



UNIVERSITÀ DI PARMA

UNIVERSITÀ DEGLI STUDI DI PARMA

Dottorato di Ricerca in Tecnologie dell'Informazione

XXXIV Ciclo

**Physical Layer Modeling for High Capacity Optical
Communication Systems**

Coordinatore:

Chiar.mo Prof. Marco Locatelli

Supervisor:

Prof. Paolo Serena

Dottoranda: *Chiara Lasagni*

Anni 2018/2019–2021/2022

To my sister Giulia

Abstract

Nowadays, ultra-wideband (UWB) and space-division multiplexing (SDM) transmissions are emerging as the best candidates to further push the capacity offered by optical communication systems. The former strategy aims at increasing the capacity by enlarging the transmission bandwidth. For these systems, a fully numerical approach for the performance estimation has a prohibitive computational time, due to the wide bandwidth. On the other hand, an SDM system exploits spatial diversity to increase the amount of transmitted data. However, the random nature of the interaction among spatial paths requires several numerical simulations to collect enough results to build the system performance statistics.

In such complex scenarios, analytical modeling stands out as a fast yet accurate tool for system performance estimation, well-suited to assist in the network design and traffic routing. In this thesis, we propose extensions of analytical models for the estimation of the system performance for UWB and SDM systems, with a particular focus on the modeling of the optical fiber nonlinearities.

In the context of UWB transmissions, we first focused on modeling the interaction between the optical fiber Kerr effect and stimulated Raman scattering (SRS). We proposed a model for the nonlinear interference (NLI) variance which takes into account the positioning of the equalizers for the SRS compensation. The model is validated against UWB numerical simulations, showing a very good agreement along with the importance of addressing the equalizer positioning along the link. Then, we moved the focus on the modeling of semiconductor optical amplifiers (SOAs). As a trade-off between simplicity and complexity, we proposed a parametric model for describing the SOA dynamics.

Regarding SDM transmissions, we first addressed the inclusion of polarization-dependent loss (PDL) in the NLI modeling. The proposed model accounting for PDL in single-mode transmissions opens the door to the more general case of mode-dependent loss in SDM systems. Then, we derived a theoretical model of the NLI in the presence of arbitrary mode dispersion among strongly coupled modes. Both models are validated against numerical simulations, repeated for several random realizations, showing excellent agreement.

Contents

List of Acronyms	vii
List of Contributions	xiii
Introduction	1
1 Review of background theory	5
1.1 Digital modulation schemes	5
1.2 Optical modulation	7
1.3 Multiplexing	9
1.3.1 Wavelength-division multiplexing	9
1.3.2 Polarization-division multiplexing	10
1.3.3 Space-division multiplexing	11
1.4 The fiber-optic channel	13
1.4.1 Scalar propagation	13
1.4.2 Dual-polarization propagation	21
1.4.3 SDM propagation	24
1.5 Approximate solutions of the Manakov equation	25
1.5.1 First-order regular perturbation	25
1.5.2 Split-step Fourier method	26
1.6 Multi-span optical-links	29
1.7 Quality of transmission metrics	31
1.7.1 Signal-to-noise ratio	32

1.7.2	Achievable information rate	34
2	Theoretical models for the nonlinear interference variance	37
2.1	Nonlinear interference	37
2.1.1	Link kernel	40
2.1.2	Nonlinear interference variance	41
2.2	The Gaussian noise model	44
2.2.1	Physical insight	47
2.2.2	Closed-form expressions	50
2.3	The Enhanced Gaussian noise model	52
2.4	Monte Carlo Integration	60
I	Ultra-wideband transmissions	67
3	EGN model in the presence of Stimulated Raman Scattering	69
3.1	Stimulated Raman Scattering	71
3.2	SNR generalization	76
3.3	The SRS-EGN model	78
3.3.1	Numerical validation	80
3.3.2	Impact of the modulation format	82
3.4	The DGE-SRS-EGN model	87
3.4.1	Integral model	88
3.4.2	Simplified model	90
3.4.3	GN-term closed-form expressions	92
3.4.4	Numerical validation	95
3.4.5	Modulation format	102
3.4.6	Self- and cross-channel effects	106
3.5	Optimization of signal power pre-emphasis	108
4	Modeling of semiconductor optical amplifiers	119
4.1	Connelly model overview	120

4.1.1	Numerical solution	122
4.2	SOA model simplifications	126
4.3	The integrated-gain model	132
4.3.1	Numerical implementation	136
4.3.2	Numerical results	138
4.4	SOA-amplified optical line	142
II	Spatial multiplexing	151
5	EGN model with polarization dependent loss	153
5.1	Polarization-dependent loss	154
5.2	The PDL-EGN model	155
5.2.1	ASE variance	160
5.2.2	NLI variance	161
5.2.3	Per-polarization PDL-aware SNR	165
5.3	Numerical results	166
5.3.1	Outage probability	169
5.3.2	NLI vs ASE	171
5.3.3	Modulation format	175
5.3.4	Number of channels	176
5.3.5	PDL profile impact	177
6	The ergodic GN model with modal dispersion	179
6.1	Modal dispersion	180
6.1.1	The PMD vector	180
6.1.2	The MD vector	181
6.1.3	The linear fiber response with MD	182
6.2	The ergodic MD-GN model	184
6.2.1	Ergodic NLI variance	186
6.2.2	FWM efficiency	190
6.3	Inter-channel MD simplification	192

6.4	Numerical results	197
6.4.1	The inter-channel MD GN model	197
6.4.2	Full MD	201
	Conclusions	211
A	Connelly model: extra equations and parameters	215
	Bibliography	219
	Acknowledgements	235

List of Acronyms

AIR achievable information rate

ASE amplified spontaneous emission

AWGN additive white Gaussian noise

BER bit error rate

CB conduction band

CLE constant local error

CNLSE coupled nonlinear Schrödinger equation

CPE carrier phase estimator

CPU central process unit

CUT channel under test

CW constant-wave

DCF dispersion compensating fiber

DGD differential group delay

DGE dynamic gain equalizer

DM dispersion managed

DSF	dispersion shifted fiber
DSP	digital signal processing
DU	dispersion uncompensated
EDFA	Erbium-doped fiber amplifier
EGN	enhanced Gaussian noise
FEC	forward error correction
FFT	fast Fourier transform
FON	fourth order noise
FWM	four-wave mixing
GN	Gaussian noise
GPU	graphical process unit
GVD	group velocity dispersion
HON	higher order noise
I	in-phase
IG	integrated-gain
i.i.d	independent and identically distributed
IIR	intensity impulse response
IS	importance sampling
MC	Monte Carlo
MCF	multi-core fiber
MCI	multi-channel interference

MD	modal dispersion
MIMO	multiple-input multiple-output
MMF	multi-mode fiber
MS	multi-stage
NLI	nonlinear interference
NLP	nonlinear phase
NLSE	nonlinear Schrödinger equation
NLT	nonlinear threshold
NZDSF	non-zero dispersion-shifted fiber
ODE	ordinary differential equation
OOK	on-off keying
PDF	probability density function
PDL	polarization-dependent loss
PDM	polarization-division multiplexing
PMD	polarization-mode dispersion
PMF	probability mass function
PSD	power spectral density
PSP	principal state of polarization
Q	quadrature
QAM	quadrature amplitude modulation
QoT	quality of transmission

QPSK	quadrature phase shift keying
ROADM	reconfigurable add-drop multiplexer
RP1	regular perturbation
RRC	root-raised cosine
SDE	stochastic differential equation
SDM	space-division multiplexing
SMD	spatial mode dispersion
SMF	single-mode fiber
SNR	signal to noise ratio
SOA	semiconductor optical amplifier
SON	second order noise
SOP	state of polarization
SPM	self-phase modulation
SRS	stimulated Raman scattering
SS	stratified sampling
SSFM	split-step Fourier method
UWB	ultra-wideband
VB	valence band
WDM	wavelength-division multiplexing
WOI	wavelength of interest
WSS	wavelength-selective switch

XCI cross-channel interference

XPM cross-phase modulation

XpolM cross-polarization modulation

List of Contributions

1. **C. Lasagni**, P. Serena, and A. Bononi, "A Raman-aware enhanced GN-model to estimate the modulation format dependence of the SNR tilt in C+L band," in *European Conference on Optical Communication (ECOC)*, Dublin, Ireland, Sep. 2019, paper W1D.2.
2. P. Serena, **C. Lasagni**, S. Musetti, and A. Bononi, "On Numerical Simulations of Ultra-Wideband Long-Haul Optical Communication Systems," *J. Lightw. Technol.*, vol. 38, no. 5, pp. 1019–1031, Mar. 2020.
3. P. Serena, **C. Lasagni** and A. Bononi, "The Enhanced Gaussian Noise Model Extended to Polarization-Dependent Loss," in *J. Lightw. Technol.*, vol. 38, no. 20, pp. 5685-5694, Oct. 2020.
4. **C. Lasagni**, P. Serena, and A. Bononi, "Impact of Periodic Gain Equalization in Presence of Stimulated Raman Scattering," Italian Conference on Optics and Photonics (ICOP) 2020. Oral presentation.
5. **C. Lasagni**, P. Serena, and A. Bononi, "Estimation of Outage Probability in the Presence of Polarization Dependent Loss," Italian Conference on Optics and Photonics (ICOP) 2020. Oral presentation.
6. **C. Lasagni**, P. Serena, and A. Bononi, "Impact of Sparse Gain Equalization in the Presence of Stimulated Raman Scattering," in *European Conference on Optical Communication (ECOC)*, Brussels, Belgium, Dec. 2020, paper We2F.4.

7. **C. Lasagni**, P. Serena, and A. Bononi, "Modeling Nonlinear Interference With Sparse Raman-tilt Equalization," *J. Lightw. Technol.*, vol. 39, no. 15, pp. 4980-4989, Aug. 2021.
8. A. Bononi, J. -C. Antona, P. Serena, A. Carbo Meseguer and **C. Lasagni**, "The Generalized Droop Model for Submarine Fiber-optic Systems," *J. Lightw. Technol.*, vol. 39, no. 16, pp. 5248-5257, Aug. 2021.
9. P. Serena, **C. Lasagni**, A. Bononi, C. Antonelli and A. Mecozzi, "A Model of the Nonlinear Interference in Space-Division Multiplexed Systems with Arbitrary Modal Dispersion," in *European Conference on Optical Communication (ECOC)*, Bordeaux, France, Sep. 2021, paper Tu2D.3.
10. **C. Lasagni**, P. Serena, A. Bononi, and J.-C. Antona, "Power Allocation Optimization in the Presence of Stimulated Raman Scattering," in *European Conference on Optical Communication (ECOC)*, Bordeaux, France, Sep. 2021, paper Tu1D.3.
11. J. Girard-Jollet, M. Lonardi, P. Ramantanis, P. Serena, **C. Lasagni**, P. Layec, and J.-C. Antona, "Estimating Network Components Polarization-Dependent Loss Using Performance Statistical Measurements," in *European Conference on Optical Communication (ECOC)*, Bordeaux, France, Sep. 2021, paper Tu1C 1.5.
12. P. Serena, **C. Lasagni**, A. Bononi, C. Antonelli, and A. Mecozzi, "The Ergodic GN Model for Space-Division Multiplexing with Strong Mode Coupling," *submitted to J. Lightw. Technol.*
13. J. Girard-Jollet, M. Lonardi, P. Ramantanis, P. Serena, **C. Lasagni**, P. Layec, and J.-C. Antona, "Estimating Network Components PDL Using Performance Statistical Measurements," *submitted to J. Lightw. Technol.*

Introduction

Optical communication systems sustain most of the current worldwide data traffic, spanning from terrestrial networks to transoceanic transmissions.

In the past decades, optical networks kept the pace with the growing traffic demand thanks to the advent of disruptive technologies, (e.g., all-optical amplifiers and cross-connects, coherent detection, and digital signal processing) and to the extensive leverage of multiplexing in the degrees of freedom offered by the standard single-mode optical fiber. Multiplexing in frequency (wavelength-division multiplexing) and polarization (polarization-division multiplexing), combined with the use of advanced modulation formats, pushed the capacity of commercial optical systems up to almost 100 Tb/s [1] over submarine distances.

Nevertheless, the disparity between the traffic growth and the system capacity growth rate ($\sim 60\%$ vs $\sim 20\%$ [2]) motivated the search for novel techniques and technologies. Nowadays, two main approaches stand out as the best candidates to meet the capacity demand: ultra-wideband (UWB) transmissions and space-division multiplexed (SDM) systems.

