrm{L}}^{\mathrm{A}}$ one also maximizes the reach. Hence, since $P_{\mathrm{L}}^{\mathrm{A}}$ scales linearly with $R$, the largest power margin is achieved at the symbol rate that maximizes the margin $P_{\mathrm{NLT}}/R$. In Sec. 2.2 we provide a thorough investigation by simulation of the *normalized margin* $P_{\mathrm{NLT},1}/R$ which provides the same information since for constrained NLT we have $P_{\mathrm{NLT},1} = P_{\mathrm{NLT}} - 1.05\,\mathrm{dB}$ [55].

## 2.2 Optimal Granularity in single-carrier WDM coherent transmissions

The issue of the best granularity in WDM long-haul coherent optical links, i.e., the per-carrier symbol rate for OFDM, and the per-channel symbol rate for single-carrier modulations, was recently discussed [45, 58–63]. The best symbol rate either minimizes the BER at a fixed distance, or maximizes the largest transparent transmission distance, i.e., the reach, at a fixed BER. In this section we tackle the issue by reporting on the 1-dB constrained NLT $P_{\mathrm{NLT},1}$ at BER=$10^{-3}$ versus symbol rate $R$, at a constant bandwidth efficiency $\eta \triangleq R/\Delta f$, where $\Delta f$ is the channel spacing, for both DM and DU systems, and for the following modulation formats: binary phase-shift keying (PDM-BPSK), quaternary phase-shift keying (PDM-QPSK), polarization switched QPSK (PS-QPSK), and 16-level quadrature amplitude modulation (PDM-16QAM). For all formats we used non-return to zero (NRZ) supporting pulses, while for PDM-BPSK and PDM-QPSK we also considered iRZ pulses with 50% duty cycle [64–66]. In each considered case, we applied the nonlinearity decoupling method described in sec. 1.5.1 to yield the contribution from each manifestation of Kerr nonlinearity to set the NLT. We present the results in terms of the normalized margin $P_{\mathrm{NLT},1}/R$ to highlight both the dominant nonlinearity for each format, and the optimal granularity for maximum reach.

### 2.2.1 Simulated Systems

Fig. 2.4 shows the optical systems simulated with the open-source software Optilux [4]. The transmitter consisted of $N_{\mathrm{ch}} = 15$ WDM channels modulated at a symbol rate of $R$ Gbaud with channel spacing $\Delta f$ and bandwidth efficiency $R/\Delta f = 0.56$ (except where otherwise stated). We considered differential phase encoding/decoding for employed all modulation formats. Before multiplexing, each channel was modulated with independent random symbols and filtered by a 2nd order super-Gaussian filter of bandwidth $\Delta f$. The number of symbols was a function of $R$, and sized to correctly reproduce the maximum walk-off cumulated by the edge channels of the WDM comb with respect to the central channel. In all cases it was confined between

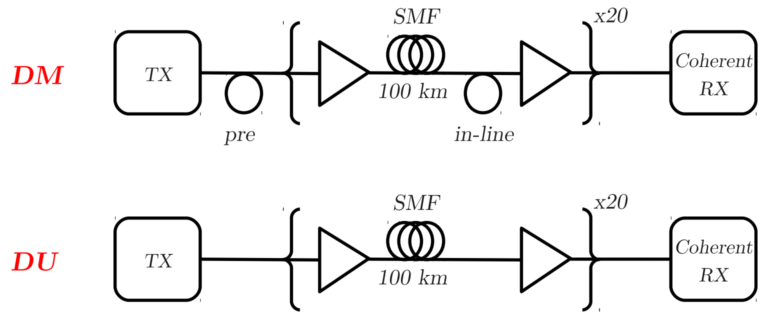


Figure 2.4: Block diagrams of DM and DU simulated links. TX block represents 15 WDM channels with spacing $\Delta f = R/\eta$, with $R$ the symbol rate and $\eta$ the bandwidth efficiency. In DM we use a residual dispersion per span of 30 ps/nm and a pre-compensation of $-540$ ps/nm.

1024 and 16384 symbols. The transmission line consisted of 20 spans of 100 km of single mode fiber (SMF) with dispersion $D = 17$ ps/nm/km, slope parameter $\beta_3 = 0$, attenuation $\alpha = 0.2$ dB/km, nonlinear coefficient $\gamma = 1.3\,W^{-1}\mathrm{k}m^{-1}$. At the end of the link, lumped electronic dispersion compensation was implemented in the coherent receiver. In the DM case, we used an in-line residual dispersion per span (RDPS) of 30 ps/nm and a straight-line rule pre-compensation of -540 ps/nm [67, 68], while in the DU case no compensation fiber was used.

Propagation used the SSFA described in sec. 1.5 with zero PMD and Manakov nonlinear step [11, 18]. The maximum nonlinear phase per step was $3 \cdot 10^{-3}$ [rad] at symbol rates lower than 80 Gbaud, and $6 \cdot 10^{-4}$ [rad] at higher symbol rates. Reasons for this choice will be provided in App. B.2. Before detection we used a 6-*th* order Butterworth optical filter of bandwidth $\Delta f$ to extract the central WDM channel. The signal was then converted to an electrical current by an ideal coherent mixer (zero laser phase noise, zero frequency offset) and sampled at 2 samples per symbol over a bandwidth of $0.6R$. The discrete sequence was then digitally processed by performing polarization recovery and equalization with a 15-tap least-squares trained equalizer, and a 27-tap carrier phase estimation (CPE). For PSK modulation formats the CPE was performed with the Viterbi&Viterbi algorithm [41], while for 16QAM we used a blind-phase search algorithm [69].

Each BER estimation, with ASE either generated at each amplifier (distributed noise) or loaded at the end of the link (noise loading), was performed by Monte Carlo simulations, by repeating several independent transmissions of the WDM comb until 100 errors were counted. The input states of polarization (SOP) of the WDM carriers were randomly oriented over the Poincaré sphere and randomly scrambled between two consecutive transmissions.

In figures of Sec. 2.2.2 and 2.2.3, in each labeled curve only the indicated nonlinearities are active, except for the curve labeled WDM, where all nonlinear effects are active and thus corresponds to the real case. The nonlinearities, introduced in Sec. 1.4, are: self-phase modulation (SPM), cross-phase modulation (XPM) and cross-polarization modulation (XPolM). The curves labeled WDM in all numerical results were obtained with the unique-field propagation (UFP) and therefore do account for linear channel crosstalk as well as all nonlinear effects, including Four Wave Mixing (FWM) and nonlinear spectral broadening, which are instead not captured by the separate-field propagation (SFP) solutions for the SPM, XPM, XPolM and XCI labeled simulations. The XCI curve includes both XPM and XPolM and is added only for results concerning DU links, to better compare cross-channel to single-channel nonlinearity. Recall that a small NLT corresponds to a strong nonlinearity, and vice-versa.

### 2.2.2 NLT in DM systems

The six plots in Fig. 2.5 shows the normalized margin versus symbol rate in the DM link. We report both the simplistic case of noise loading at the receiver (solid lines) and the true case of distributed ASE noise (dashed lines), where nonlinear signal-ASE interactions are fully accounted for. We will call *noiseless* SPM/XPM/XPolM those effects with ASE noise loading, and *noisy* SPM/XPM/XPolM those effects with distributed ASE. The six DM plots in Fig. 2.5 show that:

**1)** XPM with PSK formats is quite sensitive to signal-ASE interactions, since noisy and noiseless XPM-NLTs widely differ. This is due to the fact that in noiseless propagation in a DM link the symbol-rate quasi-periodic intensities of the transmitted

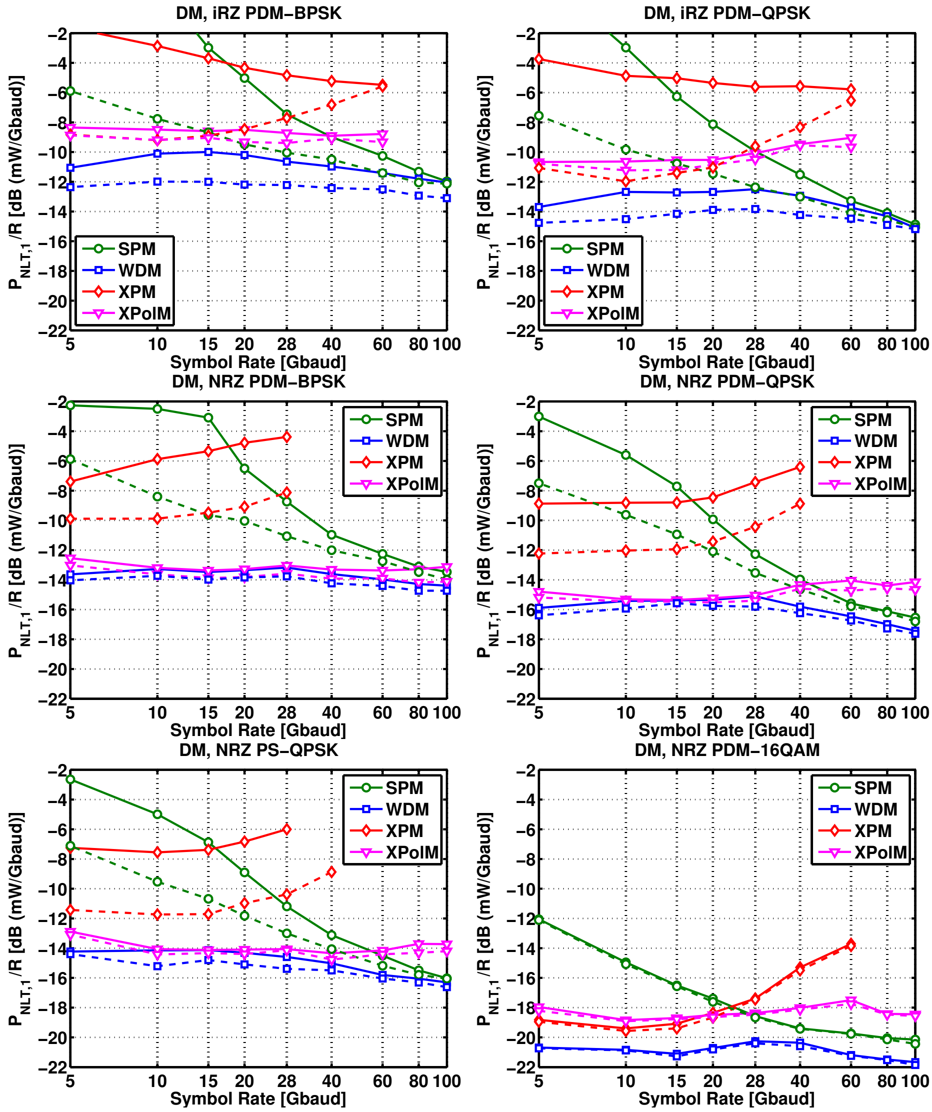


Figure 2.5: Normalized margin $10\log_{10}(P_{\mathrm{NLT},1}/R)$ vs symbol rate $R$ [log scale] for a 15-channel homogeneous WDM comb with $R/\Delta f = 0.56$ over a 20x100km SMF DM link with RDPS = 30 ps/nm. Solid lines: ASE loading. Dashed lines: distributed ASE. Labels: "SPM"=self-phase modulation only; "XPM"=scalar cross-phase modulation only; "XPolM"=cross-polarization modulation only; "WDM"=all nonlinearities.

PSK WDM fields are preserved, hence also the noiseless XPM field is symbol-rate quasi-periodic, and is thus well suppressed by the CPE block [70]. When in-line ASE is added, the intensity periodicity is broken and the XPM suppression in the CPE is way less effective, leading to a stronger residual XPM and thus a lower NLT.

Also, the low-pass filtering action introduced by fiber walk-off on the field intensity enhances XPM correlation on neighboring symbols, thus helping the phase tracking of the CPE [70]. In this case, shortening the CPE window may further improve the BER, however at the expense of an increase of the ASE-induced phase noise [71]. Since at the 1-dB NLT the ASE noise is largely dominant, then the number of CPE taps is generally set to a value large-enough (27 in our simulations) to well suppress ASE-induced phase fluctuations, and small-enough not to interfere with the frequency estimation block.

**2)** With homogeneous NRZ-PSK WDM formats, the dominant nonlinearity in the lower symbol rate range is XPolM. According to the XPolM model in [18], this is attributed to the fast modulation-induced random re-orientations of the mean WDM Stokes vector (called the *pivot*) in Stokes space, which cannot be tracked by the polarization demultiplexer at the receiver. ASE impact on XPolM is negligible, since ASE contribution to the pivot re-orientation is a second-order effect [72].

The XPolM-NLT is seen to be almost symbol-rate independent for all formats over the shown range. As symbol rate $R$ is increased at constant $\eta = R/\Delta f$, the channel spacing $\Delta f$ is also increased, and so is channel walk-off. XPM rapidly weakens and thus XPM-NLT increases[1] with $R$ because of the low-pass intensity filtering due to walk-off [70]. XPolM is affected also by channels with a huge spectral distance from the channel of interest [73,74] since it weakens with walk-off (i.e., with $R$) much more slowly than XPM [75]. Hence the XPolM-NLT on the shown $R$ range appears almost flat.

**3)** iRZ pulses in PDM-PSK formats were proposed because they quite effectively suppress XPolM [64–66]. The XPolM-NLT curves in a homogeneous WDM system with iRZ-PDM-PSK formats are in fact seen to shift to a much higher value than with NRZ pulses. The physical reason can be attributed to the symbol-time periodic

[1]For the anomalous behavior of noiseless XPM-NLT for iRZ see point **4)**.

alternation of each signal's polarization between two antipodal SOPs, which partly compensates the XPolM polarization spread [64].

**4)** PSK formats with iRZ pulses are also less sensitive to XPM in the low symbol rate regime due to a more constant transmitted PDM power than with NRZ pulses. For instance, a phase jump by $\pi$ from one symbol to the next one both in X and Y polarizations makes an NRZ-PSK format cross the zero PDM intensity, while such a zero crossing does not occur for iRZ format because at each phase transition on X polarization the Y polarization has constant power, and vice-versa. This translates into weaker intensity-dependent scalar nonlinearities, namely, SPM and XPM [76].

In the DM link, for iRZ pulses we note an anomalous decrease with $R$ of the initially very high noiseless XPM-NLT. This is ascribed to an increase with $R$ of the XPM-enhancing intensity fluctuations induced by GVD on the high-power, narrow RZ pulses, which overcome the XPM-suppressing action of walk-off. However, the realistic noisy XPM-NLTs are much lower and do recover an increasing behavior with $R$.

**5)** with iRZ-PDM-PSK formats, the suppression of XPolM makes noisy XPM no longer negligible. Also, the noise loading method may be quite inaccurate on the overall WDM-NLT, with a maximum difference of 2dB between noisy and noiseless propagation over the shown $R$ range.

**6)** SPM-NLT decreases at increasing symbol rates for all formats. SPM manifests at small $R$ as a quasi-periodic nonlinear phase rotation induced by the quasi-periodic signal intensity, which is effectively suppressed by the CPE block, hence SPM-NLT is initially large. As the symbol rate increases, the SPM perturbation field evolves into an almost signal-independent circular noise, much like in DU links [48], and the CPE block cannot suppress the intensity fluctuations of the SPM field. Noisy SPM manifests at small $R$ as an ASE-induced nonlinear phase noise: ASE destroys the noiseless-SPM periodicity so that the CPE is less effective in suppressing SPM (Cfr point **1** for XPM). With standard single mode fibers, the nonlinear signal-noise interactions are limited to bandwidths not exceeding a few tens of GHz [28], so that, as the symbol-rate increases, a larger and larger fraction of ASE is white as in linear propagation and thus noise loading simulations yield the same results as with distributed noise. Hence noisy and noiseless SPM-NLTs merge at large $R$.

In NRZ-BPSK, the plateau at which the noiseless SPM-NLT levels-off as $R$ is decreased is due to the large powers tolerated by this format. Such powers cause a strong SPM-induced bandwidth enlargement, and a significant fraction of the signal energy is cut off by the receiver filter.

**7)** the WDM-NLT of PS-QPSK is almost 1 dB higher than that of NRZ-PDM-QPSK, but still lower than that of NRZ-PDM-BPSK. With only XPolM, PS-QPSK has a NLT very close to PDM-BPSK, since they both display a binary SOP pattern on the Poincaré sphere (Cfr point **3)**). With only XPM, PS-QPSK has a similar NLT to PDM-QPSK since they both have a similar intensity profile and the same quaternary modulation in each polarization.

**8)** signal-noise interactions in 16QAM are a second-order effect, not only for XPolM, but also for scalar SPM and XPM, because of the large modulation-induced intensity fluctuations of the QAM format. The same reason makes XPM dominant at low symbol rates.

**9)** the decreasing trend of SPM with $R$ leads to a cross point where single-channel NLI dominates over cross-channel NLI. This point, for a 15-channel transmission, is placed around 100Gbaud for NRZ-BPSK format, while it lies in the range 40-60Gbaud for the other NRZ formats; it is around 25Gbaud for the iRZ-PSK formats.

**10)** The overall WDM-NLT is roughly constant (within 1.5dB) over the shown $R$ range, with a very shallow maximum in the range 10-28Gbaud. It is worth noting that the influence of signal-ASE interactions on the WDM-NLT is significant only for the iRZ-PSK formats. This implies that NLT at BER=$10^{-3}$ in WDM DM links for all other formats can be assessed by using the faster noise-loading method.

### 2.2.3 NLT in DU systems

The six plots of Fig. 2.6 show the corresponding normalized margin $P_{\mathrm{NLT},1}/R$ versus symbol rate in the DU link. Here ASE was loaded at the receiver since signal-noise interactions are negligible [72]. These six plots show that:

**11)** with all formats, scalar XPM slightly dominates over XPolM. We ascribe this behavior to the GVD-induced pseudo-random intensity variations in DU links, which

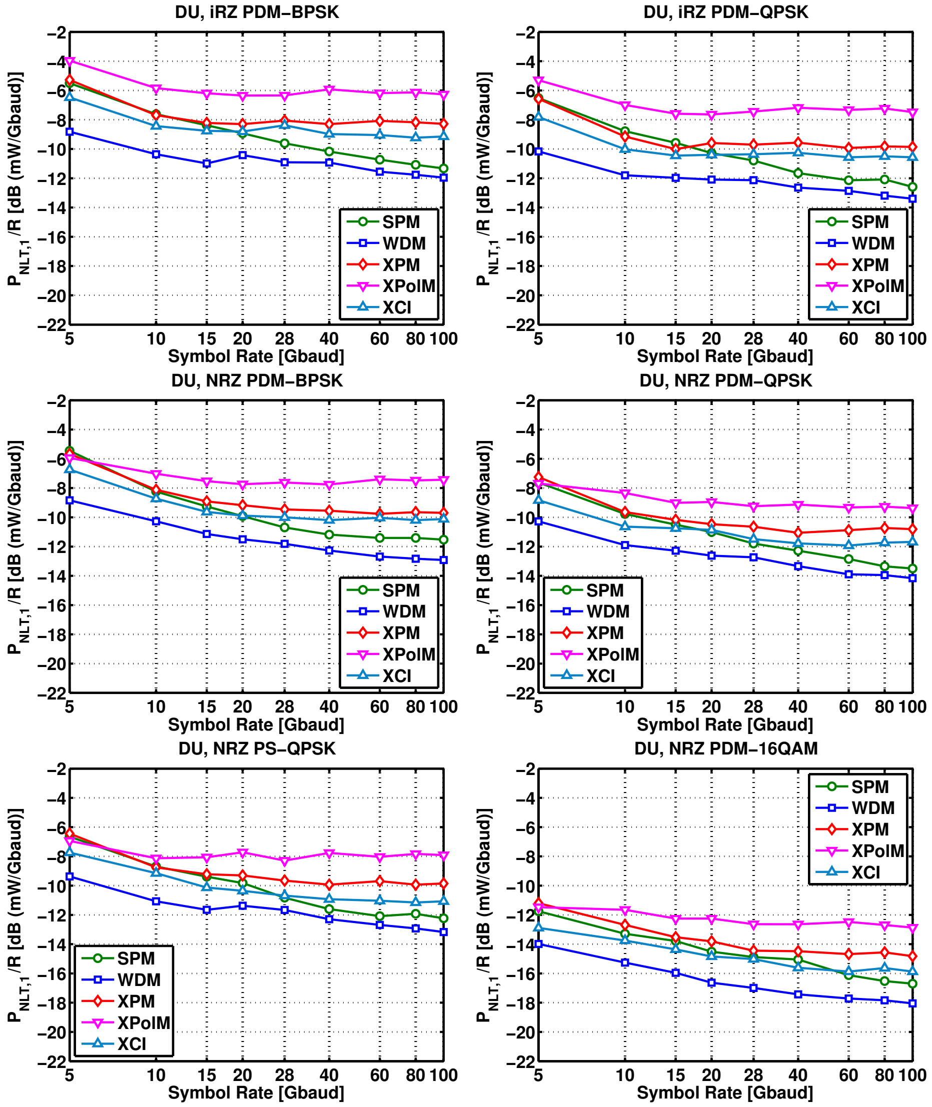


Figure 2.6: Normalized margin $10\log_{10}(P_{\mathrm{NLT},1}/R)$ vs symbol rate $R$ [log scale] for a 15-channel homogeneous WDM comb with $R/\Delta f = 0.56$ over a 20x100km SMF DU link. ASE loading. Labels: “SPM”=self-phase modulation only; “XPM”=scalar cross-phase modulation only; “XPolM”=cross-polarization modulation only; “WDM”=all nonlinearities; “XCI”=XPM+XPolM.

both tend to enhance XPM variance and weaken XPM-induced symbol correlations observed in the DM case, thus reducing the CPE XPM suppression effectiveness. However, both XPM and XPolM are much better suppressed than in DM links by the huge walk-off of the line;

**12)** in the presented 15-channel WDM system, the XPM- and XPolM-NLTs are quite close, hence to better compare cross- and single-channel NLTs it is useful to also visualize the XCI-NLT where both effects are active. We see that at lower $R$ XCI dominates, while at larger $R$ SPM dominates. The cross-point is around 20Gbaud for most formats, except for 16QAM where it is around 40-60Gbaud. When nonlinearity $i$ is active, with $i \in \{\text{XPM,XPolM}\}$, for all links where it manifests as a circular complex Gaussian field, the NLT $P_{\text{NLT,1-}i}$ is related to the corresponding nonlinear interference coefficient $a_i$ as [55, p. B211]:

$$P^2_{\text{NLT,1-}i} = 1/(3c^2 S_0 a_i) \tag{2.12}$$

where $c = 1.27$ for the NLT at 1dB penalty, and $S_0$ is the electrical SNR at the reference BER. If XPM and XPolM fields were independent, their variances would add up, hence $a_{\text{XCI}} \cong a_{\text{XPM}} + a_{\text{XPolM}}$ and we would get an XCI-NLT given by:

$$P_{\text{NLT,1-XCI}} = 1/\sqrt{1/P^2_{\text{NLT,1-XPM}} + 1/P^2_{\text{NLT,1-XPolM}}}. \tag{2.13}$$

We verified that such a formula is very close (within 0.3dB for all formats, except for PS-QPSK where the error can be up to 1dB) to the XCI simulated NLT in Fig. 2.6, hence we conclude that XPM and XPolM are practically uncorrelated in DU links for most formats. We also verified that (2.13) very accurately predicts the XCI-NLT even in DM links.

**13)** the GN reference formula [52, eq. (1)] treats the cross-nonlinearity as a whole, and cannot discriminate between XPM and XPolM. At large $R$, the XCI coefficient is predicted to scale as [77, eq. (20)]: $a_{\text{XCI}} \propto 1/(R\Delta f) = \eta/R^2$ (consistent with [78, eq. (57)] for Nyquist-WDM systems). Using (2.12) we have $P_{\text{NLT,1-XCI}} \propto 1/\sqrt{a_{\text{XCI}}}$, hence the GN model leads us to conclude that $P_{\text{NLT,1-XCI}}/R$ is a constant at large $R$ and constant $\eta$. This confirms the observed flattening of the simulated XCI-NLTs versus $R$ for all formats.

**14)** XPolM is much weaker than in the DM case. In both DM and DU links, XPolM operates in the same way within each span, with an efficiency related to the number of pivot re-orientations along propagation [18]. The difference is that the periodic map of the DM link induces similar XPolM scattering between input/output of each span, thus forcing a long trajectory of the output SOP vector over the Poincaré sphere [75, p. 97]. In the DU link, instead, the direction of the pivot is almost uncorrelated span by span, thus producing brownian-like wiggled output SOP trajectories that drift less away from the input SOP [18]. Since each symbol experiences a different SOP trajectory, overall we observe polarization scattering, and since the drift from the input SOP is smaller in DU links, then we also have a weaker XPolM.

**15)** with all formats, the overall WDM $P_{\mathrm{NLT},1}/R$ curves decrease with $R$, i.e., the best symbol rate at fixed bandwidth efficiency and with 15 channels is below the lower $R$ limit, i.e., 5Gbaud.

### 2.2.4 NLT at large bandwidth efficiency

Fig. 2.7 shows the normalized margin $P_{\mathrm{NLT},1}/R$ versus symbol rate either for NRZ PDM-QPSK (left) and signal NRZ PDM-16QAM (right) in both a DM link (top) and a DU link (bottom), with a larger bandwidth efficiency $\eta = 0.85$ w.r.t. that of Figs. 2.5–2.6. Comparing the figure with the corresponding plots in Figs. 2.5–2.6, i.e., (center, right) and (bottom, right), we observe that an increase of $\eta$ (hence a reduction of the spacing $\Delta f$ among channels at same $R$ and thus a decrease of the walk-off) enhances cross-channel effects, although in different ways in DM and DU.