UWB transmissions aim at empowering the frequency multiplexing by increasing the transmission bandwidth. This solution offers the possibility to increase the capacity exploiting the wide bandwidth over which modern single-mode fibers exhibit low attenuation values, beyond the usually adopted C band (~ 5 THz). In particular, a widespread solution adopted by vendors is the deployment of multi-band C+L band systems (~ 13 THz bandwidth). Despite the potential scaling of the system capacity with the transmission bandwidth, the design of such multi-band systems needs to

face several challenges. From the technological point of view, one of the main issues is represented by the amplification of such wide bandwidths signals. For instance, semiconductor optical amplifiers are gaining interest due to their wide amplification bandwidth. Moreover, the optical fiber exhibits a different nonlinear response over wide bandwidths which leads to stimulated Raman scattering. Such a nonlinear process causes the amplification of low-frequency components while depleting high frequencies. This phenomenon becomes more severe as the transmission bandwidth increases, thus requiring to be addressed properly for a reliable estimation of the system performance.

On the other hand, SDM systems offer the possibility to unleash the spatial dimension. One possibility to achieve spatial diversity is the deployment of parallel single-mode fibers. While this solution allows relying on the same technologies adopted in common optical communication systems, extensive research has been done on SDM systems based on optical fibers that intrinsically accommodate different spatial paths. To be cost-effective, the latter systems need to fully implement SDM. In particular, one of the main challenges is related to optical amplifiers able to uniformly amplify all the spatial paths carried by SDM fibers. Even in this scenario, the system performance estimation becomes non-trivial due to the linear and nonlinear interaction among the spatial paths during propagation along with the optical fiber. For instance, due to random perturbation in the optical fibers, the different modes mix during propagation and experience different delays, i.e., mode dispersion.

Both types of systems need a quick and reliable way to estimate the overall performance, to assist the design stage. While a fully numerical approach is too time-consuming for such complex systems, analytical models represent the ultimate tool for a fast system performance estimation along with a physical understanding of the underlying impairments. To achieve an accurate estimate, the analytical models from the single-mode and single-band literature must be extended to account for the physical impairments arising in SDM and UWB transmissions.

This thesis will propose extensions to analytical models for the performance estimation of UWB and SDM scenarios, with a particular focus on fiber nonlinearities. The remainder of the thesis is structured as follows.

Chapter 1 reviews the basic concepts of optical communications and introduces the notation that will be adopted in the remainder of the thesis.

Chapter 2 introduces the Gaussian noise model [3] and the enhanced Gaussian noise model [4] for the estimation of the fiber nonlinearities, with a detailed description of the supporting proofs. Strategies for the efficient implementation of such models are discussed.

The remainder of the thesis is divided into two parts. Part I is devoted to ultra-wideband transmissions, and is composed of Chapters 3 and 4, while Part II focuses on SDM systems, and contains Chapters 5 and 6.

Chapter 3 proposes an extension of the enhanced Gaussian noise model to include stimulated Raman scattering, accounting for an arbitrary positioning of equalizers in the link for the compensation of stimulated Raman scattering on the signal power. Novel extensions of closed-form expressions available in the literature [5] to sparse equalizers positioning are presented.

In Chapter 4 we discuss the modeling of semiconductor optical amplifiers. This chapter aims to propose a simple model for the amplifier to be used in the estimation of the overall system performance. As a first step, in this chapter, we exploit such a model for fast and accurate numerical simulations. This work was carried out during a six months internship at Nokia Bell Labs, France.

Chapter 5 proposes an extension of the enhanced Gaussian noise model to include random polarization-dependent losses. The model serves as a first step towards the inclusion of mode-dependent losses in an SDM system.

In Chapter 6 we propose the ergodic Gaussian noise model, i.e., an original extension of the Gaussian noise model to account for mode dispersion in SDM transmissions.

Finally, we will draw our conclusions.

Chapter 1

Review of background theory

In this chapter, we review basic concepts of background theory on optical communications. The structure of the chapter follows the structure of a generic optical communication system, as sketched in Fig. 1.1, with a particular focus on the fiber-optic channel.

1.1 Digital modulation schemes

We assume the transmitted signal to be a linearly modulated digital signal, which can be expressed as:

$$s(t) = \sum_{k=-\infty}^{\infty} a_k p(t - kT) \quad (1.1)$$

where a_k is the digital symbol at time k , while $p(t - kT)$ is the supporting pulse at the k th symbol time of duration T . The pulse is normalized such that the average transmitted power is given by $P = \mathbb{E} \{ |a_k|^2 \}$, with \mathbb{E} indicating expectation.



Figure 1.1: Block diagram of a generic optical communication system.

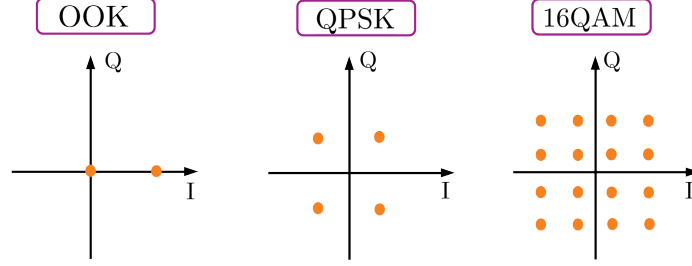


Figure 1.2: Constellation diagrams. From left to right: on-off keying (OOK), quadrature phase shift keying (QPSK), and 16 quadrature amplitude modulation (QAM).

We suppose each transmitted symbol to take values within an alphabet of cardinality M , according to the chosen modulation format, and with a uniform probability distribution. The most popular modulation formats are amplitude-only, phase-only, or phase and amplitude modulation. The latter modulation formats are of particular interest in modern communication systems since they carry information symbols in both in-phase (I) and quadrature (Q) components. A notable example is quadrature amplitude modulation (QAM) formats. In the low-pass description, the symbols are thus, in general, complex numbers, that can be reported in the complex plane (IQ plane), describing a constellation. Figure 1.2 shows an example of three polar constellations. First, we observe on-off keying (OOK) modulation, which modulates only the in-phase component, in a binary alphabet. A second example is quadrature phase-shift keying (QPSK), composed of four symbols shifted in the phase of $\pi/2$. The last example reported in Fig. 1.2 is a QAM modulation with cardinality $M = 16$.

Root-raised cosine (RRC) pulses are typically used as supporting pulses, being immune to inter-symbol interference once detected by a matched filter [6] and because of their band-limited property. An example of such pulses is sketched in Fig. 1.3 for different roll-off factors, which govern the pulse shape, as a function of time expressed as multiples of the symbol time. A small roll-off factor near 0 is typically adopted in modern communication systems since it yields a small bandwidth occupation in the frequency domain. For the sake of completeness, Fig. 1.3 reports in the inset the modulus of the RRC in the frequency-domain $|\tilde{P}(f)|$. We recall that in this

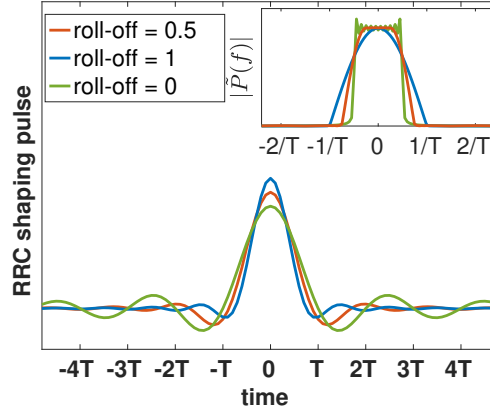


Figure 1.3: Root-raised cosine (RRC) pulse with roll-off factor 0, 0.5 and 1. Inset: frequency-domain representation. T : symbol time.

thesis, a generic function $p(t)$ is related to its Fourier transform $\tilde{P}(\omega)$ by:

$$\begin{aligned}\tilde{P}(\omega) &= \mathcal{F}\{p(t)\} = \int_{-\infty}^{+\infty} p(t) e^{-j\omega t} dt \\ p(t) &= \mathcal{F}^{-1}\{\tilde{P}(\omega)\} = \int_{-\infty}^{+\infty} \tilde{P}(\omega) e^{j\omega t} \frac{d\omega}{2\pi}\end{aligned}\tag{1.2}$$

where ω is the angular frequency, related to the frequency f by $\omega = 2\pi f$. As usual in the literature, in this thesis, we will call both ω and f simply frequency since the meaning is clear from the notation.

1.2 Optical modulation

The digital signal carrying the desired information is then used to modulate an optical carrier generated by a laser source. In fiber-optic communication systems, the modulated light is then confined and guided by an optical fiber. The optical fiber is a waveguide composed of two concentric cylinders made of silica. The guiding is achieved through internal reflection, thanks to the refractive index difference between

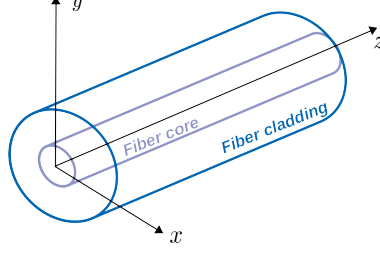


Figure 1.4: Sketch of the structure of a generic optical fiber.

the inner and the outer structures, respectively the fiber core and cladding. Fig. 1.4 sketches the structure of a generic optical fiber.

The dimension of the core and the refractive index difference discriminate the number of *modes* that can be guided by the optical fiber. A single-mode fiber (SMF) is an optical fiber that guides only the fundamental mode of light [7].

Assuming that the light is polarized along the x axis, the electric field for the fundamental mode at time t and generic coordinate z along the optical fiber length can be written as [7]

$$\vec{\mathcal{E}}(x, y, z, t) = \hat{x} \left\{ \text{Re} \left[F(x, y) E(z, t) e^{j(\beta_0 z - \Omega_c t)} \right] \right\} \quad (1.3)$$

where Ω_c is the central optical frequency, \hat{x} is the polarization unit vector, and β_0 is the propagation constant. $E(z, t)$ is the slowly-varying pulse envelope, also known as low-pass equivalent, whose instantaneous power is $|E(z, t)|^2$. For communication purposes, at coordinate $z = 0$ it is $E(0, t) = s(t)$, with $s(t)$ the linearly modulated signal in Eq. (1.1).

The term $F(x, y)$ in Eq. (1.3) indicates the transverse field distribution, which fades exponentially outside the fiber core [7], and is independent of the longitudinal coordinate z . For the fundamental model, such transverse distribution is typically approximated as Gaussian [7].

1.3 Multiplexing

In this section, we briefly review multiplexing techniques and introduce the notation that will be used for the remainder of this thesis.

1.3.1 Wavelength-division multiplexing

The fiber-optic channel offers a much wider bandwidth compared to electronics. Therefore, it is natural to exploit it by multiplexing techniques. A popular technique is wavelength-division multiplexing (WDM). WDM is a technique that multiplexes together independent signals modulated at different carrier frequencies, thus increasing the system throughput without stressing the modulator bandwidth. A WDM signal resulting from the multiplexing of N_{ch} channels can be written as

$$E(0, t) = \sum_{h=1}^{N_{\text{ch}}} \sum_{k=-\infty}^{\infty} a_{kh} p_h(t - kT_h) e^{j\omega_h t} \quad (1.4)$$

where a_{kh} is the digital symbol carried by channel h at time k , while ω_h is the low-pass channel central frequency in the band-pass reference system centered at Ω_c . In the most general case, an arbitrary channel-dependent phase-shift can be included in Eq. (1.4). The frequency-domain representation of this WDM signal is given by:

$$\begin{aligned} \tilde{E}(0, \omega) &= \sum_{h=1}^{N_{\text{ch}}} \sum_{k=-\infty}^{\infty} a_{kh} \tilde{p}_h(\omega - \omega_h) e^{-j(\omega - \omega_h)kT_h} \\ &\triangleq \sum_{h=1}^{N_{\text{ch}}} \sum_{k=-\infty}^{\infty} a_{kh} \tilde{G}_{kh}(0, \omega - \omega_h) \end{aligned} \quad (1.5)$$

where $\tilde{G}_{kh}(0, \omega - \omega_h)$ is the Fourier transform of the following basis function:

$$G_{kh}(0, t) \triangleq p(t - kT_h) e^{j\omega_h t}. \quad (1.6)$$

Figure 1.5(a) sketches the WDM mechanism for a simple case of 3 channels having carrier frequencies Ω_1, Ω_2 , and Ω_3 . The three independent signals are multi-

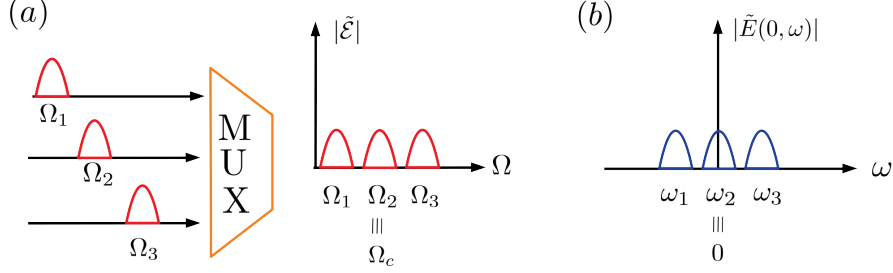


Figure 1.5: Sketch of the WDM technique. (a): three independent signals having different carriers are multiplexed together into a WDM signal. (b): frequency domain low-pass equivalent of the field.

plexed together to compose a WDM field. The low-pass equivalent of such a WDM field, which is defined in Eq. (1.5), is depicted in Fig. 1.5(b).

It can be seen from Eq. (1.4) that the WDM field can be written compactly as a superposition of N_{ch} channels:

$$E(0, t) = \sum_{h=1}^{N_{\text{ch}}} E_h(0, t) e^{j\omega_h t} \quad (1.7)$$

where each E_h writes as in Eq. (1.1).

1.3.2 Polarization-division multiplexing

Polarization-division multiplexing (PDM) is a technique that exploits the two orthogonal polarizations of light to transmit information. Each polarization tributary is modulated as in Eq. (1.1). The complex envelope can thus be written as a 2D complex vector, collecting both x and y polarizations, hence a Jones vector [8]. We adopt the bra-ket notation for vectors, namely:

$$|E\rangle \triangleq \begin{bmatrix} E_x \\ E_y \end{bmatrix}, \quad \langle E| \triangleq [E_x^*, E_y^*] \quad (1.8)$$

where $|\cdot\rangle$ is the *ket* and identifies a column vector, and its transpose-conjugate $\langle\cdot|$ is called *bra*, with superscript $*$ indicating the complex conjugate.

The PDM field can thus be written in a vectorial form as:

$$|E(0,t)\rangle = \sum_{j \in x,y} \sum_{h=1}^{N_{\text{ch}}} \sum_{k=-\infty}^{\infty} \frac{a_{khj}}{\sqrt{2}} |G_{kh}(0,t)\rangle \quad (1.9)$$

where a_{khj} is the digital symbol at time k , WDM channel-index h , and polarization j , while the basis function generalizes to the following expression:

$$|G_{kh}(0,t)\rangle \triangleq p_h(t - kT_h) e^{j\omega_h t} |j\rangle \quad (1.10)$$

being $|j\rangle$ a two entry unit-vector identifying the j th polarization. The factor $1/\sqrt{2}$ in Eq. (1.9) accounts for the fact that each polarization carries 1/2 of the average transmitted power P .

The field in Eq. (1.8) can also be expressed to highlight the so-called field's state of polarization (SOP) with the following formalism $|E\rangle = E_0 \hat{J}$, where $E_0 = \sqrt{|E_x|^2 + |E_y|^2}$ is the field amplitude while the vector \hat{J} is a unit Jones vector representing the field SOP [8].

1.3.3 Space-division multiplexing

Multiplexing in frequency and polarization, exploiting both in-phase and quadrature components, allows increasing the amount of transmitted data. A possible solution to further increase the capacity is to unleash the spatial dimension in space-division-multiplexed systems. There are several technical approaches to achieve space-division multiplexing [9], as depicted in a schematic way in Fig. 1.6.

The most intuitive approach to perform spatial multiplexing is to use a bundle of optical fibers, each supporting a single-mode, as represented in Fig. 1.6(a). However, the cost per bit of this solution remains unchanged, thus making the total investment too expensive. A first SDM solution that may reduce the cost per bit is a multi-core fiber (MCF), see Fig. 1.6(b). In a multi-core fiber, spatial multiplexing is achieved by physically distinct cores integrated directly into the fiber cross-section. Another SDM option is represented by multi-mode fibers (MMFs), as sketched in Fig. 1.6(c). Here the core is usually larger than in MCF, thus supporting more spatial modes. In this thesis, SDM is intended as achieved through the latter type of fibers.

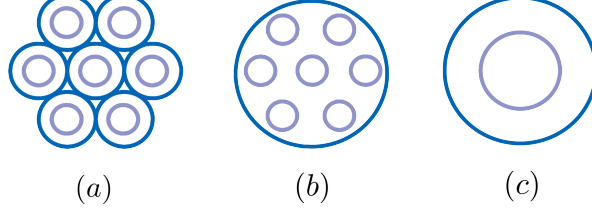


Figure 1.6: Different techniques for space-division multiplexing. Fiber bundle (a), multi-core fiber (b), and multi-mode fiber (c).

MCF and MMF may reduce the cost per bit once used with an optical amplifier that integrates the spatial dimension as well. However, the major problem of such systems is the presence of crosstalk among the modes, which can arise either by linear or nonlinear effects.

For an SDM transmission, the field complex envelope can be generalized as

$$|E(0, t)\rangle = \sum_{j \in \mathcal{D}} \sum_{h=1}^{N_{\text{ch}}} \sum_{k=-\infty}^{\infty} \frac{a_{khj}}{\sqrt{2}} |G_{kh}(0, t)\rangle \quad (1.11)$$

being \mathcal{D} the set of modes supported by the optical fiber. The set of modes has dimension $2N_{\text{m}}$, with the factor 2 accounting for the two polarizations per spatial mode. In this scenario, the field envelope $|E\rangle$ having $2N_{\text{m}}$ entries is called *hyper-polarization* vector [10] and generalizes the two-dimensional Jones vector of the PDM transmission. The average power P is intended as power per-mode, hence the factor $1/\sqrt{2}$ appears in Eq. (1.11) as in Eq. (1.9).

Atomic notation

We now introduce a useful notation that compacts together the time, frequency, and polarization/mode index as follows:

$$\sum_{n_3 \in \mathcal{D}} \sum_{n_2=1}^{N_{\text{ch}}} \sum_{n_1=-\infty}^{\infty} a_{n_1 n_2 n_3} \rightarrow \sum_{\mathbf{n}} a_{\mathbf{n}} \quad (1.12)$$

where $a_{\mathbf{n}}$ is the generic information symbol per (time, frequency, polarization/mode) channel use, i.e., an *atom* of source information, and the vector \mathbf{n} should be read as:

$$\mathbf{n} = \underset{\text{time}}{\uparrow} [n_1, \underset{\text{frequency}}{\uparrow} n_2, \overset{\text{mode}}{\downarrow} n_3] \quad (1.13)$$

According to this atomic notation, the field envelope of a PDM or SDM transmitted field writes as:

$$|E(0, t)\rangle = \sum_{\mathbf{n}} a_{\mathbf{n}} |G_{\mathbf{n}}(0, t)\rangle \quad (1.14)$$

with basis function:

$$|G_{\mathbf{n}}(0, t)\rangle \triangleq p(t - n_1 T_{n_2}) e^{j\omega_{n_2} t} |n_3\rangle \quad (1.15)$$

where the dimension of the vector $|n_3\rangle$ depends on the number of supported modes.

1.4 The fiber-optic channel

An optical fiber supports the propagation of an electromagnetic wave confined within it. Such a wave experiences many effects during propagation, as we show now. We start with the simplest case of pure scalar propagation, i.e., without any polarization effect.

1.4.1 Scalar propagation

The propagation of a field envelope along the optical fiber is governed by the following nonlinear Schrödinger equation (NLSE), which is derived from the Maxwell's equations [7]:

$$\begin{aligned} \frac{\partial E(z, t)}{\partial z} = & -\frac{\alpha}{2} E(z, t) - j\beta_0 E(z, t) - \beta_1 \frac{\partial E(z, t)}{\partial t} + j\frac{\beta_2}{2} \frac{\partial^2 E(z, t)}{\partial t^2} + \frac{\beta_3}{6} \frac{\partial^3 E(z, t)}{\partial t^3} \\ & - j\gamma \left((1 - f_R) |E(z, t)|^2 + f_R \int_0^t h_R(t - t') |E(z, t')|^2 dt' \right) E(z, t). \end{aligned} \quad (1.16)$$

The first row on the right-hand side accounts for linear effects experienced by the electric field, while the second row accounts for nonlinear effects. We describe them separately as follows.

Linear effects

The first term on the right-hand side of Eq. (1.16) is the attenuation which accounts for the vanishing of the signal along z due to the fiber loss. The fiber attenuation coefficient α , which is wavelength-dependent, is typically expressed in dB/km.

An example of fiber attenuation versus wavelength for single mode-fiber is reported in Fig. 1.7, taken from [11]. The wavelength-dependent fiber loss is caused by both intrinsic and extrinsic phenomena. In particular, intrinsic mechanisms such as Rayleigh scattering and infrared absorption [7] set the fundamental theoretical minimum value of attenuation to about 0.16 dB/km. On the other hand, extrinsic factors such as manufacturing imperfections can enhance fiber losses. For instance, Fig. 1.7 shows the so-called OH peaks around 1244 and 1390 nm, which occur due to the absorption of water molecules residual from the manufacturing process.

The wavelength dependence of the attenuation profile implies the existence of wavelength ranges more favorable for communication purposes, called telecommunication windows. The current window adopted for optical communication is centered at 1550 nm, where the optical fiber exhibits a minimum value in the attenuation of ≈ 0.2 dB/km for SMF. Such a conventional bandwidth (C-band) covers the interval 1530 – 1565 nm. Despite the slightly higher attenuation coefficient, the long wavelengths (L-band) in the range 1565 – 1625 nm are currently adopted to further exploit the wide bandwidth of the optical fiber. However, in modern optical fibers, the OH peak can be successfully removed by a proper manufacturing process [9]. As a result, low-attenuation transmissions are feasible over wider bandwidths. Other standardized bandwidths that exhibit an attenuation coefficient lower than 0.4 dB in OH peak-free fiber are the S-band (1460 – 1530 nm), the E-band (1360 – 1460 nm), and the O-band (1260 – 1360 nm).

If the bandwidth of $E(z, t)$ is confined within a telecommunication window, the wavelength dependence of α can be safely neglected, as in Eq. 1.16.

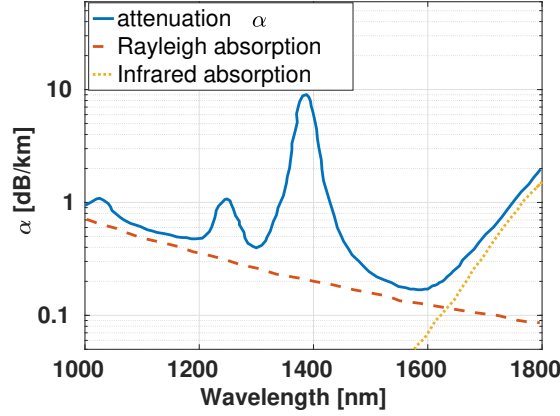


Figure 1.7: Attenuation profile of a single-mode fiber. Data from [11].

The other coefficients related to linear effects occurring during the field propagation along z are β_n , $n = 1, 2, 3$, which are the coefficients of the Taylor expansion of the propagation constant $\beta(\omega)$, usually truncated up to the third order:

$$\beta(\omega) = \beta_0 + \beta_1\omega + \frac{1}{2}\beta_2\omega^2 + \frac{1}{6}\beta_3\omega^3 \quad (1.17)$$

Formally, the ℓ th coefficient can be found as:

$$\beta_\ell \triangleq \left. \frac{d^\ell \beta(\omega)}{d\omega^\ell} \right|_{\omega=0} \quad (1.18)$$

where $\omega = 0$ is the low-pass frequency corresponding to the carrier frequency.