We start analyzing the PDM-QPSK format. In the DM link, XPolM remains the dominant nonlinearity, with a NLT decrease from the $\eta = 0.56$ case by about 1-1.5 dB at symbol rates $R > 28$Gbaud. Overall, the WDM curve shifts downwards by more than 1 dB, and keeps its concave shape with a shallow maximum in the 10-40Gbaud range, with signal-ASE interactions still negligible over the whole symbol rate range.

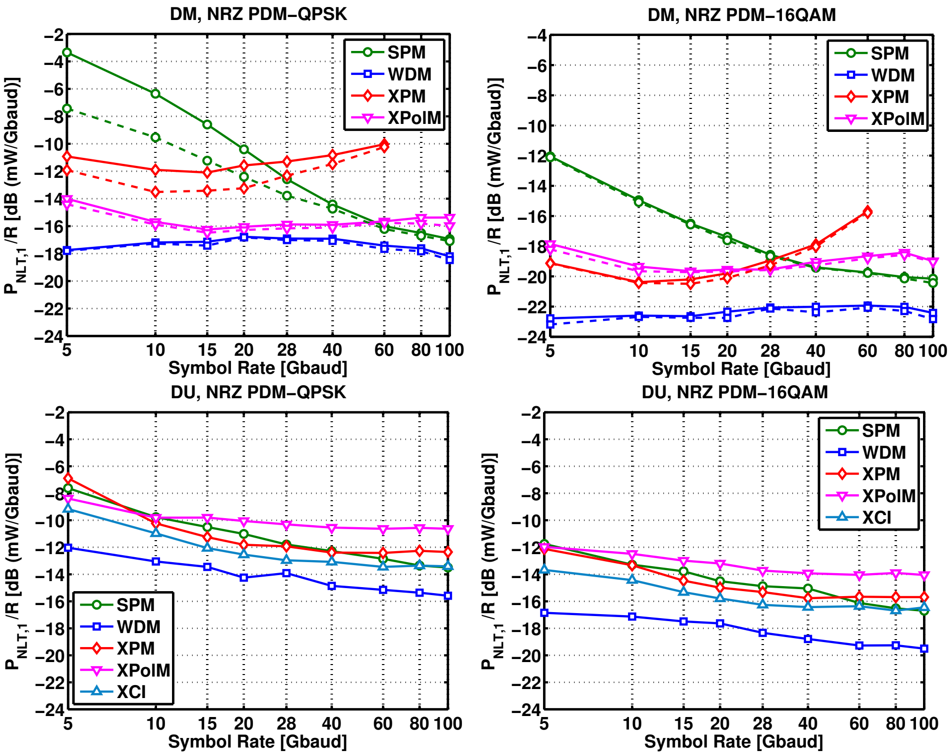


Figure 2.7: Normalized margin $P_{\mathrm{NLT},1}/R$ at reference BER=$10^{-3}$ vs symbol rate for 15-channel NRZ-PDM-QPSK (left) and NRZ-PDM-16QAM (right) homogeneous WDM comb with $\eta = 0.85$ over a 20x100km SMF link. (Top) DM link, (Bottom) DU link.

In the DU link, for QPSK format, we observe that both XPM- and XPolM-NLTs at large rates decrease by $\sim$1 dB when increasing $\eta$ by $\sim$2 dB, and so does the WDM-NLT curve. This is consistent with the GN model large-rate prediction that $a_{\mathrm{XCI}} \propto 1/(R\Delta f) = \eta/R^2$ [77]. Using (2.12) we have $P^{\mathrm{dB}}_{\mathrm{NLT,1\text{-}XCI}} \propto -\frac{1}{2}\eta^{\mathrm{dB}}$. Now the XPM-NLT is quite close to the SPM-NLT, while the XPolM-NLT is almost 2 dB larger than the XPM-NLT around 60 Gbaud. At lower rates, FWM due to next neighbors starts to be noticeable even in DU links. For instance, at $R = 5$Gbaud the XPM- and XPolM-NLTs, obtained by the separate-field propagation at $\eta = 0.85$, vary very little w.r.t. the $\eta = 0.56$ case; however the WDM-NLT (obtained by the unique-field propagation, that captures FWM) decreases by 1.8 dB. To double-check that next-neighbor FWM is indeed the cause, in a separate test of the same DU link at $R = 5$Gbaud we suppressed the 2 next neighbors of the central channel in the 15 channel comb and verified that the WDM-NLT (obtained by the UFP) now coincides (to within 0.5dB) with the SFP-obtained NLT with XPM+XPolM+SPM activated.

For the 16QAM format we observe similar NLTs reductions noted for the QPSK format, with the exception of the noiseless XPM-NLT in the DM link, which decreases for QPSK by about 2 dB more than the respective NLT for 16QAM. We ascribe this behavior to the different dynamics of XPM between 16QAM and QPSK in absence of ASE-signal interaction.

### 2.2.5 Checks of NLT behavior with GN theory

NLT simulations are extremely time-consuming hence the number of WDM channels was here limited to 15. It would be very useful to have an analytical model to cross-verify the general physical trends found by simulation and to extrapolate the behavior when the WDM channel count is in the order of 100, as in typical commercial WDM systems. Fortunately, for DU systems the GN model [52] does the job. In this section, we test the predictions of the GN model and compare them with those obtained from the NLT data in Figs. 2.5–2.6.

If the NLI coefficient $a_i$ for the $i$-th nonlinearity were format-independent, as postulated by the GN model, from (2.12) we would expect $\sqrt{S_0}P_{\mathrm{NLT,1}\text{-}i}$ to also be format-independent. If this were the case, the simulated $P_{\mathrm{NLT,1}\text{-}i}/R$ versus $R$ curves reported in Fig. 2.6 for the DU link, multiplied by the $\sqrt{S_0}$ of the used format, should all fall on top of each other. We performed this check in Fig. 2.8(a-b). We used the same $S_0$ values as in the simulations, namely, $S_0 = [7.77, 10.77, 17.07]$ dB for the three NRZ formats (BPSK, QPSK, 16QAM in that order), while iRZ formats required 0.2dB less SNR than their corresponding NRZ formats. Dashed curves refer to XPolM, dash-dotted to XPM, and solid lines to the all-nonlinearities WDM case.

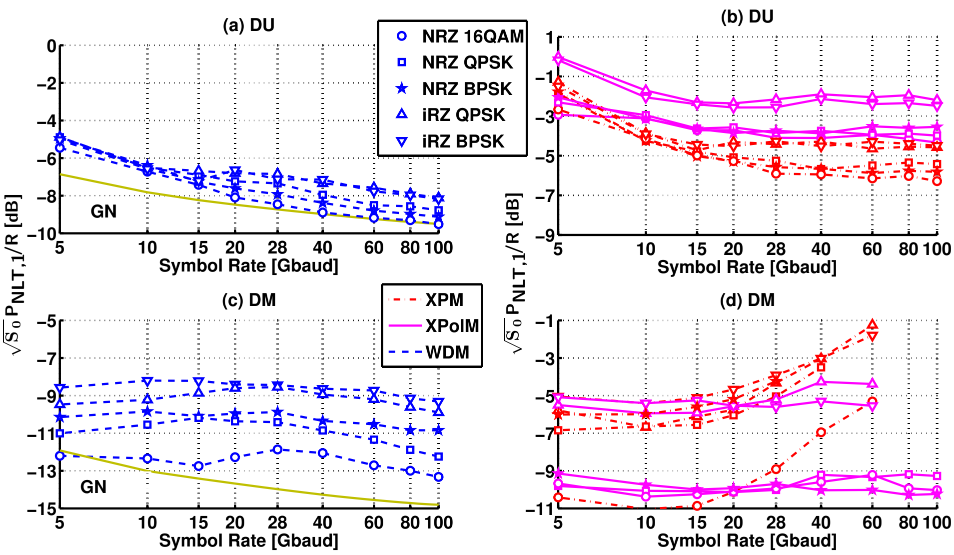


Figure 2.8: Symbols: simulated $\sqrt{S_0}P_{\mathrm{NLT},1}/R$ [dB(mW/Gbaud)] versus $R$ [Gbaud] for DU (a-b) and DM (c-d) 20x100km SMF link with 15 WDM channels at $\eta = 0.56$. $P_{\mathrm{NLT},1}/R$ simulated values are those reported in Figs. 2.5–2.6. GN theory plotted in solid line as $\sqrt{S_0}P_{\mathrm{NLT},1}/R = 1/\sqrt{3c^2 a_{\mathrm{NL}}}/R$, which should be format-independent.

We note that the $\sqrt{S_0}P_{\mathrm{NLT},1}/R$ XPM/XPolM curves reasonably merge, with different values between NRZ and iRZ curves. The WDM curves with NRZ pulses show instead a spread of about 1.5dB across the formats at larger $R$, which we verified to be also due to the spread of the corresponding SPM curves (not shown to avoid clogging the figure). As pointed out in Sec. 2.1.2, $P_{\mathrm{NLT}}/R$ is proportional to the reach, hence one can read off the WDM decreasing curves in Fig. 2.8(a) also the reach behavior versus symbol rate for each modulation format. A similar decreasing reach vs. $R$ behavior for a 5-channel WDM DU system was reported in [62].

We also report in Fig. 2.8(a-c) the solid-line curve labeled GN, which plots the theoretically equivalent quantity $1/\sqrt{3c^2 a_{\mathrm{NL}} R^2}$, as per eq. (2.12), with $a_{\mathrm{NL}} = \int G_{\mathrm{NLI}}(f)\mathrm{d}f$ calculated according to the GN reference formula [52, eq. (1)] with the efficient algorithm detailed in [79], with the same normalized WDM input power spectral density $G_{\mathrm{NLI}}$ used in the simulations of Fig. 2.6. The GN curve theoretically confirms the monotone decreasing trend versus $R$ in DU links with $\eta = 0.56$ and 15 channels. Among the NRZ formats, we note that the 16QAM format is the closest to the GN curve, because of its better agreement with the key GN model assumption of a Gaussian input process [80]. Such findings about a non-negligible modulation-format dependence of the NLI variance are in agreement with recent extensions of the GN model [81–84]. Also, note that iRZ formats are the farthest from the GN model, with a distance of over 2dB, and their weaker NLI may suggest that iRZ formats do not converge to a Gaussian-like field as fast as their NRZ counterparts.

In Fig. 2.8(c-d) we tried the same rescaling with the DM link data in Fig. 2.5, where we used the distributed-noise NLTs only. We see that the XPM curves now show a more marked format dependence (note the dramatic decrease of the 16QAM XPM), while the XPolM curves for the NRZ formats do reasonably merge together and are relatively flat, as in the DU link. This is an indication that XPolM is the nonlinear effect that behaves most similarly in DM and DU links. The large spread of WDM curves across the modulation formats now comes not only from a large spread of the SPM curves (not reported to avoid confusion), but also from that of the XPM curves. For completeness, the curve labeled GN reports the predictions of the GN model for the DM link, which are clearly unsatisfactory. A recent extended GN model seems promising in reducing such a spread also for DM links [85].

To conclude this section, we present GN model extrapolations of the $\sqrt{S_0}P_{\mathrm{NLT},1}/R$ versus $R$ curves to large WDM channel numbers, where simulations are unfeasible. We first note that the monotonically decreasing behavior versus $R$ of both GN theory and simulations observed in Fig. 2.8(a) can in part be explained by the increase of the total WDM bandwidth, since the number of channels was fixed at a fixed $\eta = R/\Delta f = 0.56$, hence by increasing $R$ we also increased the spacing $\Delta f$. In Fig. 2.9 we now more realistically fixed the WDM aggregate symbol rate $B = N_{\mathrm{ch}}R = 2.8$ Tbaud, which corresponds to increasing the number of channels $N_{\mathrm{ch}}$ when reducing $R$. The total occupied bandwidth was fixed at $W = B/\eta$. We observe that even in this case the $\sqrt{S_0}P_{\mathrm{NLT},1}/R$ (hence the reach!) plot versus $R$ keeps a monotonically decreasing trend at any $\eta < 1$, and converges to the naturally flat behavior of the Nyquist-WDM case ($\eta = 1$). The physical message from these curves is that, for

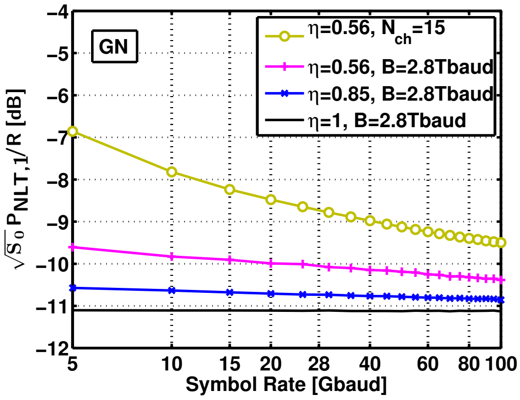


Figure 2.9: Plot of $\sqrt{S_0}P_{\mathrm{NLT},1}/R$ versus $R$ for DU 20x100km SMF link obtained from GN reference formula at varying bandwidth efficiency $\eta$ and fixed aggregate symbol rate $B = N_{\mathrm{ch}}R$. The $\eta = 0.56$, $N_{\mathrm{ch}} = 15$ curve is the same labeled GN in Fig. 2.8(a-c). These curves display the same behavior versus $R$ as the reach.

reach maximization, smaller symbol rates are slightly preferable, which means the WDM energy should be most uniformly spread across the frequency axis. However, from the "reach" curves of Fig. 2.9, we conclude that the granularity of the WDM spectrum may be of some importance only away from the Nyquist limit, and at small channel counts, where smaller symbol rates are slightly preferable. Otherwise, in all practical cases, the per-channel symbol rate need not be optimized to maximize the reach.

We cannot replicate such curves by SSFM simulation. However, to give a feeling to the reader, for NRZ QPSK and 16QAM we replicated at $R$=10 and 80 Gbaud the WDM NLTs of Figs. 2.5–2.6,2.7 at the aggregate bandwidth $B_{\mathrm{A}}$ of the 15-channel system at 28 Gbaud namely, $B_{\mathrm{A}} = 750$ GHz at $\eta = 0.56$ and $B_{\mathrm{A}} = 500$ GHz at $\eta = 0.85$, as reported in Table 2.1. We thus used 43 channels at 10 Gbaud and 5 channels at 80 Gbaud. The results are reported in Fig. 2.10. Solid lines are the WDM NLTs in the respective cases of Figs. 2.5–2.6,2.7. Red dashed lines indicate the results at a fixed aggregate bandwidth. We note a counterclockwise tilt of each curve as expected, since we are enhancing cross-nonlinearities at 10 Gbaud while reducing them at 80

| $R$ [Gbaud] | $N_{ch}$ | $\eta = 0.56$ | | $\eta = 0.85$ | |
|---|---|---|---|---|---|
| | | $\Delta f$ [GHz] | $B_A$ [GHz] | $\Delta f$ [GHz] | $B_A$ [GHz] |
| 10 | 43 | 17.86 | 768 | 11.77 | 506 |
| 28 | 15 | 50.00 | 750 | 32.94 | 494 |
| 80 | 5 | 142.86 | 714 | 94.12 | 471 |

Table 2.1: Values at fixed aggregate bandwidth $B_A$.

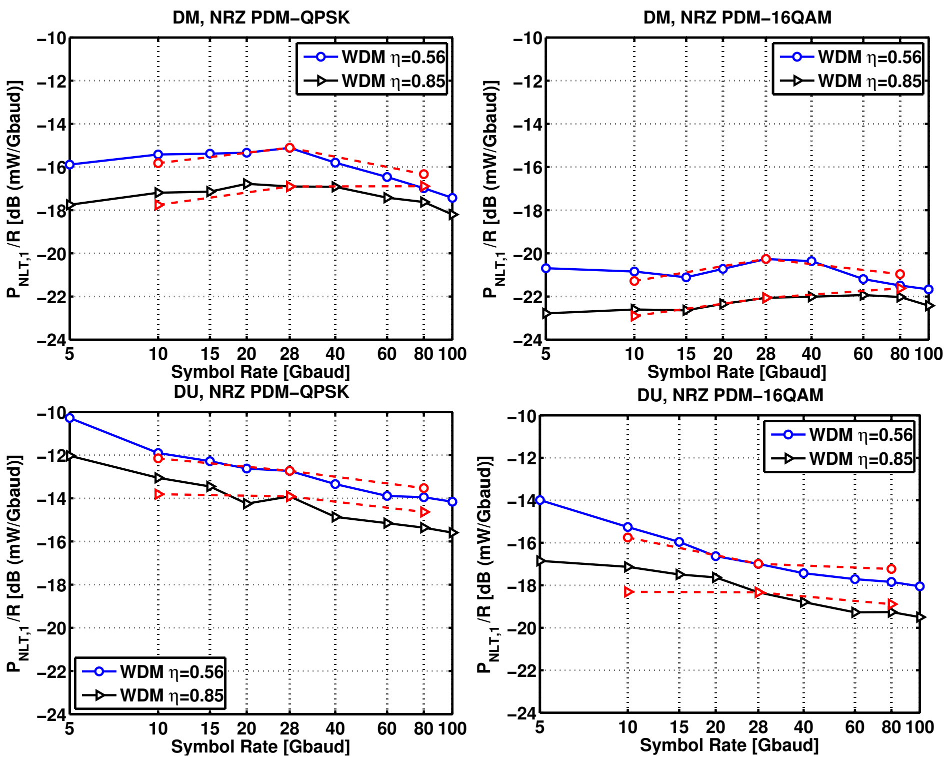


Figure 2.10: Normalized margin $P_{NLT,1}/R$ at reference BER=$10^{-3}$ vs symbol rate over a 20x100km SMF link. (Solid) Fixed number of channels, (Dashed) Fixed aggregate bandwidth. (Top) DM link, (Bottom) DU link. (Left) NRZ PDM-QPSK, (Right) NRZ PDM-16QAM.

| | 10 Gbaud | 28 Gbaud | 80 Gbaud |
|---|---|---|---|
| DM, 16QAM | −2.90 | −2.32 | −3.55 |
| DU, 16QAM | 2.68 | 0.95 | −0.38 |
| DM, QPSK | 9.17 | 9.48 | 7.60 |
| DU, QPSK | 12.69 | 11.86 | 10.35 |
| DM, BPSK | 14.42 | 14.52 | 13.41 |
| DU, BPSK | 17.42 | 15.88 | 14.83 |

Table 2.2: Amplifiers Noise figure $F_{\text{NLT},1}$ [dB] yielding 1-dB of SNR penalty at BER=$10^{-3}$ and power $P_{\text{NLT},1}$ estimated in Figs. 2.5–2.6. $20 \times 100$km SMF link with 15 WDM channels. ASE noise loaded at the receiver.

Gbaud. Regarding the DU case we verified that, such results agree with the GN model also used for the predictions of Fig. 2.9. Overall we observe that the tested symbol rates change by at most 0.76 dB in the DM case (80Gbaud QPSK at $\eta = 0.85$) and 1.18 dB in the DU case (10Gbaud QAM at $\eta = 0.85$).

## 2.3 Reach prediction from NLT

The procedure to obtain the 1-dB NLT $P_{\text{NLT},1}$ consists of tuning the lumped amplifiers noise figure together with the launch power in order to achieve the combination such that at an SNR $S = S_0 + 1$ dB the BER is $10^{-3}$, as described in Sec. 2.1.1. As a by-product of the $P_{\text{NLT},1}(N)$ estimation for an $N$-span link, we thus also get the corresponding amplifiers noise figure $F_{\text{NLT},1}(N)$. Table 2.2 reports the values of $F_{\text{NLT},1}$ at $N = 20$ spans found in the same WDM-NLT simulations yielding the $P_{\text{NLT},1}$ values used to build Figs. 2.5–2.6 for the three selected modulation formats at $R$=10, 28 and 80Gbaud in both DM and DU links, with receiver noise loading (since Fig. 2.5 indicates that, for NRZ pulses, signal-noise interactions are negligible). Whenever $F_{\text{NLT},1}$ takes unrealistic values, it means that a practical system is not able to achieve the target BER at the selected distance. For instance, PDM-16QAM achieved NLT at $N = 20$ spans with a negative noise figure, which physically means that 20 spans are unachievable with lumped amplification and does require distributed amplification.

Although such a noise figure may take unrealistic values, it can be used to estimate the system reach $N_0$ at a practical noise figure $F$ from NLT measurements over a much shorter $N$-span link. In fact, in Sec. 2.1.2 we expressed in (2.11) the reach $N_0$ as a function of the power budget at distance $N$. Eq. (2.11) is consistent with the reach predictions in [78] when $\varepsilon = 0$. Since by definition $N_0(F_{\mathrm{NLT}}(N)) = N$, then by multiplying and dividing the fraction on the right-hand side of (2.11) by $F_{\mathrm{NLT}}(N)$ and by using (2.3) and (2.4) we get:

$$\frac{N_0(F)}{N} = \left(\frac{F_{\mathrm{NLT}}(N)}{F}\right)^{\frac{2}{3+\varepsilon}} \tag{2.14}$$

which relates the maximum reach obtained at an arbitrary noise figure $F$ to the constrained noise figure $F_{\mathrm{NLT}}$ measured at $N$ spans. Eq. (2.14) is preferable to (2.11) since it does not involve the precise knowledge of the receiver electrical bandwidth, which is $R$ only with matched-filtering but in practice is set by the butterfly equalizer that also implements polarization demultiplexing. In our simulations, instead of estimating $F_{\mathrm{NLT}}$, we estimated $F_{\mathrm{NLT},1}$, i.e., the tweaked noise figure at the 1-dB $P_{\mathrm{NLT},1}$ at $N$ spans. The ratio $x_y \triangleq F_{\mathrm{NLT},1,y}/F_{\mathrm{NLT}}$ for a penalty of $y$ [dB] is plotted in [55, Fig. 6] as a function of $y$. At $y = 1$dB we read off the value $x_1 = 0.94$, hence substitution in (2.11) leads to

$$\frac{N_0(F)}{N} = \left(\frac{F_{\mathrm{NLT},1}(N)}{x_1 \cdot F}\right)^{\frac{2}{3+\varepsilon}} \tag{2.15}$$

It can be proved that in general $x_y = \frac{3}{2c(y)} - \frac{1}{2c(y)^3}$, where $c(y) = 1/\sqrt{3(1-10^{-y/10})}$ (hence $c(1) = 1.27$). The fitting factor $x_1$ is exact if the NLI is Gaussian additive, and may change for different NLI statistics. Of course one also needs the knowledge of $\varepsilon$ to perform such an estimation. $\varepsilon$ can be obtained either by theoretical models (e.g., the GN model predicts $\varepsilon \sim 0$ when cross-channel nonlinearity dominates [52, 77]) or by numerical fitting.

To double-check the accuracy of formula (2.15) for both our DU and DM links, we performed extra UFP SSFA simulations for three NRZ modulation formats (PDM-BPSK, PDM-QPSK and PDM-16QAM) at three sample symbol rates ($R$=10, 28 and

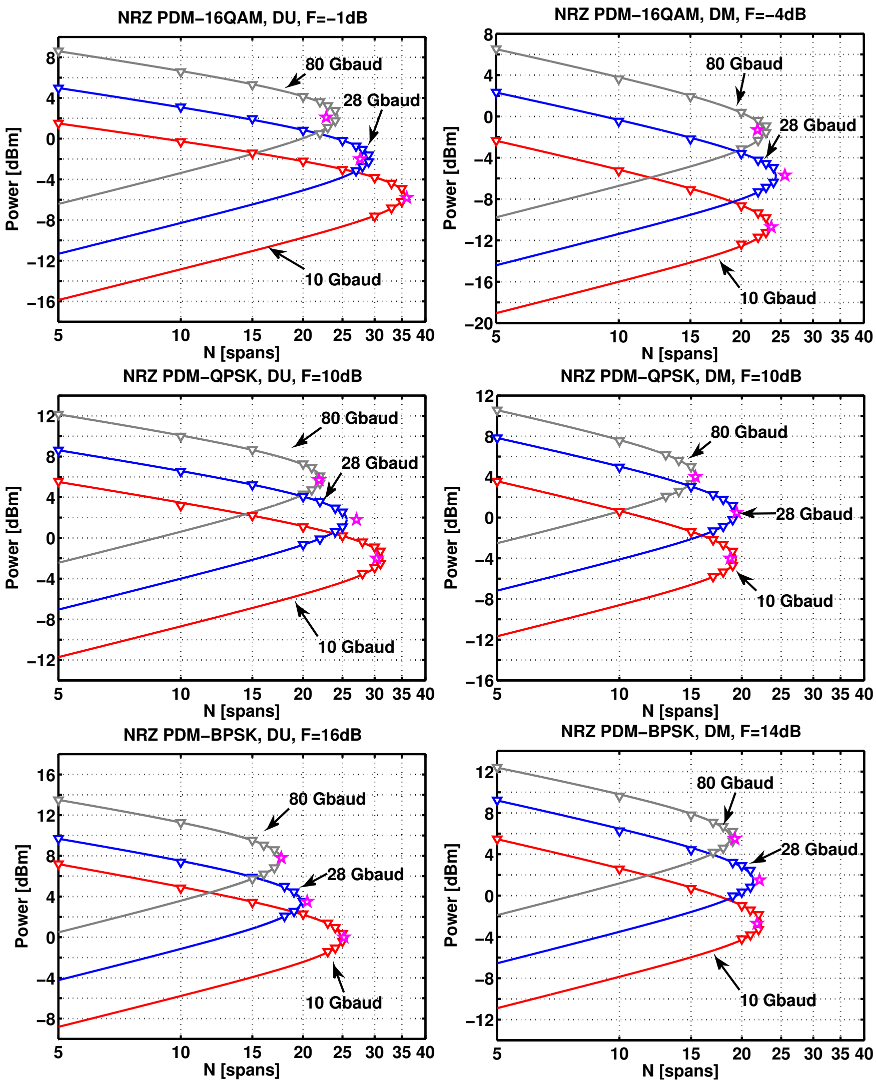


Figure 2.11: Triangles: SSFA-simulated power versus span number $N$ ($\log_{10}$ scale) at reference BER=$10^{-3}$ for a 15-channel homogeneous WDM comb with $\eta = R/\Delta f = 0.56$ over a $N$x100km SMF link. ASE noise loaded at the receiver. (Top row) PDM-16QAM, (Center row) PDM-QPSK, (Bottom row) PDM-BPSK. (Left column) DU link, (Right column) DM link. Stars: prediction by (2.15) using $F_{\mathrm{NLT},1}$ [dB] of Table 2.2. Solid lines: least-squares cubic interpolations of triangles.

| | 10 Gbaud | 28 Gbaud | 80 Gbaud |
|---|---|---|---|
| DM, 16QAM | 0.77 | 0.68 | 0.71 |
| DU, 16QAM | 0.11 | 0.19 | 0.19 |
| DM, QPSK | 0.80 | 0.72 | 0.71 |
| DU, QPSK | 0.29 | 0.27 | 0.26 |
| DM, BPSK | 0.77 | 0.70 | 0.56 |
| DU, BPSK | 0.40 | 0.30 | 0.29 |
| GN model [52, Sec. XI-D] | 0.072 | 0.013 | 0.005 |

Table 2.3: Estimated $\varepsilon$ from simulated data in Fig. 2.11.