To each coefficient, we can associate a physical action. For instance, the term β_0 in Eq. (1.16) is a constant phase rotation, while β_1 induces a group velocity v_g with $\beta_1 = 1/v_g$. The group velocity describes the velocity at which the field envelope travels. Hence, because of β_1 , a signal after z km experiences a group delay, or walk-off, of $\beta_1 z$. The second-order coefficient β_2 accounts for the group velocity dispersion (GVD), i.e., the dependence of group velocity with frequency. Such a frequency dependence causes pulse broadening since different pulse spectral components arrive at different time instants at the fiber output. The delay spread ΔT of two frequencies spaced $\Delta\omega$ due to GVD after L km of optical fiber can be found as [12]:

$$\Delta T \triangleq L\beta_2\Delta\omega \quad (1.19)$$

with $\Delta\omega$ the width of the transmitted pulse. Such a phenomenon, also known as chromatic dispersion [7], can be highly detrimental if not compensated properly since pulse broadening generates inter-symbol interference on the digital signal.

A measure of chromatic dispersion is typically provided in terms of the fiber dispersion coefficient D [ps/nm/km], which is related to β_2 through the following:

$$D \triangleq -\frac{2\pi c}{\lambda_0^2}\beta_2 \quad (1.20)$$

where λ_0 is the wavelength at which we evaluated β_2 , with c the speed of light in the vacuum. If we change the reference wavelength, the dispersion coefficient changes value because of third-order dispersion, usually called fiber slope S ps/nm²/km. The slope is proportional to the β_3 parameter by

$$S = \left(\frac{2\pi c}{\lambda_0^2}\right)^2 \beta_3 + \left(\frac{4\pi c}{\lambda_0^3}\right) \beta_2. \quad (1.21)$$

As a result, the fiber dispersion is wavelength-dependent, as depicted in Fig. 1.8 for a standard single-mode fiber [12]. The figure shows that the SMF exhibits a dispersion coefficient value ≈ 17 ps/nm/km at the reference wavelength 1550 nm associated with the minimum of fiber attenuation. The zero-dispersion wavelength for the SMF is near 1300 nm. Since the dispersion curve in Fig. 1.8 is due to both waveguide and material dispersion, it is possible to design optical fibers such that the dispersion curve is shifted. In particular, fibers having zero dispersion at 1550 nm were manufactured to avoid pulse broadening, called dispersion-shifted fibers (DSF). While the zero-dispersion coefficient yields benefits from the linear propagation point of view, it is extremely detrimental when nonlinear effects arise. For this reason, non-zero dispersion-shifted fibers (NZDSF) with low dispersion in the range 4-8 ps/nm/km around 1550 nm were designed.

In the presence of only chromatic dispersion and third-order dispersion, the solution of the NLSE is:

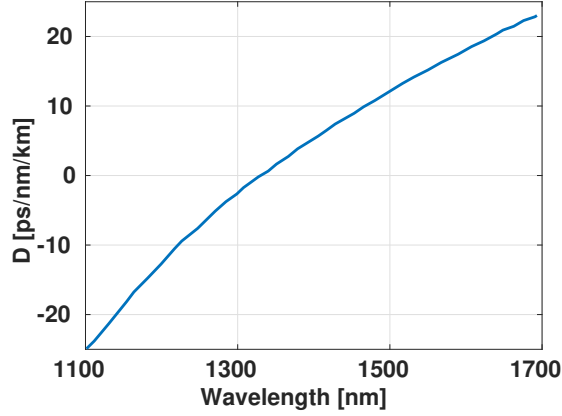


Figure 1.8: Dispersion coefficient vs. wavelength of a single-mode fiber. Data from [12].

$$E(z, t) = \mathcal{F}^{-1} \left\{ \tilde{E}(0, \omega) e^{-j \left(\frac{\beta_2}{2} \omega^2 + \frac{\beta_3}{6} \omega^3 \right) z} \right\} \quad (1.22)$$

hence a pure phase-shift in the frequency domain, and a pulse broadening in the time domain [7]. For ultra-wideband transmission even α is a function of ω . It can be included in the frequency domain solution Eq. (1.22) by $e^{-\frac{\alpha(\omega)}{2} z}$.

Nonlinear effects

The optical fiber manifests an instantaneous nonlinear response to an intense electric field, resulting in the so-called Kerr effect. Such an effect exhibits a memoryless nature coming from the instantaneous response of the glass electrons to the field. As a consequence, the fiber refractive index shows an electric field dependence which, in turn, causes a transfer of energy among frequencies. The memoryless contribution is described by $-j\gamma|E(z, t)|^2 E(z, t)$ in Eq. (1.16), with γ the nonlinear coefficient of the optical fiber:

$$\gamma \triangleq \frac{2\pi n_2}{\lambda_0 A_{\text{eff}}} \quad (1.23)$$

with n_2 the fiber nonlinear index and A_{eff} the fiber effective area.

It is convenient to work in the following phase-shifted, pre-attenuated, and retarded reference system:

$$A(z, t) = E(z, t - \beta_1 z) e^{-\frac{\alpha}{2} z - j\beta_0 z} \quad (1.24)$$

such that the NLSE in Eq. (1.16), with $f_R = 0$ (i.e., in the absence of the Raman effect), rewrites as:

$$\frac{\partial A(z, t)}{\partial z} = j \frac{\beta_2}{2} \frac{\partial^2 A(z, t)}{\partial t^2} + \frac{\beta_3}{6} \frac{\partial^3 A(z, t)}{\partial t^3} - j\gamma e^{-\alpha z} |A(z, t)|^2 A(z, t). \quad (1.25)$$

In the presence of only the Kerr effect and attenuation, the NLSE can be easily solved in a closed-form expression. The solution is found to be a signal power-dependent phase shift, namely:

$$A(z, t) = e^{-j\gamma L_{\text{eff}}(z) |A(0, t)|^2} A(0, t) \quad (1.26)$$

with $\gamma L_{\text{eff}}(z) |A(0, t)|^2$ a nonlinear phase. L_{eff} is the effective length of the Kerr effect, defined as:

$$L_{\text{eff}}(z) \triangleq \int_0^z e^{-\alpha \xi} d\xi = \frac{1 - e^{-\alpha z}}{\alpha}. \quad (1.27)$$

As a reference, the effective length of an SMF fiber of 100 km having $\alpha = 0.2$ dB/km is $L_{\text{eff}} \approx 21.5$ km, meaning that the Kerr effects are more significant at the fiber beginning due to the signal power loss along distance caused by the fiber loss. It is worth noting that, although the effect of Kerr nonlinearity in the time domain is a pure phase-shift, it yields a spectrum broadening in the frequency domain [7].

In the presence of wavelength-division multiplexing, a frequency disaggregated or *separate-field* point of view [7] can be adopted to simplify the analysis. The electric field can thus be written as the superposition of the electric fields associated with each channel in a WDM system. Each WDM channel travels with its phase shift, group velocity, and GVD, hence with coefficients:

$$\beta_{\ell i} \triangleq \left. \frac{d^\ell \beta(\omega)}{d\omega^\ell} \right|_{\omega=\omega_i} \quad (1.28)$$

where $\beta_{\ell i}$ is the ℓ th order Taylor coefficient relative to channel i . More explicitly, they are related to the coefficients β_i , introduced in Eq. (1.18) at the reference frequency $\omega = 0$, by:

$$\beta_{0i} \triangleq \beta_1 \omega_i + \frac{\beta_2}{2} \omega_i^2 + \frac{\beta_3}{6} \omega_i^3 \quad (1.29)$$

$$\beta_{1i} \triangleq \beta_1 + \beta_2 \omega_i + \frac{\beta_3}{2} \omega_i^2 \quad (1.30)$$

$$\beta_{2i} \triangleq \beta_2 + \beta_3 \omega_i \quad (1.31)$$

$$\beta_{3i} \triangleq \beta_3. \quad (1.32)$$

The separate-field point of view thus requires expressing the propagation of the WDM signal along the optical fiber through a set of differential equations, one for each channel. For instance, the evolution along z of the generic channel A_i centered at frequency ω_i writes in terms of separate-field NLSE as:

$$\begin{aligned} \frac{\partial A_i(z, t)}{\partial z} = & -\beta_{1i} \frac{\partial A_i(z, t)}{\partial t} + j \frac{\beta_{2i}}{2} \frac{\partial^2 A_i(z, t)}{\partial t^2} + \frac{\beta_{3i}}{6} \frac{\partial^3 A_i(z, t)}{\partial t^3} \\ & - j \gamma e^{-\alpha z} \sum_{h, m, n \in \mathcal{B}} A_h^*(z, t) A_m(z, t) A_n(z, t) e^{-j \Delta \beta_{hmni} z} \end{aligned} \quad (1.33)$$

where the summation in the nonlinear term is limited to the set:

$$\mathcal{B} = \{(h, m, n) : \omega_m + \omega_n - \omega_h - \omega_i = 0\} \quad (1.34)$$

ensuring the conservation of energy [7], and $\Delta \beta_{hmni}$ is the *phase matching coefficient* defined as:

$$\Delta \beta_{hmni} \triangleq \beta_{0m} + \beta_{0n} - \beta_{0h} - \beta_{0i}. \quad (1.35)$$

Depending on the values of h, m, n, i , we can identify different implications of the Kerr effects on the channels. The effect that accounts for intra-channel nonlinearity is

self-phase modulation (SPM) [7] and it occurs when $h = m = n = i$. On the other hand, when the terms in Eq. (1.35) are pair-wise degenerate, e.g., $(h = m) \neq (n = i)$ or $(h = n) \neq (m = i)$, the nonlinear Kerr effect manifests as cross-phase modulation (XPM) which accounts for the nonlinear interaction between two non-degenerate frequency channels. In both cases, a phase-matching condition is satisfied, resulting in $\Delta\beta = 0$.

Other valid configurations of the frequency indices generally resulting in $\Delta\beta \neq 0$ yield to the four-wave mixing (FWM) effect [7]. FWM is a nonlinear process that transfers energy among four frequencies satisfying the principle of energy conservation in Eq. (1.34). As a consequence of FWM, energy can arise on an idle frequency if the other three involved frequencies have a non-zero power at $z = 0$.

The glass nuclei respond to the electric field in a non-instantaneous way. This response is known as the vibrational or Raman response. The second nonlinear term in Eq. (1.16) models the Raman effect, caused by the non-instantaneous fiber response:

$$-j\gamma f_R \int_0^t h_R(t-t') |E(z,t')|^2 dt' E(z,t) \quad (1.36)$$

where f_R is the fractional Raman contribution, whose value is ≈ 0.2 for single-mode fibers [10, 13]. It can be seen that Eq. (1.36) expresses a convolution between the function h_R and the signal power, where h_R is the Raman response function [7]. Of particular interest is the Fourier transform of the Raman response $\tilde{h}_R(\omega)$, which exhibits a real and imaginary part, namely:

$$\tilde{h}_R(\omega) = \tilde{r}_R(\omega) + j\tilde{g}_R(\omega) \quad (1.37)$$

where the real part is responsible for Raman-induced index changes [14] while $\tilde{g}_R(\omega)$ is the Raman gain spectrum [14, 15]. The Raman gain is responsible for the so-called Stimulated Raman Scattering (SRS), a nonlinear process that amplifies lower frequencies while depleting the higher ones [7]. While the real part of the Raman response is typically overlooked because it yields minor implications [16], the Raman gain represents a non-negligible impairment for wide bandwidth transmissions. In such cases, it is possible to approximate the NLSE expression by replacing the fiber losses with a more generic term $\mathcal{A}(z, \omega)$. Namely, in the frequency domain:

$\alpha \rightarrow \mathcal{A}(z, \omega)$, thus requiring a convolution in the time domain. Such a frequency-dependent term $\mathcal{A}(z, \omega)$ accounts for the losses/gains experienced along the optical fiber due to both fiber loss and Raman gain. For the sake of completeness, the NLSE in the frequency domain is then:

$$\begin{aligned} \frac{\partial \tilde{E}(z, \omega)}{\partial z} \simeq & -\frac{\mathcal{A}(z, \omega)}{2} \tilde{E}(z, \omega) - j\beta(\omega) \tilde{E}(z, \omega) \\ & - j\gamma(1 - f_R) \tilde{E}(z, \omega) \otimes \tilde{E}^*(z, -\omega) \otimes \tilde{E}(z, \omega) \end{aligned} \quad (1.38)$$

with \otimes denoting the convolution operation. Hence, the signal power at a given coordinate z is related to the input power by the following power profile:

$$\rho(z, \omega) = e^{-\int_0^z \mathcal{A}(\zeta, \omega) d\zeta} \quad (1.39)$$

which reduces to $e^{-\alpha z}$ in the absence of Raman scattering. The Raman effect will be treated in more detail in Chapter 3.