80 Gbaud). Fig. 2.11 plots the power yielding the reference BER=$10^{-3}$ vs number of spans $N$ for a 15-channel homogeneous WDM comb for the selected modulation formats and symbol rates. The noise figure (reported in each plot) was arbitrarily set to keep the reach between 15 and 35 spans in all cases, and receiver noise loading was used. Triangles show simulations, while solid lines are least-squares cubic interpolations of the simulated values, according to (2.6), where $a_{\mathrm{NL}} = \alpha_{\mathrm{NL}} N^{1+\varepsilon}$, whose only fitting parameters are $\alpha_{\mathrm{NL}}$ and $\varepsilon$. From such "contours" the system reach can be easily visualized. The least-squares fitting NLI slope parameter values are reported in Table 2.3, along with the much smaller GN model predictions for $\varepsilon$ in DU links [52, Sec. XI-D]. For our 15-channel DM and DU links we note that $\varepsilon$ is very slowly decreasing with symbol rate $R$, and can be safely considered as a constant for reach maximization w.r.t. symbol rate, as discussed in Sec. 2.1.2. For DM links, it is $\varepsilon \sim 0.7$ for most formats, except for BPSK which has $\varepsilon \sim 0.6$; for DU links the value is much smaller, and close to 0.3, except for QAM which shows an $\varepsilon < 0.2$. For the PDM-QPSK DU link, the values in Table 2.3 are consistent with the experimental measurements in [46][2].

The predictions of formula (2.15) can now be calculated using Tables 2.3 and 2.2 and compared to the true reach. They are marked with stars in Fig. 2.11 and are

[2]For larger WDM systems where cross-nonlinearity dominates the NLT, it is theoretically predicted by the GN model that $\varepsilon \sim 0$ [52, 77].

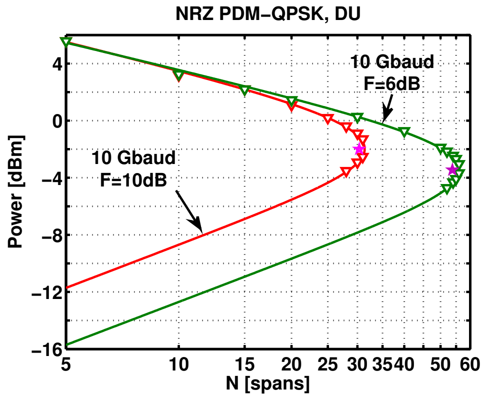


Figure 2.12: Triangles: SSFA-simulated power versus span number $N$ ($\log_{10}$ scale) at reference BER=$10^{-3}$ for a 15-channel PDM-QPSK homogeneous WDM comb with $\eta = R/\Delta f = 0.56$ over a $N$x100km SMF DU link. ASE noise loaded at the receiver. Stars: prediction by (2.15) using $F_{\mathrm{NLT},1}$ [dB] of Table 2.2. Solid lines: least-squares cubic interpolations of triangles.

quite close to the actual reach, with a worst-case error of only 1.5 spans. Although the used noise figures may be unrealistic and were used to keep the reach within 15 and 35 spans, yet the accuracy of (2.15) is not expected to degrade with practical noise figures and longer reaches, since such an equation is accurate for all physical links for which the NLT scales as $N^{-(1+\varepsilon)/2}$ as expressed in (2.9), with $\varepsilon$ a suitably estimated factor. For instance, we verified that a QPSK system in an SMF DU link with F=6 dB reaches 56 spans, as depicted in Fig. 2.12, while eq. (2.15) forecasts 53 spans.

As a sanity check, we replicated the 16QAM reach curve at 10 Gbaud for both DM and DU links and at $B_{\mathrm{A}} = 750$ GHz (43 channels). Such curves, are reported in Fig. 2.13 for both the DM and the DU case. We observe an almost rigid shift of the nonlinear asymptote with no implications on the accuracy of eq. (2.15). We believe that this conclusion holds for all the other investigated cases.

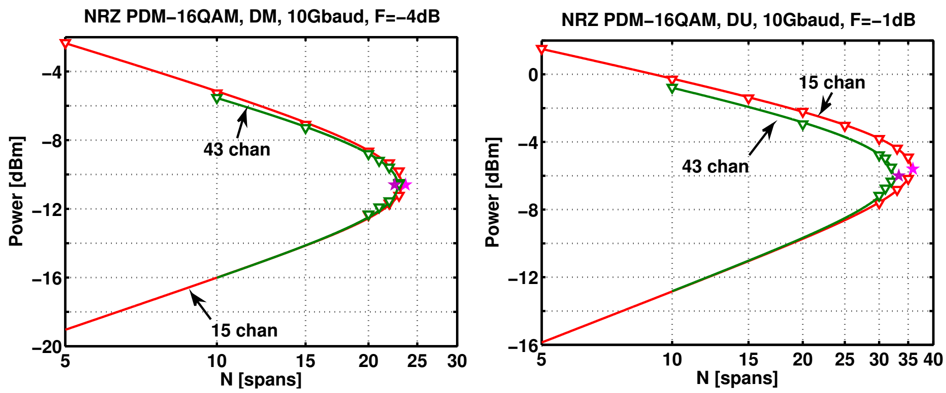


Figure 2.13: Triangles: SSFA-simulated power versus span number $N$ ($\log_{10}$ scale) at reference BER=$10^{-3}$. 10Gbaud 16QAM with bandwidth efficiency $\eta = R/\Delta f = 0.56$. $N$x100km SMF link with DM (left) and DU (right). ASE noise loaded at the receiver. Stars: prediction by (2.15) using $F_{\mathrm{NLT},1}$ of the corresponding WDM comb. Solid lines: least-squares cubic interpolations of triangles.

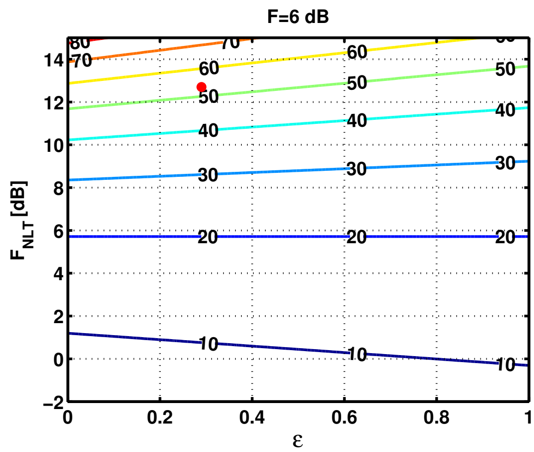


Figure 2.14: Contour levels of the maximum reach $N_0$ given in (2.15) vs. $\varepsilon$ and $F_{\mathrm{NLT},1}$ [dB] at a fixed $F = 6$ dB.

### 2.3.1 Tolerance to estimation errors

In Fig. 2.11 we showed that with the knowledge of $F_{\text{NLT},1}$ and $\varepsilon$ it is possible to estimate the system reach with a very good accuracy in a wide range of symbol rates and modulation formats. $F_{\text{NLT},1}$ comes for free from NLT simulations at a fixed distance $N$. Knowledge of the NLI slope parameter $\varepsilon$ must instead be acquired from several BER measurements versus distance, as those in Fig. 2.11, whose fit lead to Table 2.3. Alternatively, one can resort to analytical approximate expressions of $\varepsilon$. The general question is thus how accurately we need to know $\varepsilon$ and $F_{\text{NLT},1}$. To answer this question, in Fig. 2.14 we plotted the contour levels of $N_0(F_{\text{NLT},1}, \varepsilon)$ obtained from (2.15) vs. both $\varepsilon$ and $F_{\text{NLT},1}$. Such contours can be used to quantify the tolerance to our allowed ignorance of the two key parameters. We note that, according to (2.15), when the ratio $F_{\text{NLT},1}/(x_1 \cdot F) = 1$ the parameter $\varepsilon$ has no impact on reach prediction. For this reason, if the distance $N$ at which we measure the NLT is not too far from the true reach, a good estimation of $F_{\text{NLT},1}$ is more important than a good one of $\varepsilon$. This observation is also a direct consequence of the reach dependence on the term $3+\varepsilon$, where the term 3 smooths out the impact of the small $\varepsilon$. For instance, for a 10Gbaud DU-QPSK link we measured $F_{\text{NLT},1} = 12.7$ dB and $\varepsilon = 0.29$, yielding $N_0 = 53$ spans if $F = 6$ dB, see the filled circle in Fig. 2.14; an error of 1 dB on $F_{\text{NLT},1}$ gives a sizable error of 7 spans in reach, while a significant error of 22% on $\varepsilon$ yields an error of just 1 span in reach.

## 2.4 Conclusions

We provided massive simulation estimations of the nonlinear threshold $P_{\text{NLT}}$ versus channel symbol rate $R$ for both DM and DU coherent systems and for several modulation formats and supporting pulse shapes of interest. By using the nonlinearity-decoupling method, we provided the $P_{\text{NLT}}$ due to each individual Kerr effect, and showed that one can gather valuable information about the dominant nonlinear effect as the rate $R$ is varied. We provided plausible explanations of the observed behavior based on the main existing analytical nonlinear models. We proved that plots of $P_{\text{NLT}}/R$ also show the symbol rate that maximizes the transmission distance, i.e., the reach. For DU systems, for which the rather accurate analytical GN model is available, we compared our NLT simulations against theory and provided theoretical extrapolations to practically large WDM systems which cannot be simulated. We finally provided a new, simple system-level formula for reach prediction that uses the noise figure obtained at NLT, and we showed that such a formula can be used for accurate reach predictions both in DU and DM systems, provided that estimations of the NLI accumulation parameter $\varepsilon$ are available.

# Chapter 3

# Statistics of Nonlinear Interference in DU Networks.

In the previous chapter we recalled the recently proposed Gaussian Noise (GN) model to characterize the Kerr-induced nonlinear interference (NLI) distortions in dispersion uncompensated (DU) coherent optical systems. Such distortions are well modeled as zero-mean additive Gaussian noise in highly dispersive systems [46, 48]. Recent papers pointed out the dependence of the NLI variance on the modulation format [81–83, 86], which is not taken into account in the GN model [48], since the latter assumes that the transmitted signal is a stationary Gaussian process. The GN model provides accurate prediction of DU point-to-point systems performance.

In this chapter we focus on NLI statistics in a network-oriented scenario, where interfering channels may have propagated over different lightpaths before propagating together with the reference channel. We show that the group velocity dispersion (GVD) cumulated by interfering channels during their upstream propagation is the critical parameter to assess the NLI statistics, while the fiber nonlinearity upstream-experienced by interfering channels can be neglected as a first approach. We show that the clouds of the received constellations may not be circular as suggested by the first order perturbative models, and show a significant difference between the variance of NLI noise quadratures, hence displaying a *phase-noise* feature [81]. We

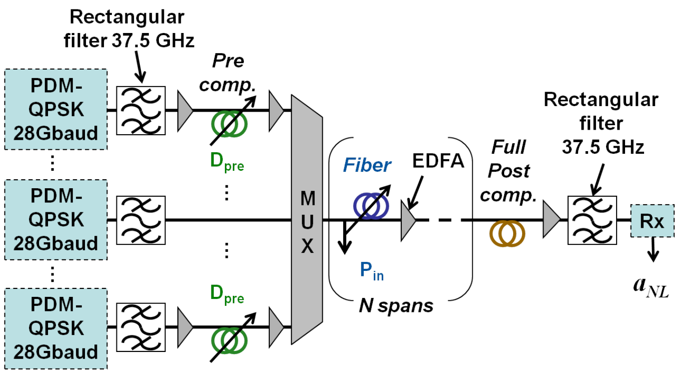


Figure 3.1: Simulation setup.

show the dominance of this feature when transmitting on low dispersion fiber and at small propagating distances. Finally, we assess the impact of NLI phase-noise feature on system Q-factor performance.

## 3.1 Numerical Setup

We simulated with the open source software Optilux [4] the transmission of 9 wavelength division multiplexed (WDM) channels with 28Gbaud non-return to zero polarization division multiplexing quadrature phase shift keying (PDM-QPSK) modulation. Each channel passed through a rectangular-shaped 37.5 GHz wide optical filter before being multiplexed with a channel spacing of 37.5 GHz, as depicted in Fig. 3.1. The transmission fiber was a standard single mode fiber (SMF) with dispersion 16.7 ps/nm/km, span length 100 km, attenuation 0.22 dB/km and nonlinear coefficient $\gamma = 1.3$ 1/W/km. Fiber propagation was modeled using the split-step Fourier algorithm (SSFA) and each channel was modulated with $2^{14}$ purely random symbols within the SSFA window. We performed 4 independent runs for a total of $2^{16}$ transmitted symbols. Amplified Spontaneous Emission (ASE) noise was not included in order to observe only the NLI noise, hence lumped noiseless EDFA amplification exactly recovered the span loss. A variable pre-distortion $D_{\text{pre}}$ was equally applied to all interfering channels before entering the link, while the central channel experienced

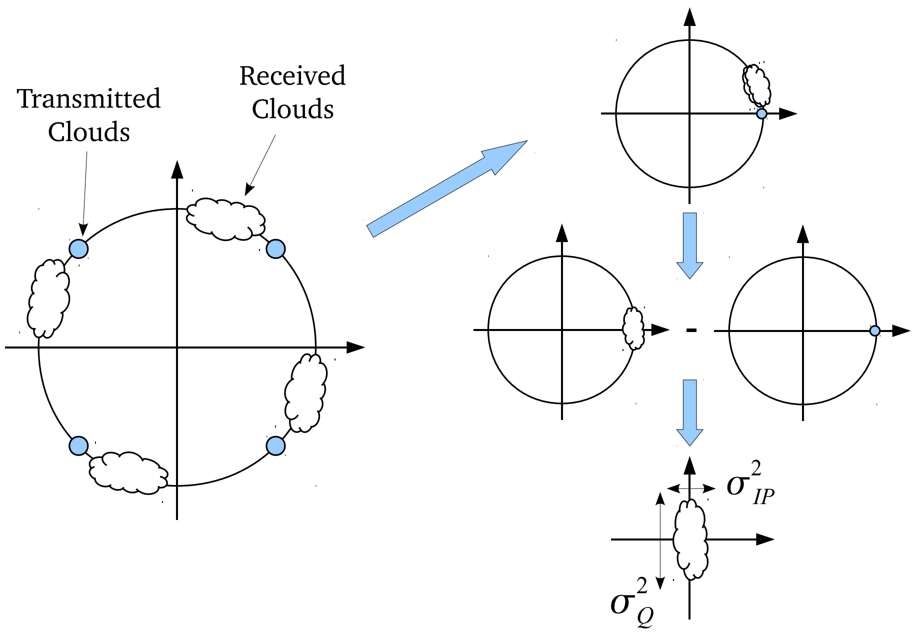


Figure 3.2: Superposition Process to evaluate the NLI variance on both quadratures (removing inter-symbol interference (ISI) generated at the transmitter side). White clouds: actual received constellation. Cyan circles: back-to-back received constellation.

no pre-distortion. The total residual dispersion before the receiver was zero and the launch power was 4 dBm. At the receiver, a rectangular 37.5 GHz wide optical filter extracted the central channel. The signal was sampled at the symbol center, and after the superposition process shown in Fig. 3.2, we measured both quadrature and in-phase variances of the NLI, by averaging on the samples of both polarizations. The steps of the superposition process are the following: i) removal of the initial modulation on both the actual and the back-to-back received constellations; ii) removal of the average phase rotation from the superposed actual received constellation; iii) subtraction of the superposed back-to-back received constellation to the superposed actual received constellation.

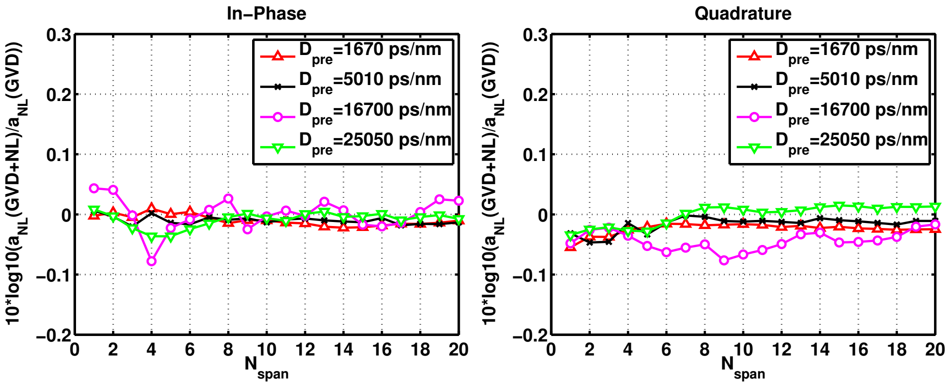


Figure 3.3: Ratio of normalized NLI noise variance $a_{\mathrm{NL}}$ vs. number of SMF spans $N_{\mathrm{span}}$, between upstream propagation accounting for both GVD and fiber nonlinearity and upstream propagation accounting only for GVD. (Left) In-Phase variance, (Right) Quadrature variance.

## 3.2 NLI Statistics with channels propagating over different lightpaths

Since NLI noise variance can be expressed [55] as $\sigma^2_{\mathrm{NLI}} = a_{\mathrm{NL}} P^3$, here we focus on $a_{\mathrm{NL}}$ which is theoretically independent of $P$. Fig. 3.3 shows the increase vs. the number of SMF spans $N_{\mathrm{span}}$ of the normalized NLI noise $a_{\mathrm{NL}}$ variance between interfering channels distortion $D_{\mathrm{pre}}$ accounting for both GVD and fiber nonlinearity and $D_{\mathrm{pre}}$ accounting for only GVD. It is worth noting that the absolute difference is always less than 0.1 dB, allowing further consideration of GVD-only pre-distortion $D_{\mathrm{pre}}$.

Fig. 3.4 shows the $a_{\mathrm{NL}}$ in-phase (left) and quadrature (right) variance vs. SMF $N_{\mathrm{span}}$ for different values of GVD-only pre-distortion $D_{\mathrm{pre}}$ of all interfering channels. It can be noted that when the interfering channels have zero pre-distortion as the central one, the curves of $a_{\mathrm{NL}}$ in-phase and quadrature variances almost coincide, suggesting a NLI circular distribution as in the GN model. The picture changes when interfering channels undergo a non-zero cumulated GVD before propagating together with the central channel. Such a scenario hints what may happen in a DU-based net-

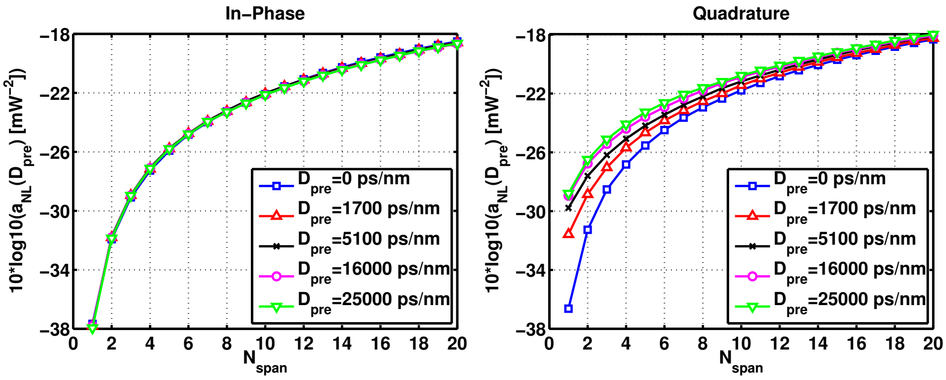


Figure 3.4: $a_{\mathrm{NL}}$ In-Phase (Left) and Quadrature (Right) variance vs. $N_{\mathrm{span}}$, over the central of 9 WDM channels, for different values of interfering channels GVD-only pre-distortion $D_{\mathrm{pre}}$.

work with the WDM comb including interfering channels that have already propagated through some hops. It is worth noting that in this scenario, while the in-phase variance is unaltered w.r.t. $D_{\mathrm{pre}}$, the quadrature variance increases for an increasing pre-distortion value. More specifically, for a fixed number of spans, $a_{\mathrm{NL}}$ quadrature variance quickly increases with non-zero cumulated GVD of interfering channels, showing a saturation for large amounts of $D_{\mathrm{pre}}$. From Fig. 3.4(right) we can observe that for the single span case, the increase of $a_{\mathrm{NL}}$ quadrature variance reaches 8 dB for the largest pre-distortion of 25000 ps/nm. This behavior can be explained by the increase of cross-channel nonlinearity, as large $D_{\mathrm{pre}}$ makes interfering channels close to Gaussian sources, hence inducing *phase-noise* NLI [81]. On the other hand, the gap w.r.t. the zero $D_{\mathrm{pre}}$ case is gradually reduced when propagating through a larger number of spans. A tentative explanation is that the NLI phase-noise feature tends to disappear when the transmitted central channel turns into Gaussian noise, i.e. for large cumulated GVD [82], as intra-channel nonlinearity accumulates at a faster rate than the interchannel one [55].

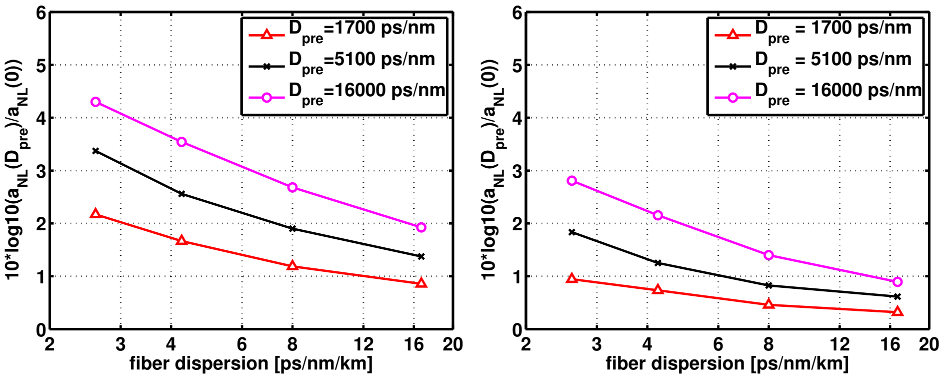


Figure 3.5: Increase w.r.t. zero $D_{\text{pre}}$ of normalized NLI noise $a_{\text{NL}}$ quadrature variance vs. transmission fiber dispersion, over the central of the 9 WDM channels, after propagation through 5 spans (left) or 10 spans (right).

### 3.2.1 Impact of downstream fiber dispersion on NLI variance

Next we assess the impact of fiber dispersion on the phase-noise feature of NLI. Hence we replaced the transmission fiber with either a Teralight-like fiber with dispersion 8 ps/nm/km and $\gamma = 1.3$ 1/W/km, or a large effective area fiber (LEAF) with $\gamma = 1.5$ 1/W/km and dispersion 4.25 or 2.6 ps/nm/km. In all cases the span length was 100 km and the fiber attenuation 0.22 dB/km. In Fig. 3.5 we show, for various $D_{\text{pre}}$, the dB increase of $a_{\text{NL}}$ quadrature variance w.r.t. zero GVD pre-distortion vs transmission fiber dispersion, after propagation through 5 and 10 spans. Please note that Fig. 3.5 also hints at the difference between in-phase and quadrature variances of $a_{\text{NL}}$, since the NLI noise in Fig. 3.4 had almost circular distribution at zero $D_{\text{pre}}$. We see that the NLI phase-noise feature is more emphasized at low fiber dispersion values, i.e., when the central channel takes more spans to converge to a Gaussian-like process. From Fig. 3.5 we may point out the worst scenario from a network perspective: propagation through a few low dispersion spans with interfering channels already propagated on high dispersion fibers, hence with high cumulated GVD. For instance, assuming interfering channels that have already propagated through 3 SMF spans, i.e., with a cumulated GVD around 5100 ps/nm, the transmission of 9 WDM channels over 5 SMF spans (16.7 ps/nm/km) returns a $a_{\rm NL}(D_{\rm pre})/a_{\rm NL}(0)$ ratio equal to 1.3 dB. The same transmission over 5 LEAF spans (4.25 ps/nm/km) returns a twice as high ratio, i.e., 2.6 dB.

## 3.3 Impact of NLI phase-noise feature on Q-performance

In order to understand the impact of the NLI phase-noise feature on bit error rate (BER) we simulated the transmission of 9 WDM 28Gbaud PDM-QPSK channels over 5 and 10 LEAF spans (2.6 ps/nm/km), with interfering channels experiencing both $D_{\rm pre} = 0$ ps/nm or $D_{\rm pre} = 16000$ ps/nm. At the transmitter side, each channel was filtered by a 2nd order supergaussian filter with bandwidth 37.5GHz. We performed noiseless propagation with additive white Gaussian noise (AWGN) loaded in front of the receiver. The amplifiers noise figure (NF) was set to 12 dB to allow feasible BER estimation. At the receiver side, coherent detection of the central channel used a digital signal processing unit including: analog to digital conversion with bandwidth 17GHz, 2 samples-per-symbol data-aided least square butterfly equalization with 7 taps, and Viterbi and Viterbi phase estimation with 15 taps. The BER was estimated with the Monte Carlo algorithm by counting at least 100 errors and averaged over 10 different random seeds. The average BER was then converted to Q-factor.

From results depicted in Fig. 3.5 for 5 spans we observe for this setup a $a_{\rm NL}$ ratio $a_{\rm NL}(16000)/a_{\rm NL}(0) = 4.3$ dB. Fig. 3.6 shows the average Q-factor vs transmitted power when interfering channels pre-distortion is set to 0 or 16000 ps/nm. It is worth noting that for the 5 spans case in the deeply-nonlinear regime the Q-penalty arising from $D_{\rm pre}$ pre-distortion is around 1.5 dB, hence roughly one third of $a_{\rm NL}$ ratio. However, the penalty on the best Q-performance is reduced to 0.4 dB, since at the optimal power ASE noise variance is twice the NLI noise variance [55]. Hence we conclude that the NLI phase-noise feature leads to a minimal penalty on best Q-factor performance, despite a sizable Q-penalty in the deeply-nonlinear regime.

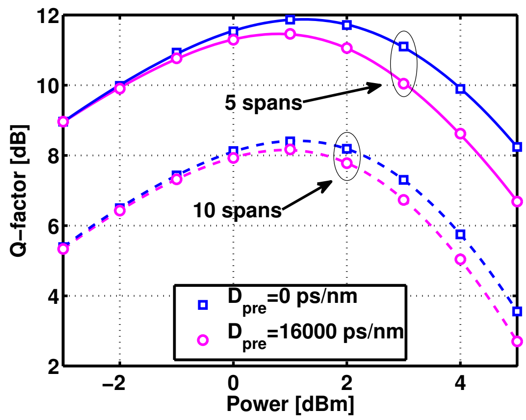


Figure 3.6: Average Q-factor vs transmitted power for a 9- channel 28Gbaud PDM-QPSK system over a DU 5x100km (solid) and 10x100km (dashed) LEAF link. Channel spacing 37.5 GHz. AWGN noise, NF=12 dB.

## 3.4 Conclusions

We showed that an upstream-cumulated GVD of interfering channels alters the NLI noise circular distribution. It turns out that the NLI quadrature variance is increased, while the in-phase variance is unaltered, giving the NLI a phase-noise nature. This phenomenon is reduced as the channel of interest experiences large cumulated dispersion, hence is more evident on low dispersion fibers and at small propagating distances. However, the increase of NLI quadrature variance produces less than half a dB penalty on best Q-factor performance, despite a sizable impact in nonlinear regime.

# Chapter 4

# PDL Impairment on PDM-QPSK Coherent Systems

In chapter 2 we reported on the optimal granularity of single carrier wavelength division multiplexed (WDM) long haul coherent transmission systems, investigating a multiplicity of scenarios including the most popular modulation formats, i.e. polarization division multiplexed (PDM)-binary phase shift keying (BPSK), PDM-quaternary phase shift keying (QPSK), polarization switched (PS)-QPSK and PDM-16 quadrature amplitude modulation (16QAM), and either dispersion managed (DM) or dispersion uncompensated (DU) links. In the study, linear impairments such as polarization mode dispersion (PMD) and polarization dependent loss (PDL) were not accounted for in order to focus only on the fiber nonlinearity main manifestations.

In this chapter we restrict the scenario by considering only one modulation format, i.e., PDM-QPSK, but adding the PDL impairment to the picture. In fact, the impact of PDL on coherent systems has only recently come to the stage, while the literature provides several studies on the impact of PMD [87] and its interaction with fiber nonlinearity [88, 89].

First studies on PDL in coherent systems investigated phenomena such as the signal-to-noise ratio (SNR) degradation in the linear regime [90, 91], while recent works [92–95] pointed out that the PDL/Kerr-effect interplay may have a significant

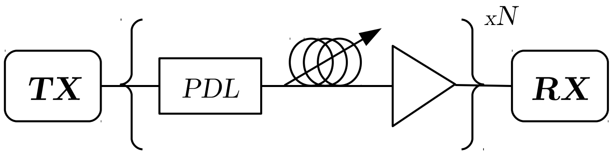


Figure 4.1: Schematic of a fiber optic transmission system with lumped PDL.

impact on terrestrial coherent links. New investigations are thus mandatory.

In this chapter we aim to add to the debate, by providing a simulation study of PDL/Kerr-effect interplay. We report on the interaction of PDL with the main nonlinear effects such as self-phase modulation (SPM), cross-phase modulation (XPM) and cross-polarization modulation (XPolM) when each one acts individually. We highlight this interplay by considering pathological manifestations of PDL. The impact of realistic PDL configurations is quantified in terms of penalty on the Q-factor average performance and on the variance in the Q-factor distribution, for both DM and DU links.

## 4.1 PDL impairment in linear regime

When studying an optical fiber link, we usually assume it as composed of *N* identical blocks, each one including the transmission fiber followed by an amplifier to recover fiber attenuation of the transmitted signal. This is the case of DU systems while with dispersion management an in-line compensating fiber is placed in front of the amplifier to set the desired residual dispersion per span (RDPS). Our reference system is composed of *N* blocks where a lumped PDL element is placed at the beginning of each block, as depicted in Fig. 4.1. In order to first assess the impairment generated by PDL on PDM-QPSK coherent systems in the linear regime, we switch off fiber nonlinearity and set linear impairments such as PMD and group velocity dispersion (GVD) to zero, hence only the signal attenuation of the transmission fiber was considered.

As described in subsec. 1.2.4, the orientation of the eigenvector of a generic PDL element is random on the Poincaré sphere. Hence the effective PDL $\rho$ experienced after $N$ blocks is a random variable with Maxwellian probability density function (1.2).

We start our study by investigating pathological cases, i.e., cases where PDL elements of consecutive blocks have the same eigenvector orientation. In this scenario the effective PDL $\rho$ after $N$ blocks is deterministic and equal to $\rho = \sum_{i=1}^{N} \rho_i$ where $\rho_i$ is the PDL contribution of the $i$-th block. For the sake of simplicity, in this work we will always consider a PDL contribution $\rho_i$ equal for all blocks. By this way, the effective PDL $\rho$ of pathological cases becomes $\rho = \rho_i N$ after $N$ blocks. To simplify, we consider pathological cases with $\varepsilon_i = 0$, $i = 1,\ldots,N$, and $\theta_i = \vartheta$, $i = 1,\ldots,N$. In such a scenario, for a PDM-QPSK signal the worst performance degradation occurs when $\vartheta = 0 + k\pi/2$, $k = 0,1,2,\ldots$. In this case in fact the optical SNR (OSNR) of one of the polarization tributaries is degraded while the OSNR of the other tributary is improved. In this case, the OSNR degradation of one tributary is more relevant than the OSNR improvement of the other one, leading to a worse overall performance. In such a configuration it is possible to derive analytically the OSNR penalty for both tributaries. Assuming that the PDM signal goes through a cascade of $N$ blocks we have:

$$\mathrm{OSNR}_{\mathrm{pen}}^{\mathrm{X}} = \frac{N}{\sum\limits_{i=1}^{N} \left(\frac{1}{1+\Gamma}\right)^i} \qquad \mathrm{OSNR}_{\mathrm{pen}}^{\mathrm{Y}} = \frac{N}{\sum\limits_{i=1}^{N} \left(\frac{1}{1-\Gamma}\right)^i} \tag{4.1}$$

with normalized PDL coefficient $\Gamma$ equal for each block $i = 1,\ldots,N$.

Fig. 4.2 shows the PDL-induced OSNR penalty on both polarization tributaries of a PDM-QPSK signal propagating through $N$ blocks as depicted in Fig. 4.1, with all PDL elements aligned with $(\theta_i, \varepsilon_i) = (0,0)$, $i = 1,\ldots,N$. Solid lines represent the penalties measured in simulation at a reference bit error rate (BER) of $10^{-3}$, while dashed ones are the predicted penalties computed using eq. (4.1). It can be noted that the OSNR improvement of one tributary has little impact on the overall penalty which is mainly set by the penalty of the impaired tributary. We also note that the same amount of effective PDL $\rho$ is less detrimental on the OSNR if distributed through more consecutive blocks. This is a consequence of the fact that in the latter

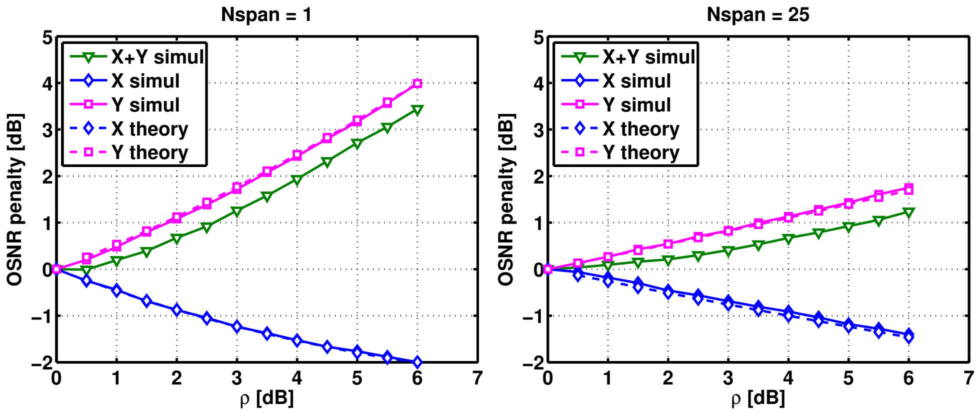


Figure 4.2: OSNR penalty on a PDM-QPSK signal versus effective PDL value $\rho$ generated by a cascade of $N$ aligned PDL blocks with $(\theta_i, \varepsilon_i) = (0, 0)$, $i = 1, \ldots, N$. (Solid) Simulation results, (Dashed) Theory eq. (4.1). (Left) $N = 1$, (Right) $N = 25$.

case also the amplified spontaneous emission (ASE) noise added to the signal from previous amplification is affected, hence reduced, in the following blocks.

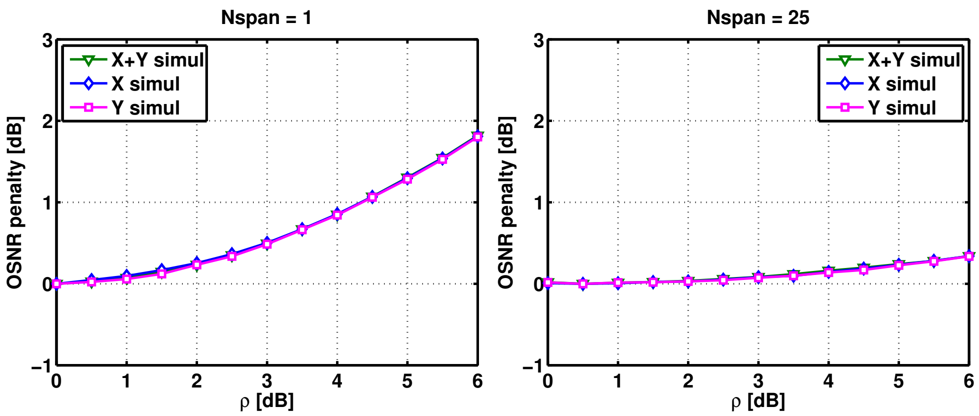


Figure 4.3: OSNR penalty on a PDM-QPSK signal versus effective PDL value $\rho$ generated by a cascade of $N$ aligned PDL blocks with $(\theta_i, \varepsilon_i) = (\pi/4, 0)$, $i = 1, \ldots, N$. (Left) $N = 1$. (Right) $N = 25$.

On the other hand, the lowest overall performance degradation in pathological cases occurs when $\vartheta = \pi/4 + k\pi/2$, $k = 0, 1, 2, \ldots$, i.e., when both polarization tributaries experience the same OSNR degradation. Fig. 4.3 shows the PDL-induced OSNR penalty on both polarization tributaries of a PDM-QPSK signal propagating through $N$ blocks as depicted in Fig. 4.1, with all PDL elements aligned with $(\theta_i, \varepsilon_i) = (\pi/4, 0)$, $i = 1, \ldots, N$.

## 4.2 Interplay between PDL and fiber nonlinearity

In the previous section we described the impact of PDL on PDM-QPSK coherent systems in the linear regime. However, the system designer is interested in the extra margin that must be allocated to the PDL penalty in realistic conditions that include the Kerr nonlinearity. Recent works [92–95] pointed out that the PDL/Kerr-effect interplay may have a significant impact on such a margin in terrestrial coherent links. Hence, we numerically investigate the interaction between PDL and Kerr effect by decoupling the fiber nonlinearities, namely, SPM, XPM and XPolM as described in Sec. 1.5.1. By means of simple setups we explain the reasons of such an interaction. In this section we quantify the impact of PDL on the average performance by considering three possible scenarios. In the first case we transmit over a DM link a homogeneous WDM system with non return to zero (NRZ) 100 Gb/s PDM-QPSK channels. In the second case we repeat the same test when surrounding the reference 100 Gb/s PDM-QPSK channel with 10 Gb/s NRZ On Off Keying (OOK) neighbors, obtaining a hybrid system. In the third case we transmit over a DU link a WDM system with 100 Gb/s PDM-QPSK channels.

### 4.2.1 Numerical Setup

We used the the open source software Optilux [4] to simulate the system depicted in Fig. 4.4, where PDL was implemented by a lumped element placed before the transmission fiber. Its input/output field relation is expressed in eq. (1.7). However, when building up the concatenation of two consecutive PDL element $i$ and $i+1$, the product $\underline{\underline{B_{i+1}}}\,\underline{\underline{B_i}}^{\dagger}$ between rotation matrices of the $i$-th and $i$+1-th PDL elements is

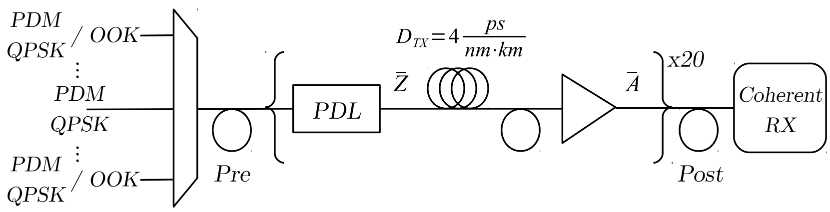


Figure 4.4: The simulated system with lumped PDL along the link.

equivalent to a new matrix $\underline{\underline{B_{i+1}}}^*$. Hence we can simplify eq. (1.7) for the $i$-th PDL element along the link as:

$$\begin{bmatrix} Z_x^{(i)} \\ Z_y^{(i)} \end{bmatrix} = \begin{bmatrix} \sqrt{1+\Gamma} & 0 \\ 0 & \sqrt{1-\Gamma} \end{bmatrix} \begin{bmatrix} \cos\theta_i & \sin\theta_i \\ -\sin\theta_i & \cos\theta_i \end{bmatrix} \begin{bmatrix} \cos\varepsilon_i & j\sin\varepsilon_i \\ j\sin\varepsilon_i & \cos\varepsilon_i \end{bmatrix} \begin{bmatrix} A_x^{(i)} \\ A_y^{(i)} \end{bmatrix} \tag{4.2}$$

where $\overrightarrow{A}^{(i)} = \left[A_x^{(i)}, A_y^{(i)}\right]^{\mathrm{T}}$ is the noisy field at the end of the $i$-th span, $i = 1, \ldots, N$.

In the DM case the optical link was composed of $N = 20$ spans of non zero dispersion shifted fiber (NZDSF) (length 100 km, attenuation 0.2 dB/km, dispersion 4 ps/nm/km, nonlinear index $\gamma = 1.5$ 1/W/km, no PMD) with residual dispersion per span of 30 ps/nm and Pre-compensation of -345 ps/nm. The DU link was identical, except for a total of $N = 25$ spans with neither in-line compensation nor Pre-compensation. In all cases, after the link, a post-compensating fiber set the overall cumulated dispersion to zero. The propagation within the optical fibers was numerically solved by the split-step Fourier algorithm described in Sec. 1.5. Each nonlinear step was implemented by activating the nonlinearity of interest, namely SPM, XPM or XPolM, as described in Sec. 1.5.1.

The nonlinear signal-noise interaction along the link was accounted for by flat-gain noisy amplifiers with 7 dB noise figure. The WDM comb consisted of 19 channels, with 50 GHz spacing. The central channel was a 112 Gb/s PDM-QPSK, surrounded by either 112 Gb/s PDM-QPSK channels or by 10 Gb/s OOK channels. In the homogeneous DM case we used for each PDM-QPSK channel 1024 random symbols, in the hybrid DM case 840, while in the homogeneous DU case 2048.

The DSP-based receiver performed polarization recovery by a 2-samples-per-symbol data-aided least-squares equalizer with 7 taps and carrier phase estimation through the Viterbi&Viterbi algorithm with 15 taps. The performance was measured in terms of the Q-factor obtained by inverting the BER of the central channel, estimated from Monte Carlo (MC) simulation stopped after counting at least 300 errors.

### 4.2.2 Interaction between pathological PDL and SPM, XPM and XPolM

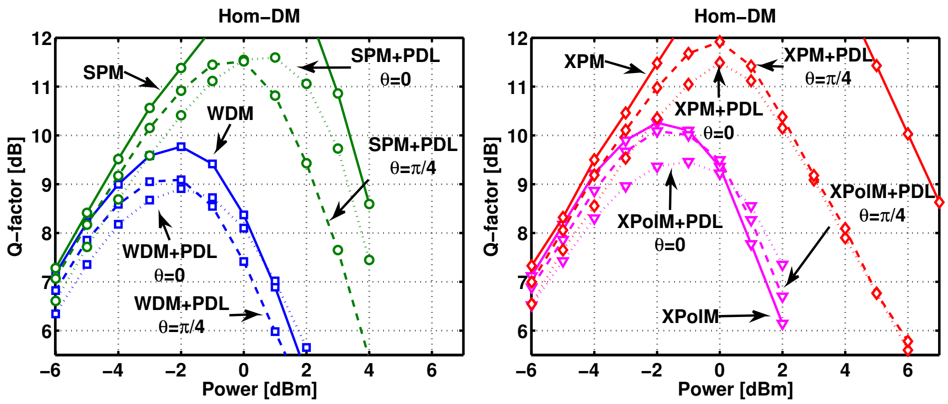


Figure 4.5: Average performance for the homogeneous setup on the $20 \times 100$ km DM link, without PDL (solid lines) or with a total PDL of $\rho = 5$ dB, $\varepsilon = 0$. (Left) Propagation of single-channel or WDM. (Right) Propagation with only selected cross-channel impairments.

Fig. 4.5 shows the average Q-factor vs. signal power using our nonlinearity decoupling method in the DM link for the homogeneous setup. The performance is averaged over 10 seeds, each corresponding to random carrier state of polarization (SOP) of the transmitted channels. In [92, 93] it has been shown that pathological cases of all-aligned PDL elements such as $(\theta_i, \varepsilon_i) = (0, 0)$ and $(\theta_i, \varepsilon_i) = (\pi/4, 0)$ are a worst case for PDL in the linear and nonlinear regimes, respectively. Hence, in order to understand the interaction of PDL with Kerr effect in a simple way, each nonlinear effect was analyzed separately in three scenarios: i) without PDL (solid lines), ii) with aligned PDL sections having $\rho_i = 0.25$ dB and $(\theta_i, \varepsilon_i) = (0, 0)$ (dotted lines) or iii) $(\theta_i, \varepsilon_i) = (\pi/4, 0)$ (dashed lines). The curve labeled WDM corresponds to the real case including all nonlinearities.

We start analyzing the linear regime, i.e., the initial increasing part of the Q-factor curves. Here in all cases the worst PDL is with $(\theta_i, \varepsilon_i) = (0, 0)$ where the polarization dependent SNR degradation induced by PDL is maximum and cannot be compensated by the DSP [92]. Moving to the nonlinear regime (decreasing part of the Q-factor curve), the behavior of SPM and XPM curves can be explained by looking at the total instantaneous power after the generic $i$-th PDL element:

$$\begin{aligned} P_{\text{out}}(t) = \left|A_x^{(i)}\right|^2 + \left|A_y^{(i)}\right|^2 + \Gamma\cos 2\theta \cos 2\varepsilon \left(\left|A_x^{(i)}\right|^2 - \left|A_y^{(i)}\right|^2\right) \\ + \Gamma\cos 2\theta \sin 2\varepsilon \left(2\mathrm{I_m}\left\{A_x^{(i)}\left(A_y^{(i)}\right)^*\right\}\right) + \Gamma\sin 2\theta \left(2\mathrm{R_e}\left\{A_x^{(i)}\left(A_y^{(i)}\right)^*\right\}\right). \end{aligned} \tag{4.3}$$

For phase modulated signals, the terms $\left|A_x^{(i)}\right|^2$ and $\left|A_y^{(i)}\right|^2$ remain almost constant along propagation in DM systems. Consequently, at $\varepsilon = 0$ only the last term in (4.3) matters, especially when $\theta = \pi/4$ [93]. Unfortunately, this term is data dependent and thus has wild variations in time [95] that enhance nonlinear distortions. Generalizing over the Poincaré sphere, this effect is maximum when the 3D Stokes vector of the PDL eigenvector lays on the polarization plane of the input PDM-QPSK signal.

The $\theta$-dependence does not appear for XPM because the transmitted SOP of the interfering channels are randomly oriented with respect to the central channel, thus removing on average the dependence on the absolute reference system of the PDL eigenvector orientation (i.e., on $\theta$).

For XPolM the interplay with PDL is different. Because each PDL element is a partial polarizer, the pattern-induced temporal fluctuations of the SOP are reduced, hence reducing XPolM. This can be easily observed from Fig. 4.6, which depicts the time trajectory of the Stokes vector of a PDM-QPSK signal before and after propagation through a PDL element with $(\theta, \varepsilon) = (0, 0)$ and $\rho = 5$ dB. The re-polarization induced by the PDL eigenvector is clearly understood by the reduced distance in

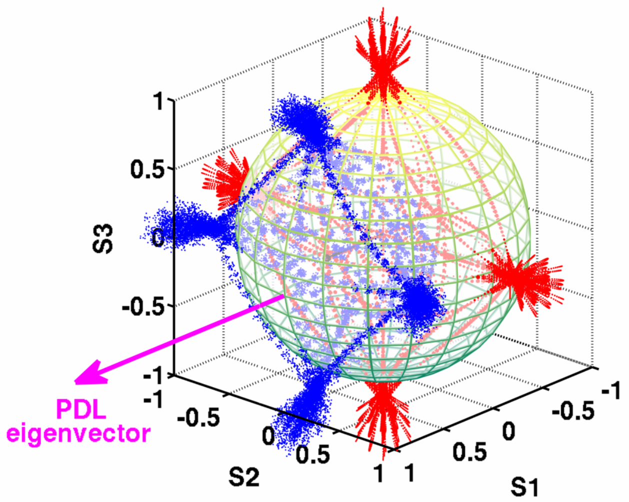


Figure 4.6: Time-sampled Stokes vector of a PDM-QPSK signal before (red) and after (blue) propagation through a lumped PDL element with $\rho = 5$ dB and $(\theta, \varepsilon) = (0, 0)$. Poincaré sphere shown for reference.

Stokes space of the four output PDM-QPSK symbol clouds. Fig. 4.5 (left) indicates that the linear penalty due to PDL is more than compensated by the reduction of XPolM at large powers. Overall, the beneficial impact of PDL on XPolM is not visible in the real WDM curve, an indication that SPM and XPM cannot be neglected when PDL is present. Moreover, the worst case Q-factor of the WDM case shows 1 dB of penalty with respect to the no PDL case at any power, although for different $\theta$, i.e., $\theta = \pi/4$ in nonlinear regime and $\theta = 0$ in linear regime.