1.4.2 Dual-polarization propagation

Exploiting the two polarizations of the field allows increasing the amount of transmitted data. However, the propagation of the two polarizations cannot be simply modeled as a pair of NLSE in Eq. (1.16), due to the presence of polarization-related impairments, such as the fiber birefringence. Birefringence is the property of exhibiting different refractive indexes for different polarizations, and it arises from broken circular symmetry of the optical fiber, due to, e.g., imperfections, external stress, or bending. Birefringence, besides changing the propagation constant of the propagating polarizations, causes an exchange of power between them over a period called beat length, which is defined as $L_B \triangleq \frac{2\pi}{|\beta_{0,x} - \beta_{0,y}|}$ [m] where $|\beta_{0,x} - \beta_{0,y}|$ is the birefringence strength.

Due to random perturbations, polarization mode dispersion (PMD) [7, 8] arises. In particular, the two polarizations are delayed one respect to the other, by random differential group delay (DGD) [17].

The evolution along the optical fiber of a PDM signal is thus governed by a pair of coupled differential equations, one describing the propagation of each polariza-

tion. The two equations can be compacted together in a unique vectorial equation, the so-called coupled NLSE (CNLSE). Neglecting the Raman effect for the sake of simplicity¹, the CNLSE writes as

$$\begin{aligned} \frac{\partial |E\rangle}{\partial z} = & -\frac{\alpha}{2} |E\rangle - j\mathbf{B}^{(0)} |E\rangle - \mathbf{B}^{(1)} \frac{\partial |E\rangle}{\partial t} + j\frac{\beta_2}{2} \frac{\partial^2 |E\rangle}{\partial t^2} + \frac{\beta_3}{6} \frac{\partial^3 |E\rangle}{\partial t^3} \\ & - j\gamma \left(\langle E|E\rangle |E\rangle - \frac{1}{3} \langle E| \sigma_3 |E\rangle \sigma_3 |E\rangle \right) \end{aligned} \quad (1.40)$$

with σ_3 the following Pauli matrix:

$$\sigma_3 = \begin{bmatrix} 0 & j \\ -j & 0 \end{bmatrix}. \quad (1.41)$$

The matrices $\mathbf{B}^{(0)}$ and $\mathbf{B}^{(1)}$ are the birefringence and PMD matrix, respectively. Mathematically, the matrix $\mathbf{B}^{(n)}$ with $n = 0, 1$ appearing in Eq. (1.40) has a deterministic and a perturbation-related term as follows:

$$\mathbf{B}^{(n)} \triangleq \mathbf{D}^{(n)} + \mathbf{P}^{(n)} \quad (1.42)$$

where the matrix \mathbf{D} collects the deterministic terms that are proper of the fiber under test. Such a deterministic matrix is a 2×2 diagonal matrix with non-zero elements equal to the propagation constant $\beta_{0,p}$ for $\mathbf{D}^{(0)}$ and the inverse group velocity $\beta_{1,p}$ in $\mathbf{D}^{(1)}$, with $p = x, y$. On the other hand, the matrix \mathbf{P} accounts for the effects arising from perturbations, such as birefringence for $n = 0$ and PMD for $n = 1$. Please note that, due to the presence of random perturbations, the CNLSE is a stochastic differential equation.

Due to the fiber imperfections, the two polarizations mix together during propagation. These dynamics occur in a length scale usually much smaller than the length scale over which the nonlinear effects change. Such a difference in the length scales allows to average the nonlinear term of Eq. (1.40) with respect to the fast rotations of

¹Note that the Raman gain can be included by substituting the attenuation coefficient with the proper profile.

$|E\rangle$, according to the method of averaging. The result of the averaging yields to the well-known Manakov-PMD equation [18]:

$$\begin{aligned} \frac{\partial |E\rangle}{\partial z} = & -\frac{\alpha}{2} |E\rangle - \mathbb{B}^{(1)} \frac{\partial |E\rangle}{\partial t} + j \frac{\beta_2}{2} \frac{\partial^2 |E\rangle}{\partial t^2} + \frac{\beta_3}{6} \frac{\partial^3 |E\rangle}{\partial t^3} \\ & - j \kappa \gamma \langle E|E \rangle |E\rangle \end{aligned} \quad (1.43)$$

where κ is the Manakov coefficient, resulting from the averaging process, and $\mathbb{B}^{(1)}$ is a rotated version of $\mathbf{B}^{(1)}$ in the reference system that follows $\mathbf{B}^{(0)}$. In single-mode fiber, the coefficient assumes the value $\kappa = \frac{8}{9}$.

In the absence of PMD, Eq. (1.43) becomes completely deterministic, and it is simply known as Manakov equation. In the phase-rotated, retarded and pre-attenuated reference systems considered in Eq. (1.24), the Manakov equation becomes:

$$\frac{\partial |A\rangle}{\partial z} = j \frac{\beta_2}{2} \frac{\partial^2 |A\rangle}{\partial t^2} + \frac{\beta_3}{6} \frac{\partial^3 |A\rangle}{\partial t^3} - j \gamma \kappa e^{-\alpha z} \langle A|A \rangle |A\rangle. \quad (1.44)$$

As for the scalar case, it is useful to split the propagation equation of the WDM signal into a system of differential equations, one for each WDM channel, according to the separate-field point of view. For instance, for a channel centered at frequency ω_i the Manakov equation writes as:

$$\frac{\partial |A_i\rangle}{\partial z} = j \frac{\beta_{2i}}{2} \frac{\partial^2 |A_i\rangle}{\partial t^2} + \frac{\beta_{3i}}{6} \frac{\partial^3 |A_i\rangle}{\partial t^3} - j \gamma \kappa e^{-\alpha z} \sum_{h,m,n \in \mathcal{B}} \langle A_h | A_m \rangle |A_n\rangle e^{-j \Delta \beta_{h m n i} z} \quad (1.45)$$

with \mathcal{B} the same set of channel indexes commented in the scalar case, and $\Delta \beta_{h m n i}$ the phase matching coefficient. Such a separate-field approach allows a meaningful insight on the Kerr effect. In particular, focusing on phase-matched combinations of the channel indexes, the nonlinear term can be expressed as the summations of different contributions. If we neglect FWM, channel i experiences the following Kerr effect:

$$\langle A_h | A_m \rangle |A_n\rangle = \underbrace{\langle A_i | A_i \rangle |A_i\rangle}_{\text{SPM}} + \sum_{n \neq i} \left(\underbrace{\langle A_n | A_n \rangle |A_i\rangle}_{\text{XPM}} + \underbrace{|A_n\rangle \langle A_n | A_i\rangle}_{\text{XPoIM}} \right) \quad (1.46)$$

where XPolM stands for cross-polarization modulation (XPolM) [19]. It is worth noting that, while scalar nonlinear effects induce a phase shift on the signal, XPolM induces a rotation of the state of polarization on the Poincaré sphere [19].

1.4.3 SDM propagation

In SDM, different spatial modes travel with different propagation constants. However, for some modes, such a difference is usually small. A set of spatial modes with an almost identical propagation constant is called a *group* of quasi-degenerate modes [10]. In these fibers, the modes couple along propagation in a random way due to refractive index perturbations [20]. On a length scale of meters, only the two polarizations of the same mode couple along propagation, as in single-mode fibers. After hundred of meters, the coupling between quasi-degenerate modes within a group becomes non-negligible. As the propagation distance increases, all the optical fiber modes couple. However, over short distances, the strength of the coupling between modes belonging to different groups may be moderate compared to the strong coupling of modes within the same group. For any sufficiently long propagation distance, all the modes supported by the fiber couple together. For the sake of simplicity, we will focus only on this latter regime of linear strong mode coupling.

In the strong coupling regime, the Manakov equation was extended to model SDM propagation [10] supporting $2N_m$ modes as:

$$\begin{aligned} \frac{\partial |E\rangle}{\partial z} = & -\mathbf{A}|E\rangle - \mathbb{B}^{(1)} \frac{\partial |E\rangle}{\partial t} + j\frac{1}{2}\mathbb{B}^{(2)} \frac{\partial^2 |E\rangle}{\partial t^2} + \frac{1}{6}\mathbb{B}^{(3)} \frac{\partial^3 |E\rangle}{\partial t^3} \\ & - j\kappa\gamma\langle E|E\rangle |E\rangle \end{aligned} \quad (1.47)$$

where $|E\rangle$ is now the hyper-polarization vector with $2N_m$ entries and the coefficient matrices are now $2N_m \times 2N_m$. In the most general case, all the coefficients can be mode-dependent, including the fiber attenuation which is now generalized through the matrix \mathbf{A} . As in dual-polarization cases, we adopted a reference system tracking $\mathbf{B}^{(0)}$ perturbations, which is now $2N_m \times 2N_m$. Within a group of modes, the quasi degeneracy makes the deterministic contribution in $\mathbb{B}^{(1)}$ well described by a scalar

effect, hence with matrices proportional to the identity matrix. On the other hand, the matrices $\mathbb{B}^{(2)}$ and $\mathbb{B}^{(3)}$ are well described by deterministic diagonal matrices, in most of the fibers supporting SDM, collecting on the diagonal elements the β_2 and β_3 coefficient of each mode.

Regarding the nonlinear term, the Manakov coefficient κ resulting from the averaging process must be re-defined in this framework, and the nonlinear coefficient should be scaled as $\gamma_{\text{SDM}} = \gamma_{\text{SMF}}/N_m$ [10]. The latter scaling can be included in the re-definition of the Manakov coefficient:

$$\kappa = \frac{4}{3} \frac{2}{2N_m + 1} \quad (1.48)$$

by keeping the nonlinear coefficient γ unchanged, as defined in Eq. (1.23).

1.5 Approximate solutions of the Manakov equation

Due to the absence of a general analytical solution for the field propagation along the optical fiber in the presence of both linear and nonlinear impairments, two strategies are viable: approximated theoretical approaches or numerical methods. We briefly review two popular methods in this framework, i.e, the first-order regular perturbation method and the split-step Fourier method algorithm. We focus on the single-mode Manakov equation in the absence of polarization-related effects.

1.5.1 First-order regular perturbation

An analytical technique to approximate the Manakov equation is represented by the first-order regular perturbation (RP1) method [21, 22], which consists in expanding the field in a first-order series with respect to the nonlinear coefficient γ . The RP1 solution of the single-mode Manakov equation is

$$|A(z, t)\rangle \simeq |A\rangle^{(0)} + \gamma |A\rangle^{(1)} \quad (1.49)$$

where the first term on the right-hand side is the zeroth-order solution, hence the solution of the linear propagation, and is given by the following expression:

$$|A\rangle^{(0)} = e^{\mathcal{L}z} |A(0, t)\rangle \quad (1.50)$$

where \mathcal{L} is the linear operator. In the considered phase-rotated, retarded and pre-attenuated reference system, such an operator is defined in the time domain as:

$$\mathcal{L} \triangleq j \frac{\beta_2}{2} \frac{\partial^2}{\partial t^2} + \frac{\beta_3}{6} \frac{\partial^3}{\partial t^3} \quad (1.51)$$

with a much more convenient frequency-domain representation:

$$\mathcal{F}\{e^{\mathcal{L}z}\} \triangleq e^{-j\left(\frac{\beta_2}{2}\omega^2 + \frac{\beta_3}{6}\omega^3\right)z}. \quad (1.52)$$

The first-order solution in Eq. (1.49) is found as follows [22]

$$\gamma |A\rangle^{(1)} = \int_0^z e^{\mathcal{L}(z-\zeta)} \mathcal{N}\left(e^{\mathcal{L}\zeta} |A(0, t)\rangle\right) d\zeta \quad (1.53)$$

where we defined the nonlinear operator \mathcal{N} applied to $|A(z, t)\rangle$ as

$$\mathcal{N}(|A(z, t)\rangle) \triangleq -j\gamma \frac{8}{9} e^{-\alpha z} \langle A(z, t) | A(z, t) \rangle |A(z, t)\rangle. \quad (1.54)$$

After performing a zero-forcing equalization of the linear effects, i.e., concatenation with $e^{-\mathcal{L}z}$, the received signal $|A_R\rangle$ is:

$$|A_R\rangle \simeq |A(0, t)\rangle + |n(z, t)\rangle \quad (1.55)$$

where $|n(z, t)\rangle$ represents the interference with respect to the desired signal due to the presence of Kerr effects, and it is known in the literature as the *nonlinear interference* (NLI). The RP1 solution is the key ingredient to evaluate the variance of the NLI, a topic that will be covered in Chapter 2.