Moving to the hybrid scenario, we replaced PDM-QPSK interfering channels with 10 Gb/s OOK having average power locked 4 dB below that of the central 112 Gb/s PDM-QPSK channel, in order to ensure equal performance for both modulation formats. Fig. 4.7 shows the average Q-factor when applying the nonlinear decoupling method. This setup has scalar XPM as the dominant nonlinear distortion because PDM-QPSK suffers the intensity fluctuations of OOK neighbors [72]. Since $A_y^{(i)} = 0$, only the first of PDL-related terms in (4.3) survives, causing a $\theta$-dependent coupling

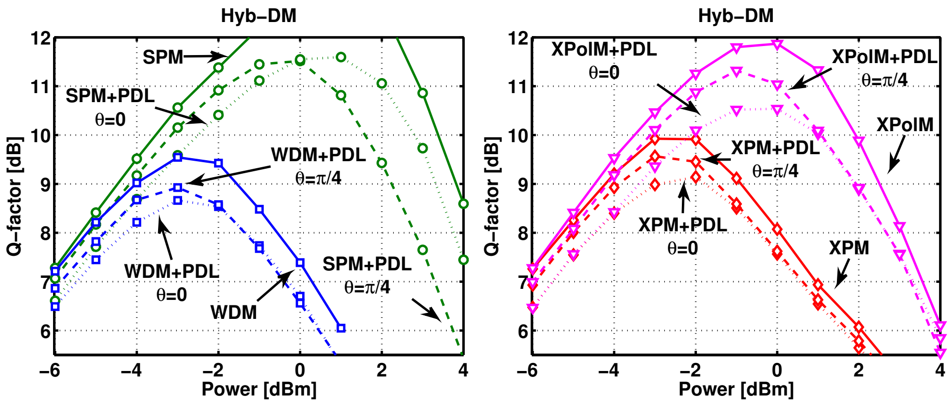


Figure 4.7: Average performance for the hybrid setup on the $20 \times 100$ km DM link, without PDL (solid lines) or with a total PDL of $\rho = 5$ dB, $\varepsilon = 0$. (Left) Propagation of single-channel or WDM. (Right) Propagation with only selected cross-channel impairments.

of XPM with PDL that can be detrimental when $\theta = 0$ or beneficial when $\theta = \pi/2$.

To support this claim we report in Fig. 4.8 the Q-factor for the hybrid setup using $\theta = 0, \pi/4, \pi/2$ and carrier SOP aligned. It is worth noting that with $\theta = \pi/2$ the Q-factor is higher than the no-PDL case, despite that $\theta = 0, \pi/4$ perform worsen. Equation (4.3) is in agreement with this observation, since for an OOK interfering channel ($A_y^{(i)} = 0$) it writes as:

$$P_{\text{out}}(t) = \left|A_x^{(i)}\right|^2 + \Gamma \cos 2\theta \cos 2\varepsilon \left|A_x^{(i)}\right|^2 .$$

With $\theta = \pi/2$ and $\varepsilon = 0$ the last term in the right-hand side is negative, thus reducing the power and hence the nonlinear effect. This observation is peculiar of the hybrid setup where $A_y^{(i)} = 0$. However, in Fig. 4.7 the PDL/XPM interaction is masked by the random orientation of the transmitted SOP of the interfering channels. XPolM is now a secondary impairment and does not show improvements with PDL, because the PDL polarizing effect has minor importance with OOK neighbors.

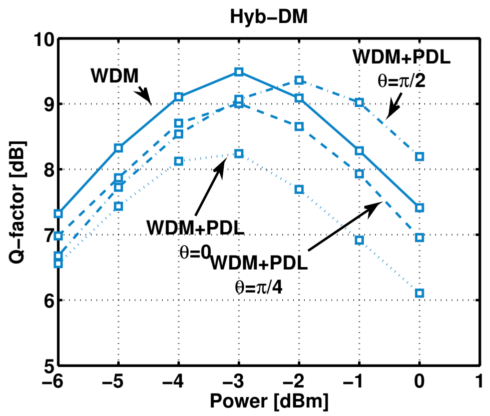


Figure 4.8: Average performance for the hybrid setup on the 20 × 100 km DM link, without PDL (solid lines) or with a total PDL of $\rho = 5$ dB, $\varepsilon = 0$. SOP of interfering channels aligned to the SOP of the reference channel.

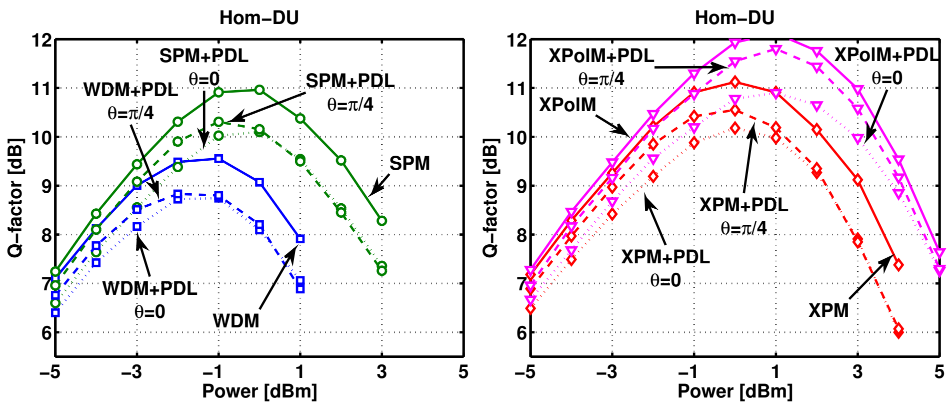


Figure 4.9: Average performance for the homogeneous setup on the 25 × 100 km DU link, without PDL (solid lines) or with a total PDL of $\rho = 5$ dB, $\varepsilon = 0$. (Left) Propagation of single-channel or WDM. (Right) Propagation with only selected cross-channel impairments.

Fig. 4.9 shows the average Q-factor for the homogeneous setup transmitted on the DU link. In this scenario, dispersion induces strong signal fluctuations, hence reducing the importance of PDL-related terms in (4.3), as shown by the $\theta$-independent Q-performance of both SPM and XPM in nonlinear regime. XPolM coupling with PDL follows similar trends as ASE noise in linear regime, suggesting that XPolM behaves like an additive white noise. Overall, PDL worsens the Q-factor by about 1 dB.

### 4.2.3 Impact of realistic PDL on Q-factor average performance

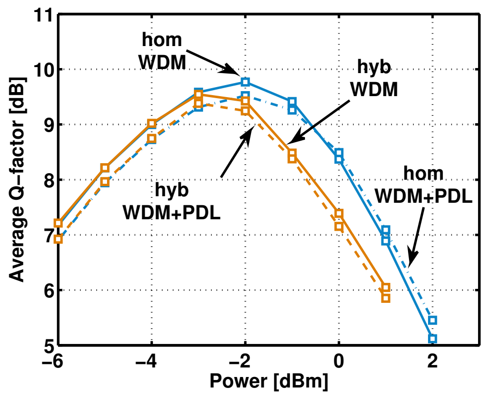


Figure 4.10: Average performance in both hybrid and homogeneous scenarios for WDM propagation on the $20 \times 100$ km DM link. Propagation without PDL or with $\rho_{\text{rms}} = 2.25$ dB (PDL elements with random orientation).

In Sec. 4.2.2 we investigated pathological cases of PDL to highlight its interplay with Kerr-effect main manifestations. However in realistic setups the eigenvectors of PDL elements are randomly oriented, hence the pathological cases analyzed in Sec. 4.2.2 are very unlikely [92, 93]. In real setups the total cumulated PDL has a Maxwellian distribution with root-mean-square value related to $\rho_i$ by $\rho_{\text{rms}} = \rho_i \cdot \sqrt{N}$, with $N$ the number of spans. Fig. 4.10 reports the impact on the average Q-factor by randomly-oriented PDL elements for both homogeneous and hybrid scenarios on DM link. The performance is averaged over 10 seeds, each corresponding to random carrier SOP and PDL orientations $(\theta_i, \varepsilon_i)$. It can be noted that, in absence of nonlinearity, PDL

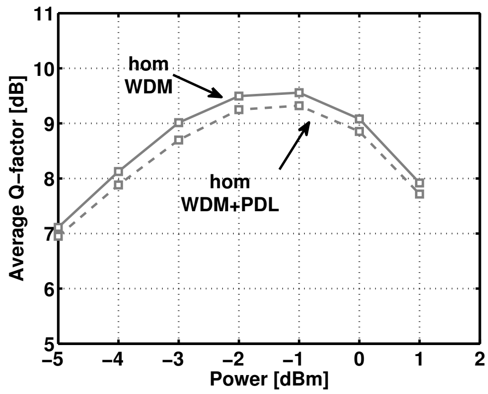


Figure 4.11: Average performance in the homogeneous scenario for WDM propagation on the $25 \times 100$ km DU link. Propagation without PDL or with $\rho_{\mathrm{rms}} = 2.25$ dB (PDL elements with random orientation).

yields a penalty of about 0.3 dB on the average Q-factor, as shown by the gap in linear regime between solid and dashed curves. Moving to the nonlinear regime, we observe that in the hybrid scenario the Q-penalty is comparable to that in the linear regime, while in the homogeneous scenario the presence of PDL slightly improves the average Q-factor as a result of the partial mitigation of the dominant nonlinearity in this setup, i.e., the XPolM effect.

Fig. 4.11 reports the impact on the average Q-factor by randomly-oriented PDL elements for the homogeneous scenario on DU link. Also in this case, the performance is averaged over 10 seeds, each corresponding to random carrier SOP and PDL orientations $(\theta_i, \varepsilon_i)$. We note that, the penalty on average Q-factor lies around 0.3 dB in both linear and nonlinear regimes, in line with what expected for a system with power-dependent dominant nonlinearities.

However, the investigation of the penalty on average performance is not fully comprehensive to assess the impact on the system of PDL-nonlinearity interplay. For this reason, in the next section we investigate the impact of this interaction on the distribution of the Q-factor.

### 4.2.4 Impact of PDL-nonlinearity interplay on Q-factor distribution

A more interesting description of the impact of realistic PDL is given by the distribution of the Q-factor. However, to obtain this distribution is a challenging task due to the great amount of simulation time necessary to collect a sufficient number of BER samples to correctly infer the distribution tails. This reason calls for simplifications of the simulated system that easily lead to wrong estimates.

#### Correct estimation procedure for BER distribution

When measuring the BER, averaging is used to summarize the impact of the stochastic parameters affecting the received signal into a unique quality of transmission parameter. For this reason, the measurement time of the simulation or experiment should be kept sufficiently long to get stable BER estimates. However, most of the times the stochastic variations of the disturbing parameters operate over widely different time scales. For instance, ASE noise has very fast dynamics of the order of [THz], while PDL/PMD vary with frequencies of the order of [kHz] [96].

From the previous discussion we understand that it is necessary to distinguish between slowly varying parameters that do not vary during a BER measurement, called *slow* random variables (RVs), and those that do, i.e., rapidly varying parameters called *fast* RVs. The distribution of the BER, and hence of the Q-factor, must account for the randomness of the *slow* RVs only, while the impact of the *fast* RVs must be averaged in the BER measurement. Among *slow* RVs we identify PDL, PMD and laser SOP, while ASE noise and information symbols represent *fast* RVs.

Thus the *correct procedure* to estimate the BER distribution requires two nested cycles: an inner one on *fast* RVs and an outer one on *slow* RVs. By this way, the outer cycle provides information about the drift of the BER due to slowly varying parameters, that eventually leads to out-of-service events occurring when the BER is larger than a reference value $\text{BER}_{\text{FEC}}$ set by the forward error correction (FEC) threshold.

This approach is extremely time consuming, because in the inner cycle it requires to simulate several times the propagation of the WDM comb along the link to obtain a single BER sample (averaged over *fast* RVs). In systems where the interaction signal-ASE noise is negligible, the simulation can be speeded up by loading, after a noiseless propagation, all the equivalent noise at the end of the link, which ensures the correct total ASE power spectral density. In these cases, we might be tempted to include information symbols among *slow* RVs in the outer cycle, with the consequence of a dramatically shorter inner cycle coming from a single WDM comb noiseless propagation and a relatively fast, w.r.t. signal propagation, noise loading at the end of the link. This approach, that we will call *naive*, leads to a large overestimation of the variance of the Q-factor distribution, as shown in App. B.3.

Besides correctly dealing with input RVs, it is necessary to avoid the Q-factor distribution being corrupted by the MC uncertainty in the inner cycle on *fast* RVs. Since MC error is independent of the *slow* RVs, it substantially adds its variance to the true variance of the Q-factor distribution. Therefore we must ensure that the MC variance in the *fast* RVs inner cycle is much smaller than the variance of the overall distribution. We verified that for our system, counting at least 400 errors in absence of PDL and 300 errors with PDL was enough for an accurate estimation of the overall distribution.

### PDL and Kerr-effect induced Q-factor distribution

In this section we investigate the Q-factor distribution induced by the presence of both realistic PDL and fiber nonlinearity, by comparing the results to when both impairments act separately. To this aim we estimate the Q-factor distribution for the three scenarios considered in Sec. 4.2.3 by working 1 dB above the optimal power, i.e., with: P = -1 dBm (OSNR = 17 dB/0.1 nm) for the homogeneous DM setup; P = -1.5 dBm (OSNR = 16.5 dB/0.1 nm) for the hybrid DM setup; P = 0 dBm (OSNR = 17 dB/0.1 nm) for the homogeneous DU setup, as can be inferred by Figs. 4.10–4.11. We present the results by the probability mass function (PMF) of the discretized Q-factor. The Q-factor, in a decibel scale, was discretized in steps of 0.15 dB. Fig. 4.12 shows the estimated PMF of the Q-factor both in presence and absence of PDL for both homogeneous and hybrid WDM systems propagated in the DM link and for the homogeneous system propagated in the DU link. The PMF was estimated by iterating

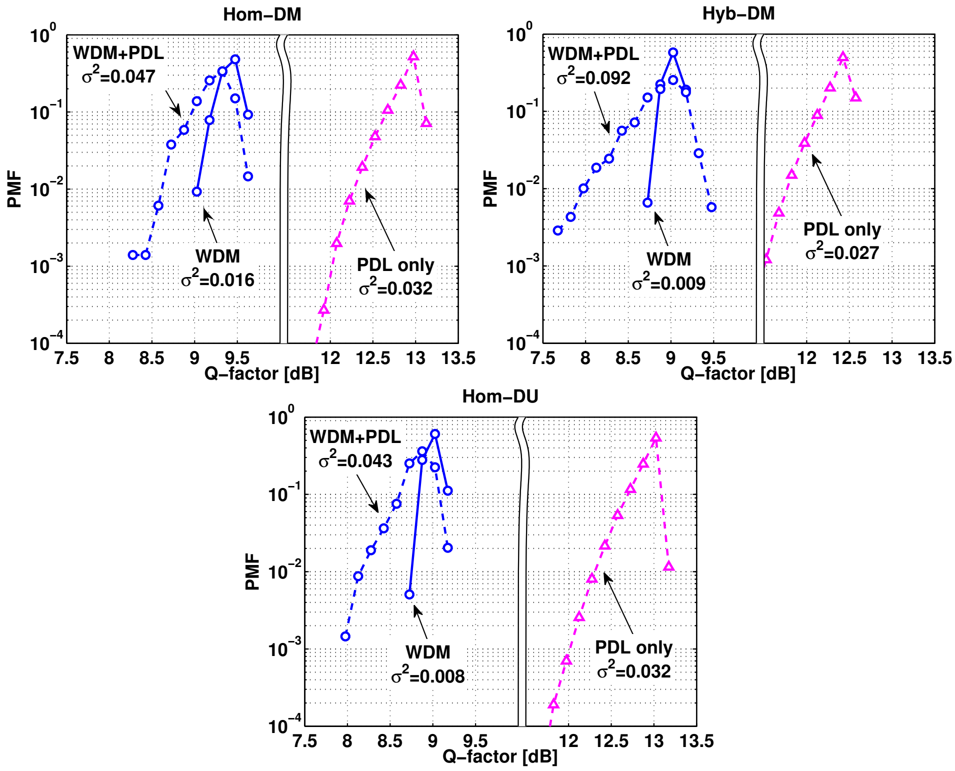


Figure 4.12: Q-factor PMF for WDM propagation, without PDL (200 seeds) or with $\rho_{\mathrm{rms}} = 2.25$ dB (700 seeds, PDL elements with random orientation), and with only PDL ($10^5$ seeds). Bin size 0.15 dB. (Top-Left) Homogeneous setup on the $20 \times 100$ km DM link with Power = -1 dBm, (Top-Right) Hybrid setup on the $20 \times 100$ km DM link with Power = -1.5 dBm, (Bottom) Homogeneous setup on the $25 \times 100$ km DU link with Power = 0 dBm.

on different random seeds of the *slow* RVs. From Fig. 4.12 we can see that without PDL the Q-factor distribution is very narrow, indicating that the carrier SOP has little impact on it. Turning PDL on remarkably enlarges such a distribution, leading to an increased outage probability. For instance, assuming a FEC Q-threshold of 8.5 dB, the probability of an outage event, i.e. Q-factor lower than 8.5 dB, is essentially zero without PDL while in presence of PDL it becomes $2.8 \cdot 10^{-3}$ for homogeneous, $1.2 \cdot 10^{-1}$ for hybrid systems on the DM link and $6.8 \cdot 10^{-2}$ for the homogeneous setup on DU link. Looking at Fig. 4.12, the outage probability at FEC Q-threshold is the sum of the PMF values below Q = 8.5 dB.

Fig. 4.12 also shows, for each setup, the Q-factor PMF in absence of nonlinearity (PDL-only) with $\rho_{\mathrm{rms}} = 2.25$ dB, obtained with the reverse channel method (RCM) [97] (and double-checked when feasible by MC simulations). We observe that the interaction between PDL and nonlinearity yields a Q-factor variance $\sigma^2$ at least 1.4 times larger with respect to the PDL-only case. We find that, contrary to [94, 95], the Q-factor PMF does not shrink with both PDL and nonlinearity, and always has a larger variance than the PDL-only case. As an explanation for this discrepancy with the conjecture in [95] we point out that the PDL-only PMF in [95, Fig. 2] is not supported by experiments, and the maximum Q-factor reported for that setup exceeds the one predicted by RCM in absence of PDL. We suspect therefore a poor accuracy in those MC estimations of the Q-factor, which could have artificially broadened the PMF and therefore led the authors to different conclusions once compared to the PMF experimentally measured in the presence of PDL.

## 4.3 Conclusions

By decoupling the Kerr nonlinearities, namely SPM, XPM and XPolM, we investigated their interplay with PDL in dispersion-managed optical systems, showing different interactions between phase and polarization effects. We also showed that the presence of realistic random PDL has little impact on average Q-factor in both dispersion-managed and dispersion-unmanaged links, but makes a remarkable difference on Q-factor distribution. As a result, while we did not observe outage events without PDL, at an rms PDL of 2.25 dB we estimated an outage probability $6.8 \cdot 10^{-2}$ for the DU link at FEC Q-threshold of 8.5 dB, $2.8 \cdot 10^{-3}$ for the DM-homogeneous setup and $1.2 \cdot 10^{-1}$ for the DM-hybrid setup. Besides, the PDL-nonlinearity interplay gives a Q-factor distribution at least 1.4 times broader than the PDL-only case.

# Chapter 5

# Stratified Sampling for PDL-induced OP estimation

In chapter 4 we showed that the presence of realistic polarization dependent loss (PDL) along an optical transmission systems has an almost negligible impact on the average Q-factor performance. However, PDL significantly affect the system outage probability (OP) since it is the major cause of Q-factor variance. Moreover, in presence of fiber nonlinearity, the interaction between PDL and Kerr effect increases the Q-factor variance. Thorough estimation of OP is unfeasible with conventional techniques as Monte Carlo (MC) algorithm at practical OP values, and advanced methods are mandatory to speed up the computation [98]. A solution for an efficient numerical simulation of PDL-induced OP in the linear regime was proposed by Tao et al. by a semi-analytical algorithm [97].

In this chapter, we present a novel adaptive Stratified-Sampling (SS) algorithm to efficiently assess the PDL-induced OP in presence of fiber nonlinearity and polarization mode dispersion (PMD). We quantify the efficiency in simulation time of the proposed algorithm in achieving a target accuracy, by comparison with the standard MC method when considering a purely linear system whose OP value makes MC estimation feasible. Later on, we apply the SS algorithm to polarization division multiplexing (PDM)-quadrature phase shift keying (QPSK) transmissions in both dispersion-managed (DM) and dispersion-unmanaged (DU) links, showing a non-negligible impact of PDL using typical ITU-T recommendations [99]. Finally, we show that estimating the OP by neglecting the interaction of nonlinearity and PDL can be largely optimistic.

## 5.1 Estimation of Outage Events

The BER is the average rate of bit errors in the detected data stream, where averaging is used to summarize the impact of the stochastic parameters affecting the received signal. However, averaged BER is not sufficient for the system designer, since it cannot give information about the rate of out-of-service events [100]. A more interesting quality of transmission parameter is given by the outage probability, defined as the probability that the averaged BER is larger than a reference value $\text{BER}_{\text{FEC}}$ set by the FEC threshold.

In Sec. 4.2.4 we described the correct procedure to estimate the BER distribution, by discriminating among stochastic parameters fast RVs on which BER must be averaged and slow RVs accounting for the drift of average BER leading to out-of-service events. We have thus the following definitions [91]:

$$\text{BER}(X) = E\left[\mathscr{I}\left(\hat{b}_k \neq b_k\right)\middle| X\right] \tag{5.1}$$

$$\text{OP} = E[\underbrace{\mathscr{I}\left(\text{BER}(X) > \text{BER}_{\text{FEC}}\right)}_{\mathscr{I}_O(X)}] \tag{5.2}$$

where: $X$ indicates the set of slow RVs, $b_k$ is the transmitted bit, $\hat{b}_k$ is the detected bit; $E[.|X]$ indicates statistical conditional expectation over the fast RVs (estimated by time-averaging over the short time-scale); $\mathscr{I}$ is the indicator function (which equals 1 when the argument is true and 0 otherwise); $\mathscr{I}_O(X)$ the indicator that BER exceeds the FEC threshold; and $E[.]$ in (5.2) indicates statistical expectation over the slow RVs $X$ (estimated by time-averaging over the long time-scale).

When a reliable semi-analytical BER model for the expectation in (5.1) w.r.t. the fast RVs is not available, the BER in (5.1) can be estimated by running a MC simulation of a data stream of $n$ bits (possibly split in parallel MC simulations of shorter length) with a fixed configuration of slow RVs. It is worth noting that, for independent errors in the $n$-bit pattern, the MC measurement variance is $\sigma^2 = \mathrm{BER}(1-\mathrm{BER})/n$. The standard deviation $\sigma$ must be much smaller than the estimated BER value while spanning the values of $X$, otherwise the OP estimated value will be corrupted by the BER measurement noise. Hence, the number of MC runs must be chosen in order to cap on $\sigma^2$, which is unknown at the beginning of the simulation. For this reason, a conservative choice with large accuracy for the BER estimation is recommended, especially for a link with small PDL and PMD leading to a small OP. In this work, we always estimated the conditional BER (5.1) by using the MC algorithm. In the next section we present an efficient algorithm to estimate the unconditional expectation (5.2) in presence of PDL along the link.

## 5.2 Stratified-Sampling for OP estimation

To introduce the ideas, let's consider for simplicity the case where the only slow RV is the PDL, and concentrate on the unconditional expectation (5.2). We plan to apply SS [98] to the estimation of such an expectation, to be performed in the outer cycle of the overall simulation. We will use a generalization of the SS algorithm proposed in [101], which can be seen as a form of IS, where the input sample space (i.e., the set of all possible slow RV realizations) is partitioned into disjoint subsets, or strata, and independent (local) MC simulations are performed within each stratum.

A block diagram describing the proposed adaptive SS importance-sampling OP estimation is shown in Fig. 5.1. The key point is the availability of a "fast system" yielding (at negligible computational cost) a control variable $C$ highly correlated with the "true system" output (the BER in our case, at high computational cost). The variable $C(X)$ is next classified into strata $\mathscr{S}_C(l)$ (whose image in the $X$ space is $\mathscr{S}_X(l)$) for $l = 1,\ldots,L$. The visiting probability of stratum $l$ is the unwarped probability $P_l = P\{C \in \mathscr{S}_C(l)\} \equiv P\{X \in \mathscr{S}_X(l)\}$. The outage probability in stratum $l$ is $\mathrm{OP}(l) = E\left[\mathscr{I}_O(X)\middle| X \in \mathscr{S}_X(l)\right]$. If the standard deviation $\sigma_l = \sqrt{\mathrm{OP}(l)\,(1-\mathrm{OP}(l))}$ of the Bernoulli indicator $\mathscr{I}_O$ in stratum $l$ is known for the given stratification for all $l = 1,\ldots,L$, then the SS algorithm biases (warps) $P_l$ according to the SS optimal rule [98, 101]:

$$P_l^* = \frac{\sigma_l P_l}{\sum_{i=1}^{L} \sigma_i P_i}, \qquad l = 1, \ldots, L. \tag{5.3}$$

This warping is achieved by the accept/reject block that operates by accepting an unwarped sample from stratum $l$ with probability:

$$W(l) = \frac{\frac{P_l^*}{P_l}}{\sum_{i=1}^{L} \frac{P_i^*}{P_i}} = \frac{\sigma_l}{\sum_{i=1}^{L} \sigma_i}, \qquad l = 1, \ldots, L. \tag{5.4}$$

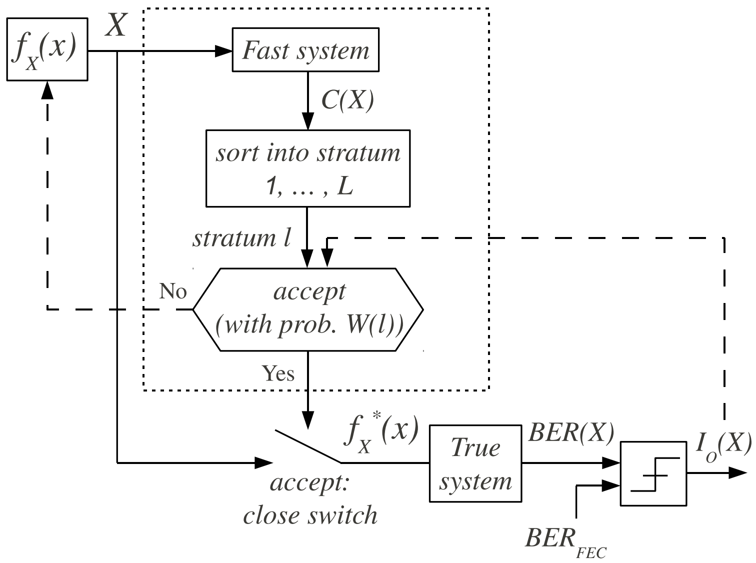


Figure 5.1: Block diagram of proposed adaptive SS importance-sampling OP estimation. The dotted box is the SS control unit of the accept/reject input sieve that implements the warped PDL input distribution $f_X^*(X)$. If the stratum test is passed (with probability $W(l)$), the corresponding PDL realization $X$ is accepted for propagation in the true system, yielding a BER sample and an outage indicator $\mathscr{I}_O$.