1.5.2 Split-step Fourier method

The most widely used numerical method to solve the field propagation along the optical fiber is the split-step Fourier method (SSFM) [7, 23, 24]. SSFM is a pseudo-spectral method which owes its name to back and forth Fourier transforms across

several spatial steps. To understand such a method, it is convenient to read Eq. (1.44) in terms of linear and nonlinear operators, namely:

$$\frac{\partial |A\rangle}{\partial z} = \underbrace{j \frac{\beta_2}{2} \frac{\partial^2 |A\rangle}{\partial t^2} + \frac{\beta_3}{6} \frac{\partial^3 |A\rangle}{\partial t^3}}_{\mathcal{L}(|A\rangle)} \underbrace{- j\gamma \frac{8}{9} e^{-\alpha z} \langle A|A\rangle |A\rangle}_{\mathcal{N}(|A\rangle)} \quad (1.56)$$

where the linear and nonlinear operators are defined in Eq. (1.51) and Eq. (1.54), respectively. Although the linear and nonlinear effects occur jointly during propagation, the key idea behind SSFM is to treat the two operators as independent across a small spatial step h . Leveraging the operators' independence, over a distance h the linear and nonlinear propagation can be solved separately. In fact, in the absence of nonlinear effects, the propagation has a solution as in Eq. (1.22). Such a *linear step* is efficiently solved in the frequency domain, in which GVD is a pure phase rotation. On the other hand, the propagation in the presence of nonlinear effects and fiber attenuation over h , called the *nonlinear step*, has a simple closed-form solution in the time domain, as reported in Eq. (1.26).

To solve the propagation along the whole fiber length, the SSFM thus splits the optical fiber length into N_h steps, each composed of a linear and nonlinear step. The concatenation of each nonlinear and linear step (and vice versa) is interleaved by a Fourier transform, to move from time to frequency domain (and vice versa). A qualitative sketch of the SSFM approach is represented in Fig. 1.9. The figure depicts the most simple SSFM scheme, in which the linear and nonlinear operators are applied over the same length h , which is the same for all the N_h steps. This scheme represents an asymmetric SSFM approach with uniform step size [25].

The step choice is a crucial point in the SSFM since it determines its solution accuracy. On this note, it was shown in [26] that applying the operators' concatenation $\mathcal{L}\mathcal{N}\mathcal{L}$, with the linear operator \mathcal{L} over a length $h/2$, yields a smaller error. Such a step choice is called symmetric step approach [27], opposite to the asymmetric step depicted in Fig. 1.9. Apart from the concatenation order of \mathcal{L} and \mathcal{N} within the same step, the other two aspects setting the accuracy of the SSFM are the length of the first step and the step updating rule, which, together, determine the number of total steps in which the fiber is discretized. The first step should be tailored to correctly capture

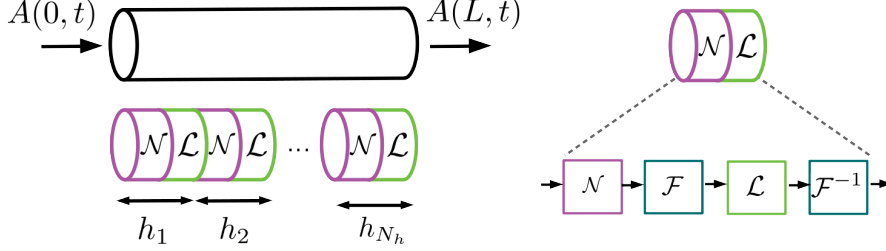


Figure 1.9: Sketch of the SSFM idea: concatenation of linear and nonlinear small steps, with back and forth Fourier transforms.

the FWM impact, as proposed in [28]:

$$h_1 = \frac{\Delta\phi_{\text{FWM}}}{|\beta_2|(2\pi B_{\text{WDM}})^2} \quad (1.57)$$

where $\Delta\phi_{\text{FWM}}$ is the maximum tolerable FWM phase shift, and B_{WDM} is the WDM signal bandwidth. Regarding the updating rule, the constant local error (CLE) criterion allows accurate SSFM implementation by keeping constant along distance the SSFM error [26, 29]. Another well-known criterion for the step update is the nonlinear phase criterion (NLP) [27], which updates the step such that the nonlinear phase accumulated within a step is below a maximum tolerable value. In both cases, the updating rule can be written as:

$$h_{k+1} \simeq h_k e^{\frac{\alpha}{q} h_k} \quad (1.58)$$

where the parameter q takes value 1 for the NLP, and 2 or 3 respectively for the CLE with symmetric or asymmetric step [28].

An accurate SSFM based simulation requires propagating an adequate number of data symbols to capture the worst-case walk-off among channels, thus avoiding artificial periodic crosstalk resonances during propagation. Since i) the walk-off scales with the inverse of the bandwidth, ii) the number of discrete points per symbol scales with bandwidth, and iii) the step scales with the inverse of the square of the bandwidth (see Eq. (1.57)), the computational cost of a simulation scales at least with the

fourth power of the bandwidth. By including also the cost of the fast Fourier transform (FFT) in the discussion, in [30] we showed that the number of multiplications scales as:

$$C_m = \mathcal{O}(|\beta_2| B_{\text{WDM}}^4 L^2 (\log_2(|\beta_2| B_{\text{WDM}}^2 L) + 1 + \kappa_e)) . \quad (1.59)$$

where $\mathcal{O}(\cdot)$ is the Big O Landau symbol, κ_e represents the cost of an exponentiation, while the cost of FFT was considered to scale with $N_{\text{FFT}} \log_2 N_{\text{FFT}}$, with N_{FFT} the number of NFFT samples.

1.6 Multi-span optical-links

The simplest point-to-point optical link can be seen as a concatenation of optical fibers, each followed by an optical amplifier recovering the intrinsic fiber loss. A lumped optical amplifier is a device that allows recovering the signal power without the necessity of optical-to-electric conversion. Nowadays, the most widely used optical amplifier is the Erbium-doped fiber amplifier (EDFA), which belongs to the category of rare-earth-doped fiber amplifiers and provides amplification in both the C and L bands. In brief, an EDFA is an optical fiber doped with Erbium ions, capable of providing gain via stimulated emission. For stimulated emission to take place, a population inversion is necessary, which is obtained by pumping at suitable wavelengths, such as 980 nm or 1480 nm [12]. In addition to desired stimulated emission of photons, the amplifier also spontaneously emits photons out of phase with respect to the signal photons. Such a spontaneous emission represents a source of noise, yielding the so-called amplified spontaneous emission (ASE) noise. The noise introduced by each amplifier thus propagates for the remainder of the link. As a consequence, the received signal in Eq. (1.55) must be extended to account for the received ASE noise $|w(z, t)\rangle$, namely:

$$|A_R\rangle \simeq |A(0, t)\rangle + |w(z, t)\rangle + |n(z, t)\rangle . \quad (1.60)$$

The ASE noise is typically modeled as an additive white Gaussian noise (AWGN).

The ASE variance at the end of a homogeneous link composed of N identical amplifiers, in the absence of Raman scattering, is equal to [12]

$$\sigma_{\text{ASE}}^2 = N \times h\nu FGB \quad (1.61)$$

where h is the Planck's constant, ν is the channel central frequency, G is the amplifier gain, here assumed frequency-independent, and B is receiver noise equivalent bandwidth. The quantity F is called the amplifier noise figure and quantifies the degradation of the system performance due to the presence of ASE. The noise figure can be related to the EDFA spontaneous emission factor n_{sp} by $F \triangleq 2n_{\text{sp}}(G - 1)/G$ [12].

The stochastic differential equation (SDE) in Eq. (6.9) can be expressed in integral form as:

$$\mathbf{U}(z) = \mathbf{U}(0) - \frac{j\omega\mu}{2N_m} \int_0^z \mathbf{U}(\zeta) \circ d\mathbf{W}(\zeta) \quad (1.62)$$

where the symbol \circ indicates the Stratonovich integral [31]. Other types of optical amplifiers based on different amplification mechanisms exist. For instance, due to their wider amplification bandwidth, semiconductor optical amplifiers (SOAs) represent an interesting, although not yet deployed, alternative to rare-earth-doped fiber amplifiers. These amplifiers will be discussed later in Chapter 4. Alternatively, optical amplification can be performed also in a distributed manner, opposite to lumped amplification, exploiting the SRS process taking place in the optical fiber. The amplification thus occurs via the transfer of power from a pump to the useful signal, in a distributed manner along the fiber length. The pump, which gets depleted by the SRS process, is transmitted at a higher frequency with respect to the useful signal for efficient SRS to take place.

In addition to optical amplification to counteract fiber losses, the other linear impairments must be addressed in the link design. A key aspect is the management of chromatic dispersion, to counteract the GVD-induced pulse broadening. The compensation of the chromatic dispersion in the optical domain is performed by means of dispersion compensating fibers (DCFs). These fibers can be designed such that the dispersion accumulated at their output compensates for the dispersion accumulated

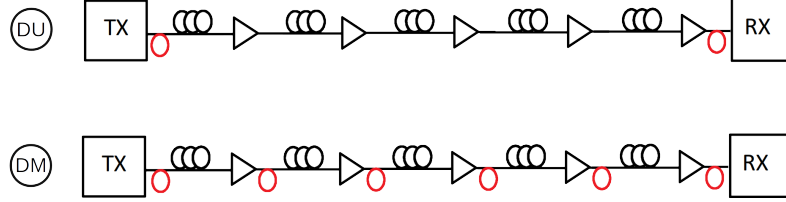


Figure 1.10: Sketch of a dispersion uncompensated (DU), and dispersion managed (DM) link. The red fibers are DCFs

up to their input $D_{\text{cum},\text{in}}$, namely:

$$D_{\text{cum},\text{in}} + D_{\text{DCF}}L_{\text{DCF}} = 0 \quad (1.63)$$

with L_{DCF} the DCF length. Dispersion compensation strategies differ from the position of the DCFs along the link. In particular, modern optical communication links are designed with a unique DCF at the link-end (or a combination of link-start/end) that compensates for the dispersion accumulated along the whole link. This type of link is called dispersion-uncompensated (DU) link.

On the other hand, the dispersion compensation can be performed on a per span basis. In this case, a DCF is placed after each span-end. These links are called dispersion-managed (DM) links. In DM links some residual dispersion per span may be used to mitigate nonlinear effects and to relax its accumulation along the distance. A sketch is reported in Fig. 1.10 for a 5 span link.

In both DM and DU the residual dispersion accumulated during the entire propagation is removed at the transmitter/receiver side, either by a DCF or by electronic dispersion compensation. The last option is currently preferred thanks to fast digital signal processing since it can be easily tuned and is lossless, contrary to DCF.

1.7 Quality of transmission metrics

In modern receivers, coherent detection is applied to retrieve the transmitted sequence on the in-phase and quadrature components of both polarization tributaries or modes.

In such receivers, after analog to digital conversion of the received signal, digital signal processing (DSP) is applied. At this stage, static and adaptive channel equalization can be applied, followed by carrier frequency and phase estimation. Finally, the samples are fed to the detector which provides the estimated sequence of transmitted symbols which is then demapped into a sequence of bits.

The estimation of the quality of transmission (QoT) can be performed based on different metrics. One of the most widely adopted metrics is the bit error rate (BER). Such a metric is typically evaluated before the forward error correction decoder (pre-FEC) and can be mapped one-to-one to the post-FEC BER when hard-decision decoding is used. However, the advent of powerful soft-decision FEC codes, such as turbo codes, made the estimation of the actual performance based on pre-FEC BER less accurate. For this reason, quality of transmission metrics related to the information theory [32] have become popular in recent years [33].

1.7.1 Signal-to-noise ratio

By definition, the signal-to-noise ratio (SNR) is the ratio between the signal and noise power. In long-haul optical transmission systems, the main impairments setting the SNR are ASE noise and Kerr effects. According to the regular perturbation model of the optical fiber, the fiber nonlinear effects are described as an NLI statistically independent of ASE, hence the SNR of a generic WDM channel can be written as

$$\text{SNR} = \frac{P}{\sigma_{\text{NLI}}^2 + \sigma_{\text{ASE}}^2} \quad (1.64)$$

with P the channel power, σ_{ASE}^2 the ASE noise variance in Eq. (1.61), and σ_{NLI}^2 the NLI variance. For fast system performance estimation, it is mandatory to rely on fast analytical expression for the latter variance. Chapter 2 of this thesis will be devoted to reviewing in the detail the theory behind the estimation of σ_{NLI}^2 .

Since the Kerr effect is cubic in the field, the NLI variance scales cubically with the channel power [34, 35] namely:

$$\sigma_{\text{NLI}}^2 = a_{\text{NL}} P^3 \quad (1.65)$$

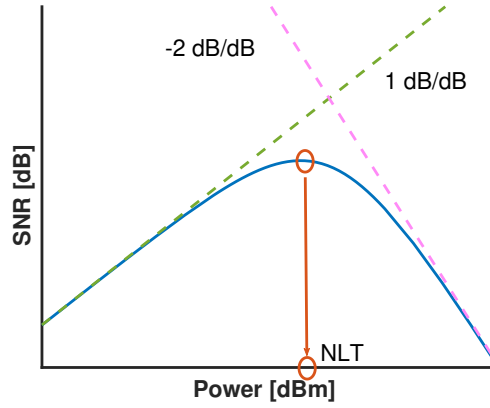


Figure 1.11: Example of the typical SNR bell-curve vs. channel power.

with a_{NL} a coefficient in $[\text{W}^{-2}]$. As a consequence, the SNR in Eq. (1.64) has two asymptotic behavior versus signal power. At a small power, where the nonlinear effects are negligible, the main impairment is given by ASE noise, namely $\text{SNR} \approx \text{SNR}_{\text{LIN}}$ with $\text{SNR}_{\text{LIN}} \triangleq \frac{P}{\sigma_{\text{ASE}}^2}$. This operation regime is hence called the linear regime. In the linear regime, an increase of 1 dB in the signal power yields 1 dB of SNR increase. On the other hand, as the signal power increases the nonlinear effects become non-negligible and eventually outweigh the ASE noise. In the nonlinear regime, i.e., at high signal power, the SNR can be approximated as $\text{SNR} \approx \text{SNR}_{\text{NL}}$, with $\text{SNR}_{\text{NL}} \triangleq \frac{1}{a_{\text{NL}} P^2}$ thus exhibiting 2 dB of decrease for 1 dB of signal power increase.

As a consequence of the two operational regimes, the SNR curve versus signal power, in a dB scale, assumes a typical bell shape. An example is sketched in Fig. 1.11, in which the slopes of the asymptotic regimes are highlighted. Due to its concave shape, the SNR curve (or the *bell-curve*) has an optimum, which is associated with a best power value [9]:

$$P_{\text{NLT}} = \left(\frac{\sigma_{\text{ASE}}^2}{2a_{\text{NL}}} \right)^{1/3} \quad (1.66)$$

and it is usually called *nonlinear threshold* (NLT) [36]. At such optimal power, the

ASE noise variance is two times the NLI variance.

The SNR in Eq. (1.64) can be found by concatenation of the inverse linear and nonlinear SNR, namely, the following relationship holds:

$$\text{SNR} = \left(\frac{1}{\text{SNR}_{\text{LIN}}} + \frac{1}{\text{SNR}_{\text{NL}}} \right)^{-1} \quad (1.67)$$

where SNR_{LIN} and SNR_{NL} are computed end-to-end. Alternatively, adopting an *incoherent* point of view, which neglects the correlation among the different spans, the end-to-end SNR can be computed as a concatenation of span-by-span metrics, as follows:

$$\text{SNR} = \left(\sum_{i=1}^N \frac{1}{\text{SNR}_{\text{LIN},i}} + \sum_{i=1}^N \frac{1}{\text{SNR}_{\text{NL},i}} \right)^{-1} \quad (1.68)$$

with N the number of spans in the link. The topic of incoherent and coherent accumulation will be treated in more details in Chapter 2 and 3.

1.7.2 Achievable information rate

The achievable information rate (AIR) provides an estimation of the bit rate that can be achieved in the presence of soft-decision decoding. In the most general case, the AIR can be estimated by evaluating the mutual information [32] through mismatched-decoding [9]. Such a metric represents a lower bound on the *channel capacity* [37], which sets the fundamental limit on the number of information bits that can be reliably transferred in the channel. In the simplest case of an AWGN channel, such a channel capacity is expressed by the elegant Shannon formula [37]:

$$C = \log_2(1 + \text{SNR}) \quad (1.69)$$

which is a function of the SNR and it is achieved by a Gaussian distribution of the transmitted symbols. Other distributions of the transmitted symbols yield lower mutual information, as depicted in Fig. 1.12 for M -QAM modulation formats with uniformly distributed symbols. This gap-to-capacity can be reduced employing of *probabilistic shaping* techniques. This technique consists in drawing the transmitted

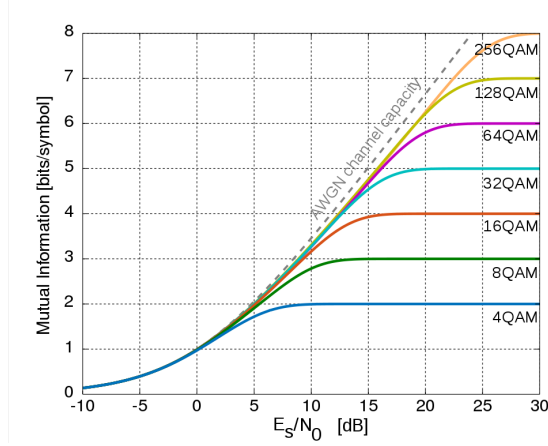


Figure 1.12: Mutual information vs. signal-to-noise ratio for uniformly distributed M -QAM modulation formats. Dashed line: Shannon formula of the AWGN channel capacity, achieved by Gaussian distributed symbols.

symbols from their alphabet with a Maxwell-Boltzmann distribution [38], instead of uniformly, hence resembling a Gaussian distribution.

Chapter 2

Theoretical models for the nonlinear interference variance

The regular perturbation analyzed in Section 1.5.1 approximates the distortion of the Kerr effect by an additive interference to the signal. As discussed in Sec 1.7.1, the variance of such NLI is a key ingredient for the computation of the SNR and thus the design and the analysis of optical links. The NLI variance can be estimated by a numerical implementation of the SSFM algorithm, at the price of observing many symbols to smooth the estimation error. On the other hand, the RP1 method opened the door to analytical models, such as the Gaussian noise (GN) model [3, 39] and the enhanced Gaussian noise (EGN) model [40, 41], also known as the nonlinear interference noise (NLIN) model [42]. In this chapter we introduce such models with a detailed description of the supporting proofs. Parts of the material presented in this chapter were published in [30, 43].

2.1 Nonlinear interference

The nonlinear interference at the receiver, after the compensation of the linear effects accumulated during propagation, is defined consistently with Eq. (1.53) as

$$|n(z, t)\rangle \triangleq \int_0^z e^{-\mathcal{L}\zeta} \mathcal{N}\left(e^{\mathcal{L}\zeta} |A(0, t)\rangle\right) d\zeta \quad (2.1)$$

with \mathcal{L} and \mathcal{N} the linear and nonlinear operators, respectively. The NLI is additive along distance, and its generic contribution generated at coordinate ζ depends on the unperturbed signal up to that coordinate and experiences only linear effects up to the end [41]. In the frequency domain, the NLI $|\tilde{n}(\omega)\rangle$ is given by the Fourier transform of Eq. (2.1):

$$\begin{aligned} |\tilde{n}(\omega)\rangle = & -j\frac{8}{9}\gamma \iint_{-\infty}^{\infty} \eta(\omega, \omega_1, \omega_2) \\ & \times \langle \tilde{A}(0, \omega + \omega_1 + \omega_2) | \tilde{A}(0, \omega + \omega_2) \rangle | \tilde{A}(0, \omega + \omega_1) \rangle \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} \end{aligned} \quad (2.2)$$

where $|\tilde{A}(0, \omega)\rangle$ is the Fourier transform of the transmitted field, while η is the link kernel weighting the inner FWM process. We will discuss the kernel in detail in the next section. Exploiting the atom notation presented in Chapter 1, such a Fourier transform can be written as

$$|\tilde{A}(0, \omega)\rangle = \sum_{\mathbf{n}} \frac{a_{\mathbf{n}}}{\sqrt{2}} |\tilde{G}_{\mathbf{n}}(0, \omega)\rangle \quad (2.3)$$

where the basis functions are defined as

$$|\tilde{G}_{\mathbf{n}}(0, \omega)\rangle \triangleq \tilde{G}_{n_1 n_2}(0, \omega) |n_3\rangle \quad (2.4)$$

with $\tilde{G}_{n_1 n_2}$ a scalar function equal to:

$$\tilde{G}_{n_1 n_2}(0, \omega) = \tilde{p}_{n_2}(\omega - \omega_{n_2}) e^{-j(\omega - \omega_{n_2})n_1 T_{n_2}}. \quad (2.5)$$

Equation (2.2) can thus be rewritten putting in evidence the data symbols as

$$\begin{aligned} |\tilde{n}(\omega)\rangle = & -j\frac{8}{9} \left(\frac{1}{\sqrt{2}}\right)^3 \gamma \sum_{\mathbf{h}, \mathbf{m}, \mathbf{n}} a_{\mathbf{h}}^* a_{\mathbf{m}} a_{\mathbf{n}} \iint_{-\infty}^{\infty} \eta(\omega, \omega_1, \omega_2) \\ & \times \langle \tilde{G}_{\mathbf{h}}(0, \omega + \omega_1 + \omega_2) | \tilde{G}_{\mathbf{m}}(0, \omega + \omega_2) \rangle | \tilde{G}_{\mathbf{n}}(0, \omega + \omega_1) \rangle \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi}. \end{aligned} \quad (2.6)$$

Such an NLI is received by the detector, which operates sampling, demodulation, and matched filtering to the received field. In our framework, the aforementioned operations correspond to the inner product $\int_{-\infty}^{\infty} \langle \tilde{G}_i(\omega) | \tilde{G}_n(\omega) \rangle \frac{d\omega}{2\pi}$.

As a consequence, neglecting ASE noise, we obtain the following equivalent discrete-time model for the fiber-optic channel:

$$u_i = a_i + n_i \quad (2.7)$$

which relates the transmitted atom a_i to the received atom u_i . The discrete-time NLI in Eq. (2.7) can be written as:

$$n_i = \sum_{\mathbf{h}, \mathbf{m}, \mathbf{n}} a_{\mathbf{h}}^* a_{\mathbf{m}} a_{\mathbf{n}} \mathcal{X}_{\mathbf{h}\mathbf{m}\mathbf{n}i}. \quad (2.8)$$

The term $\mathcal{X}_{\mathbf{h}\mathbf{m}\mathbf{n}i}$ is the function weighting the mixing at the symbol level, namely:

$$\begin{aligned} \mathcal{X}_{\mathbf{h}\mathbf{m}\mathbf{n}i} = & -j \frac{8}{9} \left(\frac{1}{\sqrt{2}} \right)^3 \gamma \iiint_{-\infty}^{\infty} \eta(\omega, \omega_1, \omega_2) \\ & \times \langle \tilde{G}_{\mathbf{h}}(0, \omega + \omega_1 + \omega_2) | \tilde{G}_{\mathbf{m}}(0, \omega + \omega_2) \rangle \\ & \times \langle \tilde{G}_{\mathbf{i}}(0, \omega) | \tilde{G}_{\mathbf{n}}(0, \omega + \omega_1) \rangle \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} \frac{d\omega}{2\pi} \end{aligned} \quad (2.9)$$

which, in the scalar case of single-polarization in a birefringent fiber, reduces to the expression:

$$\begin{aligned} \mathcal{S}_{\mathbf{h}\mathbf{m}\mathbf{n}i} = & -j \frac{8}{9} \gamma \iiint_{-\infty}^{\infty} \eta(\omega, \omega_1, \omega_2) \\ & \times \tilde{G}_{h_1 h_2}(0, \omega + \omega_1 + \omega_2) \tilde{G}_{m_1 m_2}(0, \omega + \omega_2) \\ & \times \tilde{G}_{i_1 i_2}(0, \omega) \tilde{G}_{n_1 n_2}(0, \omega + \omega_1) \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} \frac{d\omega}{2\pi} \end{aligned} \quad (2.10)$$

which benefits from useful symmetric properties such as

$$\mathcal{S}_{\mathbf{h}\mathbf{m}\mathbf{n}i} = \mathcal{S}_{\mathbf{m}\mathbf{h}\mathbf{n}i}^*, \quad \mathcal{S}_{\mathbf{h}\mathbf{m}\mathbf{n}i} = \mathcal{S}_{\mathbf{h}\mathbf{n}\mathbf{m}i}. \quad (2.11)$$

In particular, applying the basis function definition in Eq. (2.4), we note that the following relationship holds between the term $\mathcal{X}_{\mathbf{h}\mathbf{m}\mathbf{n}\mathbf{i}}$ and its scalar version:

$$\begin{aligned}\mathcal{X}_{\mathbf{h}\mathbf{m}\mathbf{n}\mathbf{i}} &= \left(\frac{1}{\sqrt{2}}\right)^3 \mathcal{S}_{\mathbf{h}\mathbf{m}\mathbf{n}\mathbf{i}} \langle h_3 | m_3 \rangle \langle i_3 | n_3 \rangle \\ &= \left(\frac{1}{\sqrt{2}}\right)^3 \mathcal{S}_{\mathbf{h}\mathbf{m}\mathbf{n}\mathbf{i}} \delta_{h_3 m_3} \delta_{i_3 n_3}\end{aligned}\quad (2.12)$$

with δ_{ij} a Kronecker's delta.