The second equality in (5.4) can be used to make the SS algorithm adaptive by updating the OP$(l)$ estimates based on the feedback of the $\mathscr{I}_O$ observable, as shown in Fig. 5.1. To clarify the expression of $W(l)$, suppose to generate $M$ samples. The number of samples in stratum $l$ after the accept/reject block is $M_{\mathrm{w}}(l) \simeq MP_l W(l)$, hence, thanks to the rejections, the relative frequency of samples in stratum $l$ at the output of the accept/reject block is:

$$\frac{M_{\mathrm{w}}(l)}{\sum_{i=1}^{L} M_{\mathrm{w}}(i)} \xrightarrow[M\to\infty]{} P_l^*, \qquad l = 1,\ldots,L$$

as desired. $W(l)$ thus changes the input distribution of PDL expressed by $f_X(X)$ into a warped PDF $f_X^*(X)$ at the input of the true system, with $X = [(\theta_1, \varepsilon_1), ..., (\theta_N, \varepsilon_N)]^T$ where $(\theta_i, \varepsilon_i)$ are azimuth and ellipticity of the $i$-th PDL element. The efficiency of the warped PDF in speeding up the OP estimation depends on the ability of the fast system to capture the true system behavior.

The crucial point of SS is the choice of the partition of the input sample space. Optimal partitioning requires finding strata having observables with zero standard deviation within each stratum, so that ideally one sample is sufficient to estimate (5.2) on each stratum. To get close to this ideal target, we suggest to stratify the slow RVs sample space according to the way PDL builds up along the link, rooting our idea on the following observations. Assuming for instance the basic equalized system of Fig. 5.2, where $M_{1,2}$ represent the PDL matrices of the first and second half of the link while $n_{1,2}$ are additive independent noise sources, the output signal is:

$$y = x + M_1^{-1} n_1 + (M_2 M_1)^{-1} n_2 \,. \tag{5.5}$$

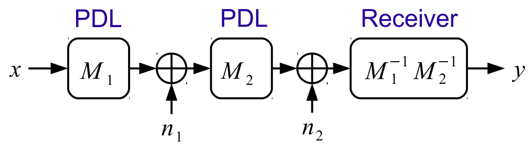


Figure 5.2: Basic transmission corrupted by PDL and noise.

The total PDL $\rho^t$ is a function of the eigenvalues of $M_1 M_2$ [9]. Such a PDL determines the contribution of $n_2$ to the variance of $y$, however it does not provide information about the contribution of $n_1$. Such information is given by the PDL of $M_1$ (referred to as $\rho^h$), i.e., by the eigenvalues of $M_1$, which give the PDL cumulated

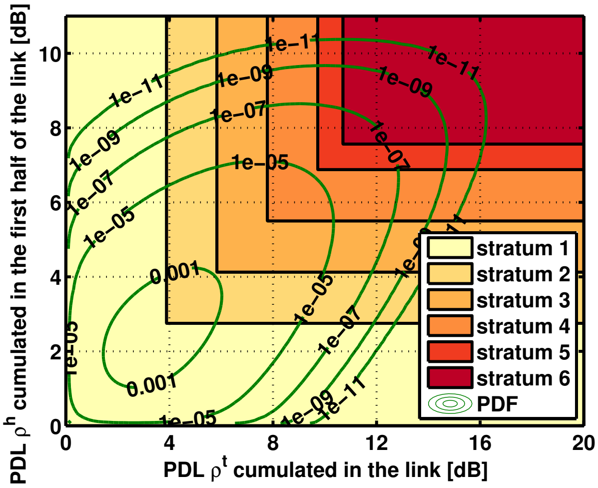


Figure 5.3: Sketch of SS idea. We divide the $(\rho^h, \rho^t)$ plane into $L$ frames (strata) identified by multiples of the median value of the corresponding PDL. We then run an MC simulation in each stratum, by increasing the number of samples in frames showing poorer accuracy. Contour levels represent an example of the joint PDF of PDL $(\rho^h, \rho^t)$, for $N = 30$ elements having PDL $\rho_i = 0.8$ dB, $i = 1, \ldots, 30$.

in the first half of the link. By considering $n_{1,2}$ as an additive distortion, either due to ASE or nonlinearity, we can summarize the impact of PDL on this basic two-section system by the pair $\left(\rho^h, \rho^t\right)$. In a more general multi-span optical system, for the sake of simplicity we may think to virtually divide it in two halves, and proceed as before. Of course this idea is not fully comprehensive of the impact of PDL. For instance, the nonlinear effect contributed by the first half might be correlated with that in the second half. Nevertheless, stratifying the sample space in terms of the values taken by the pair $(\rho^h, \rho^t)$ is expected to be a reasonable SS strategy. Hence, in our case the fast system was a two-trunk approximation of the true link, yielding a control variable $C = (\rho^h, \rho^t)$.

We proceeded by subdividing both the $\rho^h$ and $\rho^t$ axes into regular grids, thus stratifying the sample space in frames as depicted by Fig. 5.3. In particular, we used the discretization $\eta_{h,t} = \rho^{h,t}/\rho_m^{h,t} = [1, 1.5, 2, 2.5, 2.75]$, where $\rho_m^h$ and $\rho_m^t$ are the median values of the PDL after propagation in the first half and in the total link, respectively. This choice leads to $L = 6$ different strata. As an example, in Fig. 5.3 we also reported the contour levels of the joint probability density function (PDF) $f_{\mathrm{HT}}(\rho^h,\rho^t)$ of $N = 30$ elements, each having PDL [9] $\rho_i = 0.8$ dB. The probability $P_l$ that a random PDL realization falls in stratum $l$ is given by the integral of $f_{\mathrm{HT}}(\rho^h,\rho^t)$ over the corresponding frame (stratum). Since $\mathrm{OP}(l)$ is the outage probability of stratum $l$, the system outage probability is $\mathrm{OP} = \sum_{l=1}^{L} P_l \cdot \mathrm{OP}(l)$.

As each PDL realization is associated with a stratum of the sample space, a standard MC algorithm to estimate outage events would sample each stratum according to its probability of occurrence. This is strongly inefficient as we are interested in rare events of PDL. The proposed SS algorithm samples instead each stratum on the basis of the stratum uncertainty. The law establishing the optimal frequency of visits for the stratum $l$ depends also on $P_l$ [101], which is non trivial to compute due to the expression of $f_{\mathrm{HT}}(\rho^h,\rho^t)$. Using the total probability law we have

$$P_l = \iint_{\mathscr{S}_C(l)} f_{\mathrm{HT}}(h,t)\mathrm{d}h\mathrm{d}t = \iint_{\mathscr{S}_C(l)} f_{\mathrm{T}}(t\,|\,H=h) f_{\mathrm{H}}(h)\mathrm{d}h\mathrm{d}t \tag{5.6}$$

where $\mathscr{S}_C(l)$ is the frame identifying stratum $l$ in the space $(h,t) = (\rho^h,\rho^t)$. It can be proved that for a finite number $N/2$ of PDL elements, each having PDL $\rho_i$, it is [8]:

$$f_{\mathrm{H}}(h) = \frac{2h}{\pi}\int_0^{\infty} r\sin(hr)\,\mathrm{Sa}^{\frac{N}{2}}(\rho_i r)\,\mathrm{d}r \tag{5.7}$$

with $\mathrm{Sa}(x) = \sin(x)/x$. For a chain of $N$ elements where the first $N/2$ are fixed, we can apply the same reasoning leading to (5.7) and recall that the chain of $N/2$ fixed elements can be collected into one equivalent element, thus obtaining:

$$f_{\mathrm{T}}(t \mid H=h) = \frac{2t}{\pi}\int_0^{\infty} r\sin(tr)\,\mathrm{Sa}(hr)\,\mathrm{Sa}^{\frac{N}{2}}(\rho_i r)\,\mathrm{d}r\,. \tag{5.8}$$

A closed-form expression of (5.7) was first reported in [8]:

$$f_{\mathrm{H}}(h) = \frac{h}{2\rho_i^2\left(\frac{N}{2}-2\right)!}\sum_{s=0}^{\left\lfloor\frac{Nm}{2}\right\rfloor}(-1)^s\binom{N/2}{s}\left(\frac{N}{2}m-s\right)^{\frac{N}{2}-2} \tag{5.9}$$

where $m = \frac{1}{2}\left(1 - \frac{2h}{N\rho_i}\right)$. An expression of (5.8) can be found after expanding $\sin(tr)\sin(hr)$ by the prosthaphaeresis formula, obtaining:

$$f_{\mathrm{T}}(t \mid H = h) = \frac{t}{\pi h \rho_i}\left[G\left(\frac{t-h}{\rho_i}, \frac{N}{2}\right) - G\left(\frac{t+h}{\rho_i}, \frac{N}{2}\right)\right] \tag{5.10}$$

where

$$G(\delta, N) \triangleq \int_0^{\infty} \cos(\delta y)\left(\frac{\sin(y)}{y}\right)^N \mathrm{d}y\,. \tag{5.11}$$

A closed form expression of (5.11) for $N > 1$ is obtained as [102]:

$$G(\delta,N) = \frac{\pi/2}{(N-1)!}\sum_{s=0}^{\lfloor \frac{N-1}{2} \rfloor}(-1)^s\binom{N}{s}\left\{\pm\left(\frac{N}{2} - s + \frac{\delta}{2}\right)^{N-1}\right.$$
$$\left. \pm\left(\frac{N}{2} - s - \frac{\delta}{2}\right)^{N-1} + \mu(-1)^{\frac{N}{2}}\binom{N}{\frac{N}{2}}\left|\frac{\delta}{2}\right|^{N-1}\right\} \tag{5.12}$$

where: the positive sign is to be taken in front of a bracket if the expression within the bracket is itself positive and the negative sign in the opposite case; $\mu = 1$ if $N$ is even, 0 otherwise. Once the PDFs (5.9), (5.10) are available, $P_l$ in (5.6) can be found by numerical two-dimensional quadrature integration.

Despite the analytical closed-form formulas, both expressions (5.9) and (5.12) are based on ratios and differences of very large numbers when $N \gg 1$ that are difficult to evaluate. In such cases, quadrature integration of (5.7) and (5.8), e.g. by the function quadgk of MATLAB, may be numerically more stable. Alternatively, following the same steps of [8], one may resort to the asymptotic approximation for large $N$:

$$f_{\mathrm{T}}(t \mid H = h) \sim \frac{t}{h\sqrt{N}\rho_i}\sqrt{\frac{3}{\pi}}\left(e^{-\frac{3(h-t)^2}{N\rho_i^2}} - e^{-\frac{3(h+t)^2}{N\rho_i^2}}\right) \tag{5.13}$$

while (5.9) can be approximated by a Maxwellian distribution [8]. A check of the correctness of (5.9)-(5.10) for $N = 10, 30$ and $\rho_i = 0.5$ dB is reported in Fig. 5.4. We chose a random setup with $\rho^h = 0.54$ dB and $\rho^h = 0.76$ dB for the $N = 10$ and $N = 30$ cases, respectively.

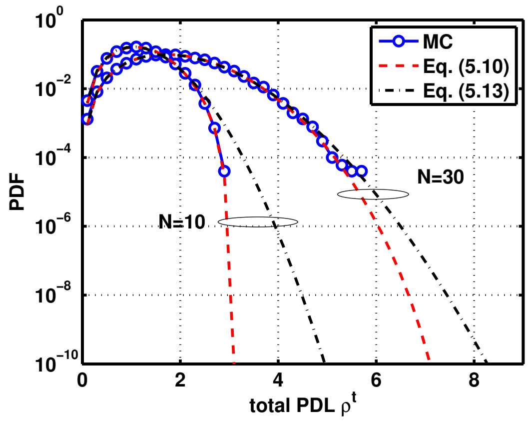


Figure 5.4: PDF $f_{\mathrm{T}}(t \mid H = h)$ of the total PDL conditioned to a random choice of the PDL cumulated in the first half of the link. Circles: MC simulations. Dashed lines: exact solution (5.10). Dash-dotted lines: asymptotic solution (5.13) for large $N$.

A pseudo-code of the SS algorithm is reported in Fig. 5.5. Before starting the SS algorithm, we performed short (transient) MC true-system runs in each stratum to roughly estimate the standard deviation $\sigma_l$, $l = 1, \ldots, L$ (vector "sigma" in Fig. 5.5). A stratum is explored by generating PDL from the unwarped PDF and accepting only samples within that stratum. We started the transient from stratum 1 and stopped it at the first stratum in which we counted five outages (OUTthr = 5 in Fig. 5.5). If less than 5 outages were counted in stratum $l$ after collecting $N_{\max}(l)$ samples, we moved to the next stratum $l+1$. After the transient, we ran the true SS algorithm, and updated the accept-reject probability $W(l)$, eq. (5.4), at the end of each BER estimation.

Since the estimated standard deviation of the outage probability is available, the SS simulation can be stopped with the same criterion as MC simulations, i.e., when the standard deviation of outage events is a specific fraction of the estimated outage probability.

```
define strata;               % L strata (see, e.g., Fig. 5.3)
P = [P1, P2, ..., PL];       % strata prob. from (5.6)
N = zeros(1,L);              % strata number of visits: vector
OUTcount = zeros(1,L);   % number of outage events
l = 1;                       % stratum index
while max(OUTcount) < OUTthr     % run transient.
   generate unwarped PDL in stratum l;  % (reject if not in l)
   Q-factor = Q_estimate(PDL); % inner cycle over fast RVs
   N(l) = N(l)+1;      % update visit counter of stratum l
   if Q-factor < FECthr,
      OUTcount(l) = OUTcount(l)+1;
   end
   if N(l) == Nmax(l),
      l = l+1; % move to next stratum
   end
end     % end transient.
OPstrata = OUTcount./N;     % vector of conditional OPs
OPest = sum(P.*OPstrata);              % System OP (scalar)
sigma = sqrt(OPstrata.*(1-OPstrata));     % std of strata OP
OPdev = sqrt(sum(sigma.*P)^2/sum(N)); % std(OPest) [98, p. 300]
ACC_OP = OPdev/OPest;                % estimation accuracy
while ACC_OP >= targetACC % outer cycle (slow RVs)
   P* = sigma.*P/sum(sigma.*P);      % update eq. (5.3)
   generate warped PDL % (accept-reject in Fig. 5.1)
   Q-factor = Q_estimate(PDL);
   update N(l),sigma(l),OPest,OPdev,ACC_OP; % like above
end

function Q-factor = Q_estimate(PDL) % true-system BER (5.1)
   ACC_Q = Inf;               % sample-estimated Q accuracy
   while ACC_Q >= targetACCerg % inner cycle (fast RVs)
      generate fast RVs;         % ASE and symbols
      run SSFA;
      estimate Q-factor;                 % pure MC
      update ACC_Q
   end
return
```

Figure 5.5: MATLAB-style pseudo-code of adaptive SS algorithm.

## 5.3 SS-estimated OP

We applied the SS strategy described in the previous section to both a DM and a DU link, each in presence or absence of PMD. In the DM case the optical link was composed of $N = 30$ spans, each of length 100 km, attenuation 0.2 dB/km, dispersion 4 ps/nm/km, nonlinear index $\gamma = 1.5$ 1/W/km and residual dispersion per span 30 ps/nm. Before transmission we added a pre-compensation of -495 ps/nm. The DU link was identical, except for a total of $N = 35$ spans with neither in-line compensation nor pre-compensation. The overall dispersion was set to zero by the receiver dispersion equalizer.

When present, PMD was emulated through 50 random waveplates per transmission fiber and a PMD coefficient of $0.13\,\mathrm{ps}/\sqrt{\mathrm{km}}$. PDL was emulated by lumped elements placed before the transmission fibers. The nonlinear signal-noise interaction along the link was accounted for by flat-gain noisy amplifiers with 6 dB noise figure. The WDM comb consisted of 15 channels with 32Gbaud PDM-QPSK modulation, with 50 GHz spacing. In the DM case we used for each PDM-QPSK channel 1024 purely random symbols, while 2048 were used in the DU case [103,104]. Fiber propagation used the SSFA applied to the Manakov-PMD equation [12]. At the receiver, coherent detection of the central channel used a digital signal processor [43] including: analog to digital conversion with bandwidth 17 GHz, 2 samples-per-symbol data-aided least squares butterfly equalization with 7 taps, and Viterbi and Viterbi phase estimation with 15 taps.

The performance was measured in terms of the Q-factor obtained by inverting the BER. We counted at least 400 errors for each BER estimation. An outage event was declared when the Q-factor was below the FEC threshold of 6.25 dB, i.e., when the BER was above $2 \cdot 10^{-2}$. The OP was estimated after observing at least 25 outage events.

Fig. 5.6 shows the average Q-factor of the central of 15 WDM channels versus signal power. The average Q-factor was obtained with the standard MC algorithm, by averaging over 10 seeds of slow RVs. In the DM scenario, Fig. 5.6 (left), a rms PDL of 2.74 dB reduces the best Q-factor by about 0.3 dB in both absence and presence of

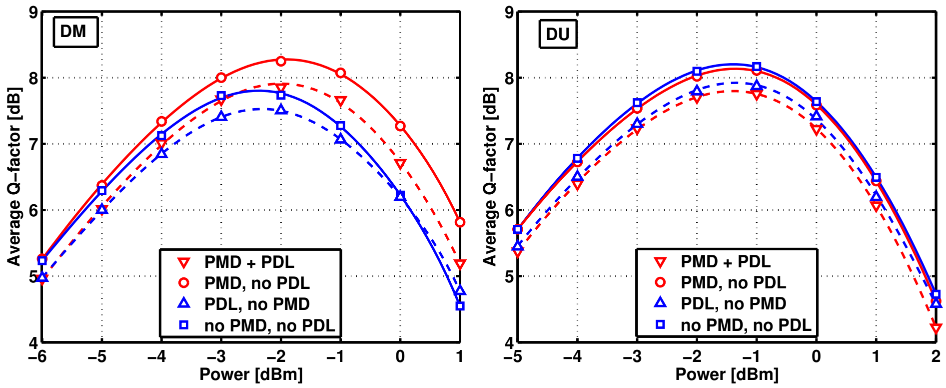


Figure 5.6: Average Q-factor of a 15-channel 32Gbaud PDM-QPSK system, in presence/absence of PMD = 0.13 ps/$\sqrt{\text{km}}$ and rms PDL 2.74 dB. (Left) 30 × 100 km DM link. (Right) 35 × 100 km DU link.

a typical PMD of 0.13 ps/$\sqrt{\text{km}}$. It is worth noting that, without PMD, the detrimental effect of PDL on the average Q-factor vanishes at large powers, as a consequence of the XPolM reduction induced by PDL. When PMD is present, the best Q-factor improves by 0.4 dB because of the polarization decorrelation induced by PMD. However, now the beneficial effect of PDL at large powers disappears, thus making the PDL penalty on the average Q slightly increase with the input power.

Fig. 5.6 (right) shows the DU case. Now, in absence of PDL, PMD does not affect the average performance because of the large amount of dispersion cumulated along the link. By turning PDL on, we observe that the PDL penalty on the Q-factor is confined within 0.4 dB both in presence and absence of PMD.

Hence by just looking at the average Q-factor we would erroneously conclude that PDL/PMD have a marginal effect on system performance. We must instead look at the OP, which will be estimated at a power 1 dB beyond that at the best average Q-factor.

In order to test the proposed SS algorithm and get a feeling of its computational gain with respect to standard MC, we first applied it to a linear system with $\rho^t_{\text{rms}} = 3.3$ dB, no PMD, at an OSNR of 12.8 dB over a bandwidth of 0.1 nm. We performed 20

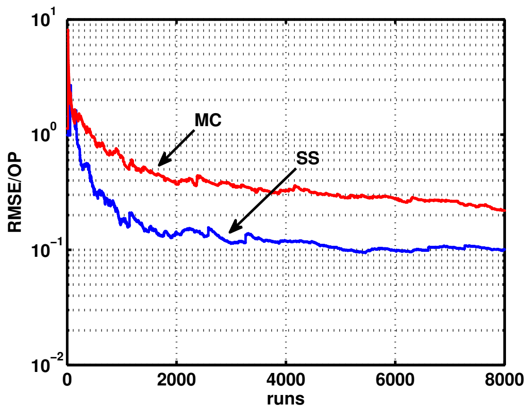


Figure 5.7: OP Root-Mean-Square Error normalized to true OP versus number of runs for MC and SS algorithms, at $\text{OP} = 2.3 \cdot 10^{-3}$, with $\rho^{t}_{\text{rms}} = 3.3$ dB and OSNR = 12.8 dB/0.1 nm. Linear propagation without PMD.

parallel MC simulations of 8000 Q-factor samples (runs) each. The MC estimation of the OP, averaged over all the 160,000 Q-samples, was $2.3 \cdot 10^{-3}$ and was declared the *true* value. Moreover, while the 20 simulations were in progress, we also measured the running OP estimate, and calculated the OP mean-square error (MSE) w.r.t. the true OP after each run. We then ran 20 SS simulations and again computed the running OP MSE. Please note that with SS we also accounted for the initial transient in the number of runs.

Fig. 5.7 shows the OP root-MSE (RMSE) over true OP versus the number of runs. We observe that the SS algorithm slashes the simulation time by more than a factor 4 compared to the MC algorithm at an RMSE/OP value of $4 \cdot 10^{-1}$, i.e., at MSE $10^{-6}$, and the gain in efficiency increases at lower RMSE.

Having verified the SS efficiency in a simple case where the *true* OP was computable with great accuracy, we next applied the SS algorithm to a realistic nonlinear DM transmission. Fig. 5.8 shows with solid lines the SS-estimated OP versus PDL, both without and with PMD (rms DGD of 7.12 ps) and at transmitted power $P = -1.5$ dBm. For all SS simulations we used the values $\eta_{h,t} = [1, 1.5, 2, 2.5, 2.75]$, which

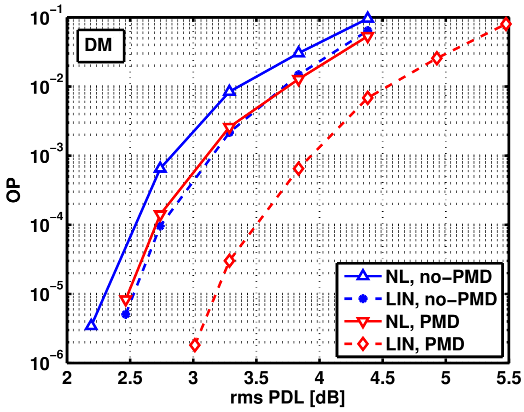


Figure 5.8: OP versus PDL rms of a 15-channel 32Gbaud PDM-QPSK system in a $30 \times 100$ km DM link with dispersion 4 ps/nm/km, in presence/absence of PMD$=$ $0.13$ ps/$\sqrt{\text{km}}$. $P = -1.5$ dBm and OSNR $= 15.7$ dB/0.1 nm.

define $L = 6$ strata, with the heuristic choice $N_{\max}(l) = l \cdot 100$ for $l = 1, \ldots, L$ as a compromise between computational speed and accuracy. It is worth noting that, in absence of PMD, the tolerable PDL to achieve an OP below $10^{-3}$ is $\rho^t_{\text{rms}} = 2.8$ dB, i.e., the PDL generated by 30 lumped elements with typical 0.5 dB of PDL each [99]. When the target is instead an OP below $10^{-5}$, the tolerable rms PDL becomes 2.3 dB. The presence of 0.13 ps/$\sqrt{\text{km}}$ of PMD significantly lowers the OP in absolute terms, but improves by only 0.2 dB the PDL tolerance at fixed OP.

We also tried to see if the true OP values could be estimated by an equivalent linear system [91]. Hence we turned off fiber nonlinearity and set the amplifiers noise figure so as to have the same average Q-factor as the nonlinear system in absence of PDL, i.e., $Q_{\text{avg}} = 7.6$ dB without PMD and $Q_{\text{avg}} = 8.2$ dB with PMD, as shown in Fig. 5.6 (left). By this way, if PDL did not interact with Kerr nonlinearity, we would observe the same Q-factor distribution, hence resulting in the same OP as the nonlinear system. However, as shown by the dashed curves in Fig. 5.8, this does not happen, thus implying a significant coupling of PDL and fiber nonlinearity that rules out the use of an equivalent linear system to guess the true OP in presence of nonlinearity.

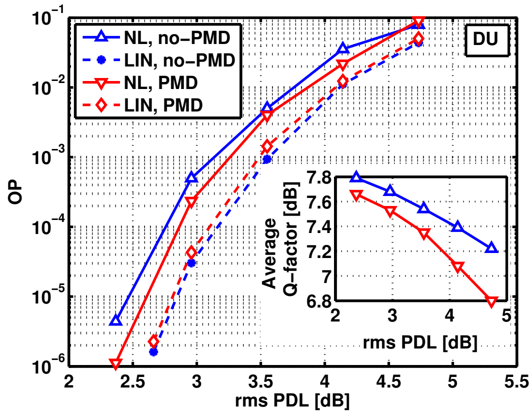


Figure 5.9: OP versus rms PDL of a 15-channel 32Gbaud PDM-QPSK system in a $35 \times 100$ km DU link with dispersion 4 ps/nm/km, in presence/absence of PMD$=$ $0.13$ ps$/\sqrt{\text{km}}$. $P = -0.5$ dBm and OSNR $= 16.1$ dB/0.1nm. Inset: Average Q-factor versus rms PDL.

We finally moved to estimate the OP in the DU scenario. Fig. 5.9 shows with solid lines the SS-estimated OP versus PDL, at an rms DGD both zero and 7.69 ps, and power $P = -0.5$ dBm. It can be noted that, in absence of PMD, the tolerable rms PDL to achieve an OP below $10^{-5}$ is $\rho^t_{\text{rms}} = 2.45$ dB, and the presence of PMD along the link increases this value by 0.1 dB. Dashed lines in Fig. 5.9 depict the OP predicted by the linear equivalent system, which significantly underestimates the true OP.

We can see that the presence of PMD along the link lowers the OP notwithstanding a lower average Q-factor, as shown in the inset of Fig. 5.9. This fact suggests that the variance of the Q-factor is reduced in presence of PMD. Please note that all OPs below $10^{-3}$ were SS-estimated with a number of Q-samples varying between 2500 and 10000, after observing a number of outage events varying between 25 and 200. It is worth noting that, when using standard MC, observing 25 outage events with 10000 Q-samples corresponds, on average, to an estimated value

$\mathrm{OP} = 25/10000 = 2.5 \cdot 10^{-3}$.

The proposed SS strategy does not control directly the PDL of each span, but just their cumulative behavior by stratifying the sample space of $\rho^t$ and $\rho^h$ in a few strata, whatever the link length. This way, it is possible to run efficient simulations even when many PDL elements are present along the link, thus by-passing the typical dimensionality problem of IS [105–107]. A comparison of the proposed algorithm with other IS-based techniques is out of the scope of this work. However, we remark the intrinsic simplicity of the code, the ability to work even in the nonlinear regime, and its efficiency at practical levels of OP.

It is worth noting that the efficiency of SS is maximum when the set of outage events looks like a rectangle in the $(\rho^h, \rho^t)$ plane. In this case, the SS algorithm is essentially a stochastic search of the boundary of such a rectangle. If additional information about the shape of the set of outage events is available, the stratification can be changed from frames to other shapes better fitting the expected contour of outage events.

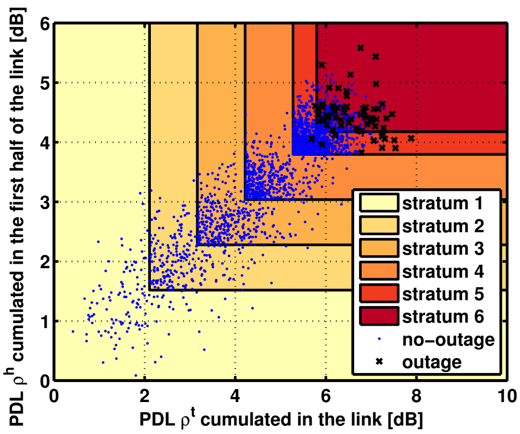


Figure 5.10: Distribution in the $(\rho^h, \rho^t)$ plane of SS-driven 2500 PDL realizations, from which we computed the OP in Fig. 5.9 for the nonlinear DU system without PMD at $\rho^t_{\mathrm{rms}} = 2.4$ dB. Cross: outage. Dot: no-outage.

For the nonlinear DU system without PMD used to compute the OP in Fig. 5.9

at $\rho^t_{\text{rms}} = 2.4$ dB, Fig. 5.10 shows the distribution in the $(\rho^h, \rho^t)$ plane of SS-driven PDL realizations. Crosses/dots show PDL realizations inducing/not-inducing outage. We see that the set of PDL realizations leading to outage overlaps pretty well with only two strata. However, the OP per stratum is not always 1 or 0, as required by the optimal zero-variance IS estimator. Insight about this behavior can be obtained by looking at eq. (5.5). Intuitively, outage events occur for "large" PDL realizations, identified by the eigenvalues of $M_1^{-1}$ and $(M_2M_1)^{-1}$ and captured by the proposed $(\rho^h, \rho^t)$ stratification. However, an uncertainty still remains because of the eigenvectors' random orientation, whence the gap between the proposed algorithm and the optimal IS.

## 5.4 Conclusions

We presented an adaptive stratified-sampling algorithm to efficiently asses the PDL-induced OP in presence of fiber nonlinearity and PMD along the link. We showed the effectiveness of the proposed algorithm compared to standard Monte Carlo, with a factor 4 savings in computation time. We applied the proposed algorithm to both DM and DU links, in presence and absence of PMD, by estimating the OP down to $10^{-5}$. Our results show that speeding up simulations by considering a linear equivalent system with the same margin over FEC threshold as the true system leads to a significant underestimation of the actual OP.

# Conclusions

In this thesis we shed light on multiple aspects of modern coherent optical transmissions. In particular, we analyzed the optimal granularity of single carrier WDM system, i.e., the symbol rate maximizing the achievable distance while assuring a desired BER, in several scenarios accounting for most the popular coherent modulation formats, in both DM and DU links. We also analyzed the impact of the PDL effect on coherent detection-based systems, highlighting the interplay between PDL and fiber nonlinearity and its impact on both the average BER and its distribution of system performance. Thanks to the possibility in the Optilux software to selectively activate the nonlinear terms of the propagation equation related to the main manifestations of fiber nonlinearity, namely, SPM, XPM and XPolM, we exploited this opportunity to highlight: i) the dominant nonlinearity as the granularity of the system changes; ii) the interaction of each nonlinear effect with PDL.

In chapter 2 we quantified, at fixed bandwidth efficiency, the constrained Nonlinear Threshold (NLT) as the per-channel symbol rate varied, in both DM and DU links, for the following modulation formats: i) NRZ and iRZ PDM-BPSK; ii) NRZ and iRZ PDM-QPSK; iii) NRZ PS-QPSK; iv) NRZ PDM-16QAM. We also provided the NLT due to each manifestation of Kerr effect, i.e., SPM, XPM and XPolM, highlighting the dominant nonlinearity as the symbol rate changed. In DM systems we also showed the importance of signal-noise interaction for each effect. In standard single-mode fiber DU systems we found, for each modulation format, that in a 15 channel system, the SPM dominates over XPM or XPolM at symbol rates above 15Gbaud. In DM systems we observed SPM as dominant nonlinearity at high symbol rates while cross-channel nonlinearities are dominant at lower rates, with the cross-point varying with the modulation format. In the low symbol rate range, we observed XPolM nonlinearity as dominant for all NRZ PSK formats and XPM slightly dominant over XPolM for the other considered formats. By presenting the results as the NLT divided by the symbol rate, we identified the optimal granularity of the system. We observed that in DU links the symbol rate that maximizes the reach at the target BER $= 10^{-3}$ is the lowest considered, i.e., 5Gbaud. In DM links, instead, the optimal symbol rate lays in the 10-40Gbaud range, according to the modulation format, with shallow reach variations within this range. For DU systems, we compared NLT results against the available Gaussian Noise theory, providing theoretical extrapolations to practically large WDM systems which cannot be simulated. We provided a new, simple formula for the reach prediction using side data of NLT estimations. Finally, for (NRZ) BPSK, QPSK, 16QAM formats and selected symbol rates, we performed reach simulations. By this way: i) we provided the nonlinear interference (NLI) accumulation parameter; ii) we verified reach predictions of the proposed formula, finding good accuracy in both DM and DU systems; iii) we verified the prediction on the optimal granularity by NLT results, finding correct forecasts.

In chapter 3 we analyzed the statistics of the nonlinear interference (NLI) distortions in a DU network scenario. Since interfering channels may propagate over different lightpaths before propagating together with the reference channel, we demonstrated that their upstream-cumulated GVD is the critical parameter to assess the NLI distribution. By considering a PDM-QPSK modulation format, we showed that the clouds of the received constellations may not be circular as suggested by first order perturbative models. In fact, in case of non-zero upstream-cumulated GVD of neighboring channels and zero pre-dispersed reference channel, the NLI distortions coming from downstream propagation become non-circular. In particular, we showed that while the NLI quadrature variance is increased, the in-phase variance is unaltered, giving the NLI a phase-noise feature. We showed that the NLI quadrature variance increases for increasing values of neighbors upstream-cumulated GVD. This phenomenon is reduced when the channel of interest experiences a large cumulated dispersion, hence is more evident on low dispersion fibers and at small propagating distances. We observed that the increase of the NLI quadrature variance produces a reduced penalty on the best Q-factor performance, despite a sizable impact in the deeply nonlinear regime.

In chapter 4 we analyzed the impact of PDL on systems with PDM-QPSK transmission and coherent detection. By considering pathological cases of PDL, i.e., the eigenvector of each PDL element aligned with respect to others, we showed that in the linear regime the most detrimental case is represented by PDL aligned with one of the two polarization tributaries. By investigating the same pathological cases of PDL, we highlighted the interplay of PDL with each individual nonlinear effect, i.e., SPM, XPM and XPolM in three scenarios: i) transmission over a DM link of a homogeneous WDM system with 100 Gb/s PDM-QPSK channels; ii) transmission over a DM link of a WDM comb with a 100 Gb/s PDM-QPSK channel surrounded by 10Gb/s OOK channels, giving a hybrid system; iii) transmission over a DU link of a WDM system with 100 Gb/s PDM-QPSK channels. We found, for the homogeneous DM system, that in the deeply nonlinear regime SPM and XPM are enhanced by PDL-induced peak-power fluctuations, while XPolM is reduced because of minor patter-induced temporal fluctuations of the SOP. Overall the beneficial effect on XPolM is more than compensated for by the detrimental effect on SPM and XPM. In both the DM hybrid and DU homogeneous, the inherent peak-power fluctuations of neighboring channels reduce the importance of PDL-induced fluctuations, while XPolM is now a secondary impairment and does not show improvements with PDL. Moving to analyze the impact of realistic PDL configurations, we found that the penalty on average the Q-factor performance was confined within 0.3 dB. However, notwithstanding the negligible average penalty, we observed a significant impact on the Q-factor distribution. In fact, by operating 1 dB above the optimal power, we observed that the presence of PDL remarkably enlarges such a distribution, leading to outage events, while in absence of PDL the system OP was essentially zero. Besides, we found that the interplay between PDL and nonlinearity yields a Q-factor variance at least 1.4 times larger with respect to the PDL-only case.

In chapter 5 we presented an adaptive Stratified-Sampling algorithm to efficiently assess the PDL-induced OP. The proposed algorithm splits the set of all possible PDL configurations into subsets according to how PDL builds up along the link. Separate Monte Carlo OP estimations are performed in each subset. The SS method allocates samples to each individual MC estimation according to the subset contribution to the overall OP. By comparing the proposed algorithm with the standard Monte Carlo method, the new SS algorithm achieved a factor 4 of savings in computational time on a purely linear system with known OP of $2.3 \cdot 10^{-3}$. Later, we applied the proposed algorithm to practical systems, estimating the OP down to $10^{-5}$ in presence of fiber nonlinearity in both DM and DU links. We showed that the presence of PMD along the link significantly lowers the OP in absolute terms, but provides only a slight improvement on the PDL tolerance at fixed OP. Finally, we demonstrated that considering a linear equivalent system with the same margin over FEC threshold as the true nonlinear system leads to a significant underestimation of the actual OP.

In the Appendix B we highlighted possible pitfalls in the simulation of WDM systems. In particular we described the correct procedure to estimate the Q-factor distribution, showing how naive expedients to speed up the estimation may introduce significant overestimation of the distribution variance. We also provided correct dimensioning of the sequence length and warnings on the length of the Split-Step Fourier algorithm nonlinear step.

# Appendix A

# Stokes Representation

We first define four matrices:

$$\sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \sigma_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad \sigma_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \sigma_3 = \begin{bmatrix} 0 & -j \\ j & 0 \end{bmatrix} \tag{A.1}$$

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 $\underline{a} = [a_0, a_1, a_2, a_3]^{\mathrm{T}}$, by defining $a_i = \overrightarrow{A}^{\dagger} \sigma_i \overrightarrow{A}$, 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:

$$\overrightarrow{A} = \begin{bmatrix} A_x \\ A_y \end{bmatrix} = \begin{bmatrix} |A_x| \, \mathrm{e}^{j\phi_x} \\ |A_y| \, \mathrm{e}^{j\phi_y} \end{bmatrix} = \mathrm{e}^{j(\phi_x+\phi_y)/2} A \widehat{\underline{A}}$$

[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 2 1. Our labeling is conventional for the Poincaré representation.

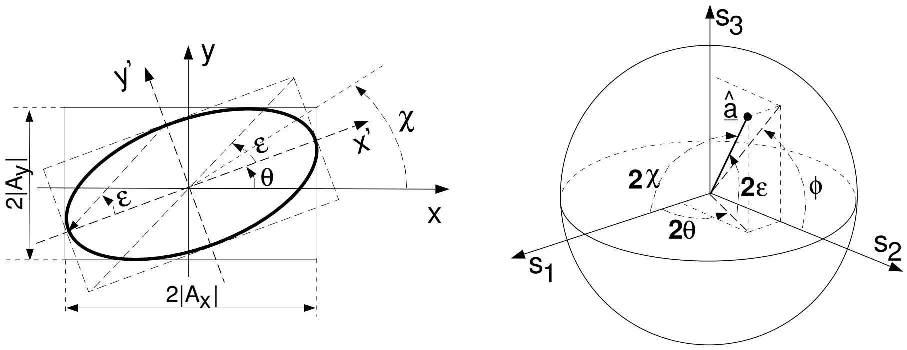


Figure A.1: (Left) Jones SOP parameters; (Right) Stokes SOP parameters on Poincaré sphere.

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

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

Is thus clear that the representation of the field $\overrightarrow{A}$ is faithful up to a common, arbitrary phase factor, i.e., the family of fields $\overrightarrow{A}\mathrm{e}^{j\delta}$, $\delta$ real, all map into the same vector $\underline{a}$. The *Poincaré representation* of the field $\overrightarrow{A}$ is the vector $\overrightarrow{a} = [a_1, a_2, a_3]$, which can be considered as the projection on the 3D space of the 4D Stokes vector $\underline{a}$. The *Poincaré representation* of its Jones SOP vector $\underline{\widehat{A}}$ is the unit vector $\underline{\widehat{a}} \triangleq [a_1, a_2, a_3]/a_0$, which can be plotted as a point on the three-dimensional unit sphere, the so-called Poincaré sphere.

Fig. A.1 shows the relation between the polarization ellipse [108] giving the parameters of the Jones SOP vector $\underline{\widehat{A}}$ and the Poincaré unit vector $\underline{\widehat{a}}$, also called the *Stokes SOP vector*. Another equivalent expression of the Jones and Stokes SOP vectors is:

$$\underline{\widehat{A}} = \begin{bmatrix} \cos\theta\cos\varepsilon - j\sin\theta\sin\varepsilon \\ \sin\theta\cos\varepsilon + j\cos\theta\sin\varepsilon \end{bmatrix} \iff \underline{\widehat{a}} = \begin{bmatrix} \cos 2\theta \cos 2\varepsilon \\ \sin 2\theta \cos 2\varepsilon \\ \sin 2\varepsilon \end{bmatrix}$$

where azimuth and elevation angles $2\theta$, $2\varepsilon$ are also indicated in Fig. A.1. Note the 2 multiplicative factor of $\theta$ and $\varepsilon$ when moving between Jones and Stokes spaces. Note also that we can write $\underline{\widehat{A}} = \underline{\underline{U}}(\theta) \begin{bmatrix} \cos\varepsilon \\ j\sin\varepsilon \end{bmatrix}$, where by $\underline{\underline{U}}(\alpha)$ we indicate the matrix that operates a *counterclockwise* rotation by an angle $\alpha$:

$$\underline{\underline{U}}(\alpha) = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}.$$

In the figure, $S_1$, $S_2$ and $S_3$ 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 the assigned Stokes SOP vector $\underline{\widehat{r}} = [r_1, r_2, r_3]^{\mathrm{T}}$, we find its associated Jones SOP vector $\underline{\widehat{R}}$, as well as its orthogonal $\underline{\widehat{R_o}}$ associated[2] with $-\underline{\widehat{r}}$ as:

$$\begin{cases} \underline{\widehat{R}} = \begin{bmatrix} 1 + r_1 \\ r_2 + jr_3 \end{bmatrix} \frac{1}{\sqrt{2(1+r_1)}} \\ \underline{\widehat{R_o}} = \begin{bmatrix} -r_2 + jr_3 \\ 1 + r_1 \end{bmatrix} \frac{1}{\sqrt{2(1+r_1)}} \end{cases}.$$

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

[2] Both $\underline{\widehat{R}}$ and $\underline{\widehat{R_o}}$ are unique up to a common phase term. For instance, by adding a common phase rotation of $\pi$ to the second eigenvector one gets the equivalent form $\underline{\widehat{R_o}} = [r_3 - jr_2, -(1 + r_1)]^{\mathrm{T}}/\sqrt{2(1+r_1)}$.

# Appendix B

# Pitfalls in simulating WDM systems

In this appendix we provide warnings on possible pitfalls in the simulation of wavelength division multiplexed (WDM) systems. In particular we assess the correct dimensioning of pattern length to correctly account for channel walk-off, the length of the nonlinear step in the split-step Fourier algorithm (SSFA) and the procedure to correctly estimate the distribution of system Q-factor performance.

## B.1 Pattern length to correctly account for channels Walk-Off

Pattern length is one critical parameter to correctly account for cross-channel nonlinearity when simulating WDM transmission systems. Because of fiber chromatic dispersion, as described in Sec. 1.2.2, pulses at different wavelength propagate at different speed inside a fiber because of a mismatch in their group velocities. The net result is that patterns of different channels walk past each other during propagation. Taking into account the circular property of the Fast Fourier Transform (FFT), which is used to efficiently implement the SSFA, we obtain a periodically repeated symbol pattern. In case of pattern length not sufficiently long to account for the largest walk-off experienced during propagation between the probe and the interfering channels, a probe symbol sees multiple times along its propagation the same interfering symbol. This introduces a correlation in the nonlinear distortion induced on the probe symbol that can alter the results. In absence of third order dispersion, the walk-off $w_{in}(z)$ between the $i$-th channel and the $n$-th one over a length of $z$ km is:

$$w_{in}(z) = (\omega_i - \omega_n) \int_0^z \beta_2(x) dx \,.$$

Hence, for a generic link the pattern length coping for the largest walk-off should be longer than [70, 75]:

$$N_{\text{symb}} \geq \left\lceil \frac{\Delta\lambda \cdot \left\lceil \frac{N_{\text{ch}}-1}{2} \right\rceil \cdot \big(D_{\text{tx}} \cdot L_{\text{span}} + \big(N_{\text{span}} - 1\big) \cdot \text{RDPS}\big)}{T_0} \right\rceil \tag{B.1}$$

where $D_{\text{tx}}$ $\left[\frac{\text{ps}}{\text{nm}\cdot\text{km}}\right]$ is the transmission fiber dispersion, $\Delta\lambda$ [nm] is the channel spacing, $L_{\text{span}}$ [km] is the span length, $N_{\text{span}}$ is the number of spans, RDPS [ps/nm] is the residual dispersion per span, $T_0$ [ps] is the symbol time and $\lceil\cdot\rceil$ is the ceiling function. Eq. (B.1) correctly takes into account of the walk-off between the middle channel of a WDM comb and the farthest edge channel. Since the required simulation time scales as $N_{\text{FFT}} \log(N_{\text{FFT}})$ with the size of the FFT window $N_{\text{FFT}}$ directly proportional to $N_{\text{symb}}$, to satisfy (B.1) for DU links may require exceedingly long sequences. In this case a trade-off between time consumption and numerical accuracy is necessary.

## B.2 Size of the SSFA Nonlinear step

In the open source software [4], the length of the nonlinear step of the SSFA is ruled by two parameters: $\Delta z_{\max}$ and $\Delta\Phi_{\max}$. $\Delta z_{\max}$ [m] is the maximum length of the nonlinear step, hence by setting $\Delta z_{\max} = L$ the length of the nonlinear step cannot exceed $L$ [m]. $\Delta\Phi_{\max}$ [rad] is the maximum allowed nonlinear phase rotation in the nonlinear step, set by default to $3 \cdot 10^{-3}$ [rad]. By fixing $\Delta\Phi_{\max}$, the length of the nonlinear step is adaptively chosen, thus allowing short steps in regions of high power (usually at the beginning of the fiber) and large steps in regions of low power, but always shorter than $\Delta z_{\max}$. $\Delta\Phi_{\max}$ is a critical parameter of the simulation since most of the times it determines the time-consumption of the simulation. The default value $3 \cdot 10^{-3}$ [rad] was chosen to ensure correct simulation in feasible time.

All results depicted in chapter 2 were obtained by setting $\Delta\Phi_{\max} = 3 \cdot 10^{-3}$ for symbol rates up to 60Gbaud, and $\Delta\Phi_{\max} = 6 \cdot 10^{-4}$ for symbol rates of 80 and 100Gbaud. This choice comes from considerations at point **13** in Sec. 2.2.3. We initially estimated $P_{\mathrm{NLT},1}$ with $\Delta\Phi_{\max} = 3 \cdot 10^{-3}$ for each symbol rate.

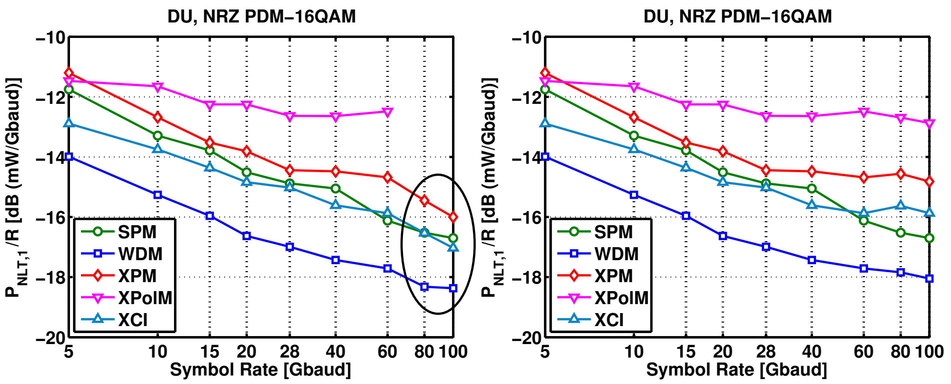


Figure B.1: Normalized margin $10\log_{10}(P_{\mathrm{NLT},1}/R)$ vs symbol rate $R$ [log scale] for a 15-channel homogeneous WDM comb with $R/\Delta f = 0.56$ over a 20x100km SMF DU link. ASE loading. Labels: “SPM”=self-phase modulation only; “XPM”=scalar cross-phase modulation only; “XPolM”=cross-polarization modulation only; “WDM”=all nonlinearities; “XCI”=XPM+XPolM. (Left) $\Delta\Phi_{\max} = 3 \cdot 10^{-3}$ [rad] for each symbol rate, (Right) $\Delta\Phi_{\max} = 6 \cdot 10^{-4}$ [rad] for 80 and 100Gbaud. Circle: anomalous behavior.

Fig. B.1(left) shows the results obtained for PDM-16QAM format in the DU link. We found a discrepancy between the behavior of the $P_{\mathrm{NLT},1}/R$ vs $R$ plot at high symbol rates and what predicted by [77]. Hence we more accurately estimated $P_{\mathrm{NLT},1}/R$ vs $R$ at 80 and 100Gbaud by using $\Delta\Phi_{\max} = 6 \cdot 10^{-4}$ (results shown in Fig. B.1(right)), now finding agreement with the GN theory predictions. We then adopted $\Delta\Phi_{\max} = 6 \cdot 10^{-4}$ whenever the symbol rate was above 60Gbaud on SMF. Since we cannot provide a general rule for the choice of $\Delta\Phi_{\text{max}}$, we suggest to tighten the constraint on $\Delta\Phi_{\text{max}}$ for very high values of the symbol rate [109, eq. (22)].

## B.3 Procedure to estimate the distribution of system performance

The estimation of the distribution of system BER, or equivalently of the Q-factor, is a very time-consuming non-trivial task. In this section we aim at demonstrating that apparently reasonable expedients to speed up the estimation may introduce significant overestimation of the distribution variance. As pointed out in Sec. 4.2.4, the averaging during BER measurements is used to summarize the impact of the system stochastic parameters. However, those disturbing parameters have different dynamics. For instance, ASE noise has very fast dynamics of the order of [THz], while PDL/PMD vary with frequencies of the order of [kHz] [96]. Hence, it is necessary to distinguish between slowly varying parameters that do not vary during a BER measurement, called *slow* random variables (RVs), and those that do, i.e., rapidly varying parameters called *fast* RVs. The distribution of the BER must account for the randomness of the *slow* RVs only, while the impact of the *fast* RVs must be averaged in the BER measurement. Among *slow* RVs we identify PDL, PMD and laser SOP, while ASE noise and information symbols represent *fast* RVs.

Thus the *correct* procedure to estimate the BER distribution requires two nested cycles: an inner one on *fast* RVs and an outer one on *slow* RVs. By this way, the outer cycle provides information about the drift of the BER due to slowly varying parameters, that eventually leads to out-of-service events occurring when the BER is larger than a reference value $\text{BER}_{\text{FEC}}$ set by the forward error correction (FEC) threshold. This approach is extremely time consuming, because in the inner cycle it requires to simulate several times the propagation of the WDM comb along the link to obtain a single BER sample (averaged over *fast* RVs until the desired accuracy is achieved). Hence we might be tempted to speed up the simulation through some expedients. For instance, in systems where the interaction signal-ASE noise is negligible, the simulation can be speeded up by loading, after a noiseless propagation, all the equivalent noise at the end of the link, which ensures the correct total ASE power spectral density. However, by this way the time gain with respect to a distributed noise simulation is limited. In noise loading simulations, we might be tempted to include information symbols among *slow* RVs in the outer cycle, with the consequence of a dramatically shorter inner cycle coming from a single noiseless propagation of the WDM comb and a relatively fast (with respect to signal propagation) noise loading cycle at the end of the link. We will refer to this approach as the *naive* procedure.

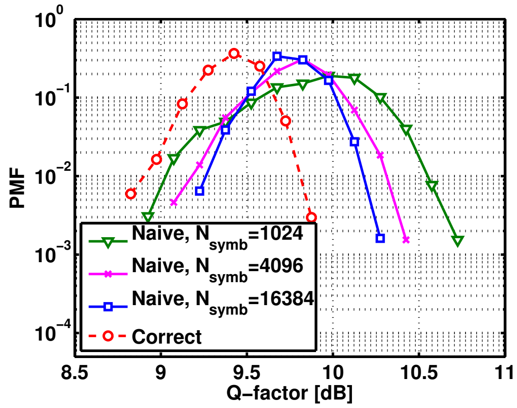


Figure B.2: Q-factor Probability Mass Function (PMF). Bin size 0.15 dB. 600 random seeds for *slow* RVs. BER samples estimated by Monte Carlo counting at least 100 errors. Homogeneous DM setup with Power -1 dBm. (Dashed) *Correct* procedure. (Solid) *Naive* procedure.

To give an idea of the overestimation on the distribution variance, we consider the homogeneous setup used to obtain the WDM labeled curve in Fig. 4.12(top-right). Hence we estimated the Q-factor PMF for a $20 \times 100$km DM link of non zero dispersion shifted fiber (NZDSF) (attenuation 0.2 dB/km, dispersion 4 ps/nm/km, nonlinear index $\gamma = 1.5$ 1/W/km, no PMD, no PDL) with residual dispersion per span of 30 ps/nm and Pre-compensation of -345 ps/nm. The amplifiers noise figure was 7 dB. The WDM comb was composed of 19 channels, all 112 Gb/s PDM-QPSK modulated, with 50 GHz spacing. Fig. B.2 depicts the estimated PMF at power -1 dBm

(1 dB above the optimal power as shown in Fig. 4.10). The dashed curve shows the Q-factor PMF estimated with the *correct* procedure, while solid curves show the estimation obtained with the *naive* procedure, for different sizes of the pattern length. It is worth noting that the *naive* procedure returns not only a 0.4 dB offset on the average Q-factor, but also a PMF with enlarged variance (up to 4 times for pattern length $N_{\text{symb}} = 1024$) with respect to the real PMF obtained with the *correct* procedure. We ascribe the 0.4 dB offset on the average Q-factor to the missing signal/noise interaction in the *naive* procedure, while reasons for the enlarged variance can mainly be ascribed to the wrong inclusion of symbol patterns into *slow* RVs, which increases the occurrence of rare events in both tails of the PMF.

Besides correctly dealing with input RVs, it is necessary to avoid the Q-factor distribution being corrupted by the MC uncertainty in the inner cycle on *fast* RVs. Since MC error is independent of the *slow* RVs, it substantially adds its variance to the true variance of the Q-factor distribution. Therefore we must ensure that the MC variance in the *fast* RVs inner cycle is much smaller than the variance of the overall distribution.

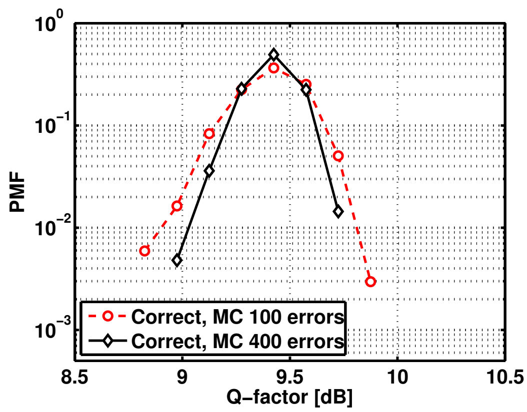


Figure B.3: Q-factor Probability Mass Function (PMF) estimated with the *Correct* procedure. Bin size 0.15 dB. 600 random seeds for *slow* RVs. BER samples estimated by Monte Carlo counting at least 100 errors (dashed) or 400 errors (solid). Homogeneous DM setup with Power -1 dBm.

Fig. B.3 shows the Q-factor PMF obtained with the *correct* procedure for the DM homogeneous setup at power -1 dBm. The dashed curve, identical to Fig. B.2, shows the result obtained by stopping MC estimation of each BER after counting at least 100 errors. The solid curve pertain result by stopping MC estimation after counting at least 400 errors. We note that the PMF variance shrank of about 1.8 times, sign that Q-factor distribution from MC stopped after 100 errors was still corrupted by the MC uncertainty.

# Bibliography


[1] L. G. Kazovsky, G. Kalogerakis, and W.-T. Shaw, “Homodyne Phase-Shift-Keying Systems: Past Challenges and Future Opportunities,” *J. Lightw. Technol.*, vol. 24, no. 12, pp. 4876–4884, Dec. 2006.

[2] E. Ip and J. M. Khan, “Digital Equalization of Chromatic Dispersion and Polarization Mode Dispersion,” *J. Lightw. Technol.*, vol. 25, no. 8, pp. 2033–2043, Aug. 2007.

[3] O. Rival, G. Villares, and A. Morea, “Impact of Inter-Channel Nonlinearities on the planning of 25-100 Gb/s Elastic Optical Networks,” *J. Lightw. Technol.*, vol. 29, no. 9, pp. 1326–1334, May 2011.

[4] P. Serena, M. Bertolini, and A. Vannucci, “Optilux Toolbox,” [Online]. Available: http://www.optilux.sourceforge.net

[5] C. R. Menyuk, “Application of multiple-length-scale methods to the study of optical fiber transmission,” *J. Eng. Math.*, vol. 36, no. 1/2, pp. 113–136, Aug. 1999.

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

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

[8] M. Karlsson, "Probability Density Functions of the Differential Group Delay in Optical Fiber Communication Systems," *J. Lightw. Technol.*, vol. 19, no. 3, pp. 324–331, Mar. 2001.

[9] J. N. Damask, *Polarization Optics in Telecommunications*. New York: Springer, 2004.

[10] B. Huttner, C. Geiser, and N. Gisin, "Polarization-induced distortions in optical fiber networks with polarization-mode dispersion and polarization-dependent losses," *IEEE J. Sel. Topics Quantum Electron.*, vol. 6, no. 2, pp. 317–329, Mar./Apr. 2000.

[11] C. Menyuk and M. 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.

[12] D. Marcuse, C. R. Menyuk, and P. K. A. 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.

[13] 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.

[14] 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.

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

[16] 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.

[17] 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.

[18] 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.

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

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

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

[22] 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, 2008, Paper Th.3.E.3.

[23] 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.

[24] P. J. Winzer and A. H. Gnauk, "112-Gb/s polarization-multiplexed 16-QAM on a 25-GHz WDM grid," in *Proc. ECOC 2008*, Bruxelles, Belgium, 2008, Paper We.7.2.2.

[25] P. J. Winzer and R.-J. Essiambre, "Advanced modulation formats for high-capacity optical transport networks," *J. Lightw. Technol.*, vol. 24, no. 12, pp. 4711–4826, Dec. 2006.

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

[27] 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 2008*, San Diego, CA, 2008, Paper OThU8.

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

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

[30] 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.

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

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

[33] M. Taylor, "Algorithms for coherent detection," in *Proc. OFC 2010*, San Diego, CA, 2010, Paper OThL6.

[34] S. J. Savory, "Digital filters for coherent optical receivers," *Opt. Exp.*, vol. 16, no. 12, pp. 804–817, Jan. 2008.

[35] 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.

[36] E. Basch and T. Brown, "Introduction to coherent optical fiber transmission," *Communication Magazine, IEEE*, vol. 3, no. 5, pp. 23–30, May 1985.

[37] 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.

[38] 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.

[39] S. J. Savory, "Digital Signal Processing Options in Long Haul Transmissions," in *Proc. OFC 2008*, San Diego, CA, 2008, Paper OTuO3.

[40] S. J. 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, 2006, Paper Th2.5.5.

[41] 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.

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

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

[44] V. Curri, P. Poggiolini, A. Carena and F. Forghieri, “Dispersion Compensation and Mitigation of Nonlinear Effects in 111-Gb/s WDM Coherent PM-QPSK Systems,” *IEEE Photon. Technol. Lett.*, vol. 20, no. 17, pp. 1473–1465, Sep. 2008.

[45] P. Poggiolini, G. Bosco, A. Carena, V. Curri, V. Miot, and F. Forghieri “Performance Dependence on Channel Baud-Rate of PM-QPSK Systems Over Uncompensated Links,” *IEEE Photon. Technol. Lett.*, vol. 23, no. 1, pp. 15–17, Jan. 2011.

[46] F. Vacondio, O. Rival, C. Simonneau, E. Grellier, A. Bononi, L. Lorcy, J.-C. Antona, and S. Bigo, “On nonlinear distortions of highly dispersive optical coherent systems,” *Opt. Exp.*, vol. 20, no. 2, pp. 1022–1032, Jan. 2012.

[47] A. Carena, G. Bosco, V. Curri, P. Poggiolini, M. Tapia Taiba, and F. Forghieri, “Statistical characterization of PM-QPSK signals after propagation in uncompensated fiber links,” in Proc. *ECOC 2010*, Turin, Italy, 2010, Paper P4.07.

[48] A. Carena, V. Curri, G. Bosco, P. Poggiolini, and F. Forghieri, “Modeling of the Impact of Non-Linear Propagation Effects in Uncompensated Optical Coherent Transmission Links,” *J. Lightw. Technol.*, vol. 30, no. 10, pp. 1524–1539, May 2012.

[49] M. Nazarathy, J. Khurgin, R. Weidenfeld, Y. Meiman, P. S. Pak, R. Noe, I. Shpantzer, and V. Karagodsky, “Phased-array cancellation of nonlinear FWM in coherent OFDM dispersive multi-span links,” *Opt. Exp.*, vol. 16, no. 20, pp. 15777–15810, Sep. 2008.

[50] X. Chen and W. Shieh, “Closed-form expressions for nonlinear transmission performance of densely spaced coherent optical OFDM systems,” *Opt. Exp.*, vol. 18, no. 18, pp. 19039–19054, Aug. 2010.

[51] E. Grellier, and A. Bononi, “Quality parameter for coherent transmissions with Gaussian-distributed nonlinear noise,” *Opt. Exp.*, vol. 19, no. 13, pp. 12781–12788, Jun. 2011.

[52] P. Poggiolini, "The GN model of non-linear propagation in uncompensated coherent optical systems," *J. Lightw. Technol.*, vol. 30, no. 24, pp. 3857–3879, Dec. 2012.

[53] M. Secondini, E. Forestieri, and G. Prati, "Achievable Information Rate in Nonlinear WDM Fiber-Optic Systems With Arbitrary Modulation Formats and Dispersion Maps," *J. Lightw. Technol.*, vol. 31, no. 23, pp. 3839–3852, Dec. 2013.

[54] A. Bononi, P. Serena, N. Rossi, E. Grellier, and F. Vacondio, "Modeling nonlinearity in coherent transmissions with dominant intrachannel-four-wave-mixing," *Opt. Exp.*, vol. 20, no. 7, pp. 7777–7791, Mar. 2012.

[55] A. Bononi, N. Rossi, and P. Serena, "On the nonlinear threshold versus distance in long-haul highly-dispersive coherent systems," *Opt. Exp.*, vol. 20, no. 26, pp. B204–B216, Dec. 2012.

[56] J.-C. Antona and S. Bigo, "Physical design and performance estimation of heterogeneous optical transmission systems," *Comptes Rendus - Physique*, vol. 9, pp. 963–984, Nov.-Dec. 2008.

[57] S. Kilmurray, T. Fehenberger, P. Bayvel, and R. I. Killey, "Comparison of the nonlinear transmission performance of quasi-Nyquist WDM and reduced guard interval OFDM," *Opt. Exp.*, vol. 20, no. 4, pp. 4198–4205, Feb. 2012.

[58] L. B. Du and A. J. Lowery, "Optimizing the subcarrier granularity of coherent optical communications systems," *Opt. Exp.*, vol. 19, no. 9, pp. 8079–84, Apr. 2011.

[59] A. D. Ellis et al., "Future Directions to Realise Ultra- High Bit Rate Transmission Systems," in *Proc. OECC 2010*, Sapporo, Japan, 2010, Paper 8B3-1.

[60] S. L. Jansen, "Multi-Carrier Approaches for Next-Generation Transmission: Why, Where and How?," in *Proc. OFC 2012*, Los Angeles, CA, 2012, Paper OTh1B.1.

[61] Y. Tang, W. Shieh, and B. S. Krongold, "DFT-Spread OFDM for Fiber Nonlinearity Mitigation," *IEEE Photon. Technol. Lett.*, vol. 22, no. 16, pp. 1250–1252, Aug. 2010.

[62] E. Palkopoulou, G. Bosco, A. Carena, D. Klonidis, P. Poggiolini, and I. Tomkos, "Nyquist-WDM-Based Flexible Optical Networks: Exploring Physical Layer Design Parameters," *J. Lightw. Technol.*, vol. 31, no. 14, pp. 2332–2339, Jul. 2013.

[63] C. Behrens, R. I. Killey, S. J. Savory, M. Chen and P. Bayvel, "Nonlinear Transmission Performance of Higher-Order Modulation Formats," *IEEE Photon. Technol. Lett.*, vol. 23, no. 6, pp. 377–379, Mar. 2011.

[64] 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.

[65] J. Renaudier, O. Bertran-Pardo, H. Mardoyan, P. Tran, M. Salsi, G. Charlet, and S. Bigo, "Impact of temporal interleaving of polarization tributaries onto 100-Gb/s coherent transmission systems with RZ pulse carving," *IEEE Photon. Technol. Lett.*, vol. 20, no. 24, pp. 2036–2038, Dec. 2008.

[66] O. Bertran-Pardo, J. Renaudier, G. Charlet, M. 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. 19902–19907, Oct. 2009.

[67] Y. Frignac, J. C. Antona, and S. Bigo, "Enhanced analytical engineering rule for fast optimization of dispersion map in 40 Gbit/s-based transmission systems," in *Proc. OFC 2004*, Los Angeles, CA, 2004, Paper TuN3.

[68] 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.

[69] T. Pfau and S. Hoffmann, "Hardware-Efficient Coherent Digital Receiver Concept With Feedforward Carrier Recovery for -QAM Constellations," *J. Lightw. Technol.*, vol. 27, no. 8, pp. 989–999, Apr. 2009.

[70] 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.

[71] E. Ip and J. M. Kahn, "Feedforward Carrier Recovery for Coherent Optical Communications," *J. Lightw. Technol.*, vol. 25, no. 9, pp. 2675–2692, Sep. 2007.

[72] A. Bononi, N. Rossi, and P. Serena, "Transmission Limitations due to Fiber Nonlinearity," in *Proc. OFC 2011*, Los Angeles, CA, 2011, Paper OWO7.

[73] 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 2011*, Los Angeles, CA, 2011, paper OWO1.

[74] 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, 2009, paper Tu3.4.5.

[75] D. Sperti, "Polarization-dependent nonlinear effects in coherent detection systems", Ph.D. dissertation, Dept. Information Engineering. Università di Parma, Parma, Italy, Jan. 2012.

[76] Y.-H. Wang and I. Lyubomirsky, "Impact of DP-QPSK pulse shape in nonlinear 100 G Transmission," *J. Lightw. Technol.*, vol. 28, no. 18, pp. 2750–2756 , Sep. 2010.

[77] A. Bononi, O. Beucher, and P. Serena, "Single- and cross-channel nonlinear interference in the Gaussian Noise model with rectangular spectra," *Opt. Exp.*, vol. 21, no. 26, pp. 32254–32268, Dec. 2013.

[78] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "The GN model of fiber non-linear propagation and its applications," *J. Lightw. Technol.*, vol. 32, no. 4, pp. 694–721, Feb. 2014.

[79] P. Serena, and A. Bononi, "An Alternative Approach to the Gaussian Noise Model and its System Implications," *J. Lightw. Technol.*, vol. 31, no. 22, pp. 3489–3499, Nov. 2013.

[80] A. Bononi and P. Serena, "An alternative derivation of Johannisson's regular perturbation model," arXiv:1207.4729v1 [physics.optics], 2012.

[81] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Properties of nonlinear noise in long, dispersion-uncompensated fiber links," *Opt. Exp.*, vol. 21, no. 22, pp. 25685–25699, Jul. 2013.

[82] A. Carena, G. Bosco, V. Curri, P. Poggiolini, and F. Forghieri, "Impact of the Transmitted Signal Initial Dispersion Transient on the Accuracy of the GN-Model of Non-Linear Propagation," in *Proc. ECOC 2013*, London, United Kingdom, 2013, Paper Th.1.D.4.

[83] P. Serena, A. Bononi, "On the Accuracy of the Gaussian Nonlinear Model for Dispersion-unmanaged Coherent Links," in *Proc. ECOC 2013*, London, United Kingdom, 2013, Paper Th.1.D.3.

[84] A. Carena, G. Bosco, V. Curri, P. Poggiolini, and F. Forghieri, "EGN model of non-linear fiber propagation," *Opt. Exp.*, vol. 22, no. 13, pp. 16335–16362, Jul. 2014.

[85] P. Serena, N. Rossi, A. Bononi, "The Impact of the Modulation Dependent Nonlinear Interference Missed by the Gaussian Noise Model," in *Proc. ECOC 2014*, Cannes, France, 2014, Paper Mo.4.3.1.

[86] A. Mecozzi, R. Essiambre, "Nonlinear Shannon Limit in Pseudolinear Coherent Systems," *J. Lightw. Technol.*, vol. 30, no. 12, pp. 2011–2024, Apr. 2012.

[87] C. Xie, “Impact of nonlinear and polarization effects in coherent systems,” *Opt. Exp.*, vol. 19, no. 26, pp. B915–B930, Dec. 2011.

[88] P. Serena, N. Rossi, O. Bertran-Pardo, J. Renaudier, A. Vannucci, and A. Bononi, “Intra- Versus Inter-Channel PMD in Linearly Compensated Coherent PDM-PSK Nonlinear Transmissions,” *J. Lightw. Technol.*, vol. 29, no. 11, pp. 1691–1700, Jun. 2011.

[89] O. Bertran-Pardo, J. Renaudier, G. Charlet, P. Tran, H. Mardoyan, M. Salsi, and S. Bigo, “Experimental assessment of interactions between nonlinear impairments and polarization-mode dispersion in 100-Gb/s coherent systems versus receiver complexity,” *IEEE Photon. Technol. Lett.*, vol. 21, no. 1, pp. 51–53, Jan. 2009.

[90] M. Shtaif, “Performance degradation in coherent polarization multiplexed systems as a result of polarization dependent loss,” *Opt. Exp.*, vol. 16, no. 18, pp. 13918–13932, Sep. 2008.

[91] C. Xie, “Polarization-Dependent Loss Induced Penalties in PDM-QPSK Coherent Optical Communication Systems”, in *Proc. OFC 2010*, San Diego, CA, 2010, Paper OWE6.

[92] O. Vassilieva, I. Kim, T. Naito, “Systematic Investigation of Interplay between Nonlinear and Polarization Dependent Loss Effects in Coherent Polarization Multiplexed Systems,” in *Proc. ECOC 2010*, Torino, Italy, 2010, Paper P4.08.

[93] Z. Tao, L. Li, T. Hoshida and J. Rasmussen, “Interaction between PDL and Intra-channel Nonlinearity in Dual Polarization Systems,” in *Proc. OECC 2011*, Kaohsiung, Taiwan, 2011, Paper 6B3_2.

[94] O. Vassilieva, I. Kim, M. Sekiya, “Statistical Analysis of the Interplay between Nonlinear and PDL Effects in Coherent Polarization Multiplexed Systems,” in *Proc. ECOC 2012*, Amsterdam, the Netherlands, 2012, Paper We.3.C.4.

[95] O. Vassilieva, S. Oda, T. Hoshida, J. Rasmussen, M. Sekiya, “Experimental Investigation of the Statistics of the Interplay between Nonlinear and PDL

Effects in Coherent Polarization Multiplexed Systems," in *Proc. OFC 2013*, Anaheim, CA, 2013, Paper OM3B.6.

[96] M. Karlsson, J. Brentel, and P. A. Andrekson, "Long-Term Measurement of PMD and Polarization Drift in Installed Fibers," *J. Lightw. Technol.*, vol. 18, no. 7, pp. 941–951, Jul. 2000.

[97] Z. Tao, L. Dou, T. Hoshida, and J. Rasmussen, "A Fast Method to Simulate the PDL Impact on Dual-Polarization Coherent Systems," *IEEE Photon. Technol. Lett.*, vol. 21, no. 24, pp. 1882–884, Dec. 2009.

[98] G. S. Fishman, *Monte Carlo: concepts, algorithms and applications*, New York: Springer-Verlag, 1999.

[99] *ITU-T Rec. G.680: Physical Transfer Functions of Optical Networks Elements*, Int. Telecommun. Union, Geneva, Switzerland, July 2007.

[100] P. J. Winzer, M. Pfenningbauer, and R.-J. Essiambre, "Coherent Crosstalk in Ultradense WDM Systems," *J. Lightw. Technol.*, vol. 23, no. 4, pp. 1734–1744, Apr. 2005.

[101] P. Serena, N. Rossi, M. Bertolini, and A. Bononi, "Stratified Sampling Monte Carlo Algorithm for Efficient BER estimation in Long-Haul Optical Transmission Systems," *J. Lightw. Technol.*, vol. 27, no. 13, pp. 2404–2411, Jul. 2009.

[102] J. O. Irwin, "On the frequency distribution of the mean of samples from a population having any law of frequency with finite moments, with Special Reference to Pearson's Type II," *Biometrika*, vol. 19, no. 3/4, pp. 225–239, Dec. 1927.

[103] J.-C. Antona, E. Grellier, A. Bononi, S. Petitrenaud, and S. Bigo, "Revisiting binary sequence length requirements for the accurate emulation of highly dispersive transmission systems," in *Proc. ECOC 2008*, Brussels, Belgium, 2008, paper We.1.E.3.

[104] E. Grellier, J.-C. Antona, A. Bononi, and S. Bigo, “Revisiting Binary Sequence Length Requirements to Accurately Emulate Optical Transmission Systems in Highly Dispersive Regime,” in *Proc. Optical Transmission, Switching, and Subsystems VI*, Hangzhou, China, Oct. 2008, SPIE vol. 7136, 713613, 2008.

[105] R. Y. Rubinstein, *Simulation and the Monte Carlo Method*, New York: Wiley, 1981

[106] J. Chen, D. Lu, J. Sadovsky, and K. Yao, “On importance sampling in digital communications. I. Fundamentals,” *IEEE J. Sel. Areas Commun.*, vol. 11, no. 3, pp. 289–299, Apr. 1993.

[107] G. Biondini, W. L. Kath, and C. R. Menyuk, “Importance Sampling for Polarization-Mode Dispersion: Techniques and Applications,” *J. Lightw. Technol.*, vol. 22, no. 4, pp. 1201–1215, Apr. 2004.

[108] S. Huard, *Polarisation de la Lumière*. Paris, France: Masson, 1994.

[109] Q. Zhang and M.I. Hayee, “Symmetrized Split-Step Fourier Scheme to Control Global Simulation Accuracy in Fiber-Optic Communication Systems,” *J. Lightw. Technol.*, vol. 26, no. 2, pp. 302–316, Jan. 2008.

# Acknowledgments

This thesis comes at the end of a path that lasted about 8 years. During this period I met many people that made it a special experience.

First of all, I would like to thank my family and my girlfriend Claudia for their unconditional and continuous support.

I want to thank my supervisor Prof. Alberto Bononi for giving me the opportunity to work in his group and for investing on me. I owe him much of my professional growth. His attitude has been an example to follow. In Parma I worked with wonderful people as Dr. Alessandra Orlandini, Dr. Marco Bertolini, and Dr. Donato Sperti. Honorable mention goes to Prof. Paolo Serena. I'm sure I will tremendously miss our daily chats, always inspiring and enjoyables.

I want to thank Prof. Leslie Rusch for welcoming me as a visiting student in her research group at Université Laval, where I acquainted wonderful people as Dr. Francesco Vacondio, Dr. Amirhossein Ghazisaeidi and Dr. Walid Mathlouthi. It has been an experience that really broadened my horizons.

I want to thank Dr. Sébastien Bigo for giving me the opportunity to work in his team, and all people at Alcatel-Lucent Bell-Labs France for putting me at my ease during those 6 months. In particular, I want to thank Dr. Petros Ramantanis and Dr. J.-C. Antona for supervising my work and for their wise suggestions.

Finally, I want to thank all friends acquainted at the Department of Information Engineering. They are so many that I can't mention them all. They have decisively contributed to make these last 8 years a wonderful period and to make it happily last so long.