2.1.1 Link kernel

The term $\eta(\omega, \omega_1, \omega_2)$ in Eq. (2.2) is the so-called link kernel (also known as link function [40]) weighting the four-wave mixing interaction among signal frequencies. For a homogeneous dispersion-uncompensated link composed of N identical spans, each of length L , such a link kernel can be written as

$$\eta(\omega, \omega_1, \omega_2) = \sum_{k=1}^N e^{j\Delta\beta(k-1)L} \int_0^L e^{j\Delta\beta\zeta} e^{-\alpha\zeta} d\zeta \quad (2.13)$$

where the phase-matching coefficient $\Delta\beta$ is defined in Eq. (1.35) and can be expanded in terms of the fiber dispersion coefficients by:

$$\begin{aligned}\Delta\beta &\triangleq \beta(\omega) - \beta(\omega + \omega_1) - \beta(\omega + \omega_2) + \beta(\omega + \omega_1 + \omega_2) \\ &= \omega_1 \omega_2 \left[\beta_2 + \frac{1}{2}(2\omega + \omega_1 + \omega_2)\beta_3 \right].\end{aligned}\quad (2.14)$$

Note that here, contrary to Eq. (1.33), we are not stressing the carrier indexes since at this level we are viewing the entire WDM as a unique channel.

The integral in Eq. (2.13) weights the FWM in a generic span, while $e^{j\Delta\beta(k-1)L}$ accounts for the accumulated phase mismatch among the frequencies before entering the k th span. The link kernel can be written in the compact form:

$$\eta(\omega, \omega_1, \omega_2) = \chi(\omega, \omega_1, \omega_2) \eta_1(\omega, \omega_1, \omega_2) \quad (2.15)$$

where η_1 is the single-span kernel, which accounts for the NLI generated by a span. In this framework, the z -integral can be expressed in a closed-form formula through

$$\eta_1 \triangleq \int_0^L e^{-\alpha\zeta} e^{j\Delta\beta\zeta} d\zeta = \frac{1 - e^{-\alpha L} e^{j\Delta\beta L}}{\alpha - j\Delta\beta}. \quad (2.16)$$

Even the summation in Eq. (2.13) can be closed yielding the so-called *phased-array* term χ [44]:

$$\chi \triangleq \sum_{k=1}^N e^{j\Delta\beta(k-1)L} = \frac{1 - e^{jN\Delta\beta L}}{1 - e^{j\Delta\beta L}} \quad (2.17)$$

which accounts for the NLI accumulation span-by-span for identical spans [3].

2.1.2 Nonlinear interference variance

The evaluation of the NLI variance on a generic (time, frequency, mode) channel use requires the calculation of the following moment:

$$\mathbb{E}[n_i n_i^*] = \sum_{\mathbf{h}, \mathbf{m}, \mathbf{n}} \sum_{\mathbf{l}, \mathbf{j}, \mathbf{o}} \mathbb{E}[a_{\mathbf{h}}^* a_{\mathbf{m}} a_{\mathbf{n}} a_{\mathbf{l}}^* a_{\mathbf{j}}^* a_{\mathbf{o}}] \mathcal{X}_{\mathbf{h}\mathbf{m}\mathbf{n}\mathbf{i}} \mathcal{X}_{\mathbf{l}\mathbf{j}\mathbf{o}\mathbf{i}}^*. \quad (2.18)$$

Assuming all the $a_{\mathbf{h}}$ to be complex zero-mean independent random variables with n -fold rotational symmetry and $n \geq 4$, only combinations with an equal number of conjugate/non-conjugate pairs are non-zero. Therefore, the only terms surviving in Eq. (2.18) are those in the form [45]: $\mathbb{E}[|a_{\mathbf{h}}|^2 |a_{\mathbf{m}}|^2 |a_{\mathbf{n}}|^2]$, $\mathbb{E}[|a_{\mathbf{h}}|^4 |a_{\mathbf{m}}|^2]$, and $\mathbb{E}[|a_{\mathbf{h}}|^6]$. Exploiting the cumulant theory as in [41], these moments can be expressed in the compact notation

$$\mathbb{E}[|a_{\mathbf{h}}|^2 |a_{\mathbf{m}}|^2 |a_{\mathbf{n}}|^2] = \kappa_{1;1}^{(\mathbf{h})} \kappa_{1;1}^{(\mathbf{m})} \kappa_{1;1}^{(\mathbf{n})} \quad (2.19)$$

$$\mathbb{E}[|a_{\mathbf{h}}|^4 |a_{\mathbf{m}}|^2] = \kappa_{2;2}^{(\mathbf{h})} \kappa_{1;1}^{(\mathbf{m})} \quad (2.20)$$

$$\mathbb{E}[|a_{\mathbf{h}}|^6] = \kappa_{3;3}^{(\mathbf{h})} \quad (2.21)$$

where the cumulants $\kappa_{1;1}$, $\kappa_{2;2}$ and $\kappa_{3;3}$ are defined as:

$$\kappa_{1;1} = \mathbb{E} [|a|^2] \quad (2.22)$$

$$\kappa_{2;2} = \mathbb{E} [|a|^2]^2 \left(\frac{\mathbb{E} [|a|^4]}{\mathbb{E} [|a|^2]^2} - 2 \right) \triangleq \mathbb{E} [|a|^2]^2 \Phi \quad (2.23)$$

$$\kappa_{3;3} = \mathbb{E} [|a|^2]^3 \left(\frac{\mathbb{E} [|a|^6]}{\mathbb{E} [|a|^2]^3} - 9 \frac{\mathbb{E} [|a|^4]}{\mathbb{E} [|a|^2]^2} + 12 \right) \triangleq \mathbb{E} [|a|^2]^3 \Psi \quad (2.24)$$

where we put in evidence the average power and implicitly defined the factors Φ and Ψ as typically done in the literature [40]. The term in Eq. (2.19) involves second-order moments of the transmitted symbols and hence weights the so-called second-order noise (SON). Such a second-order noise terms involves three non-degenerate pairs of conjugate/non-conjugate symbols. Among these six possible indexes combinations, only two configurations yields a valid pair's triplet, namely:

$$\text{SON} = \begin{cases} \mathbf{h} = \mathbf{l}, \mathbf{m} = \mathbf{j}, \mathbf{n} = \mathbf{o} & \rightarrow \mathcal{X}_{\mathbf{h}\mathbf{m}\mathbf{n}\mathbf{i}} \mathcal{X}_{\mathbf{h}\mathbf{m}\mathbf{n}\mathbf{i}}^* \\ \mathbf{h} = \mathbf{l}, \mathbf{m} = \mathbf{o}, \mathbf{n} = \mathbf{j} & \rightarrow \mathcal{X}_{\mathbf{h}\mathbf{m}\mathbf{n}\mathbf{i}} \mathcal{X}_{\mathbf{h}\mathbf{m}\mathbf{n}\mathbf{i}}^* \end{cases} \quad (2.25)$$

It is worth noting that, in the scalar case, the two contributions in Eq. (2.25) are degenerate since $\mathcal{X}_{\mathbf{h}\mathbf{m}\mathbf{n}\mathbf{i}} \rightarrow \mathcal{S}_{\mathbf{h}\mathbf{m}\mathbf{n}\mathbf{i}} = \mathcal{S}_{\mathbf{h}\mathbf{m}\mathbf{n}\mathbf{i}}$.

The four discarded combinations in the SON term are those where a pair conjugate/ non-conjugate yielding a modulus square belongs to the same time-discrete NLI in Eq. (2.18), either n_i or n_i^* . These terms can be safely neglected as their contribution is an average phase rotation that can be removed by a carrier phase estimator (CPE).

The term described by Eq. (2.20) weights the fourth-order noise (FON) contribution to NLI. It involves a quadruplet of conjugate/non-conjugate symbols and one pair, associated with a different index. There are nine possible configurations of indexes for this scenario. Four of these combinations will be discarded since they are removed by the CPE. Among the remaining configurations, we identify two different contributions, called the F4 and Q4 term following the notation of [41], as follows:

$$\text{FON} = \begin{cases} \mathbf{h} = \mathbf{m} = \mathbf{l} = \mathbf{j}, \mathbf{n} = \mathbf{o} & \rightarrow \mathcal{X}_{\mathbf{hhni}} \mathcal{X}_{\mathbf{hhni}}^* & \text{F4} \\ \mathbf{h} = \mathbf{m} = \mathbf{l} = \mathbf{o}, \mathbf{n} = \mathbf{j} & \rightarrow \mathcal{X}_{\mathbf{hhni}} \mathcal{X}_{\mathbf{hnhi}}^* & \text{F4} \\ \mathbf{h} = \mathbf{n} = \mathbf{l} = \mathbf{j}, \mathbf{m} = \mathbf{o} & \rightarrow \mathcal{X}_{\mathbf{hmhi}} \mathcal{X}_{\mathbf{hhmi}}^* & \text{F4} \\ \mathbf{h} = \mathbf{n} = \mathbf{l} = \mathbf{o}, \mathbf{m} = \mathbf{j} & \rightarrow \mathcal{X}_{\mathbf{hmhi}} \mathcal{X}_{\mathbf{hmhi}}^* & \text{F4} \\ \mathbf{m} = \mathbf{n} = \mathbf{j} = \mathbf{o}, \mathbf{h} = \mathbf{l} & \rightarrow \mathcal{X}_{\mathbf{hnni}} \mathcal{X}_{\mathbf{hnni}}^* & \text{Q4.} \end{cases} \quad (2.26)$$

We thus have four F4 contributions and one Q4 combination. This discrimination F4/Q4 is useful since in Section 2.3 we will show that the F4 terms have all the same variance, which is different from the Q4 term.

Finally, the moment in Eq. (2.21) acts as a weight for the so-called higher-order noise (HON) term, which involves sixth-order moments of the symbols. We call such a contribution the Q6 term, which is associated with all-degenerates indexes in Eq. (2.18), namely:

$$\text{HON} : \quad \mathbf{h} = \mathbf{m} = \mathbf{n} = \mathbf{l} = \mathbf{j} = \mathbf{o} \rightarrow \mathcal{X}_{\mathbf{hhhi}} \mathcal{X}_{\mathbf{hhhi}}^* \quad \text{Q6.} \quad (2.27)$$

Combining now the pieces from Eqs.(2.25)-(2.27), the NLI variance in Eq. (2.18) can be written as [43]:

$$\begin{aligned} \mathbb{E}[n_{\mathbf{i}} n_{\mathbf{i}}^*] &= \sum_{\mathbf{h}, \mathbf{m}, \mathbf{n}} \kappa_{1;1}^{(\mathbf{h})} \kappa_{1;1}^{(\mathbf{n})} \kappa_{1;1}^{(\mathbf{m})} \mathcal{X}_{\mathbf{hmni}} (\mathcal{X}_{\mathbf{hmni}}^* + \mathcal{X}_{\mathbf{hnmi}}^*) & (\text{SON}) \\ &+ \sum_{\mathbf{h}, \mathbf{n}} \kappa_{2;2}^{(\mathbf{h})} \kappa_{1;1}^{(\mathbf{n})} \left(|\mathcal{X}_{\mathbf{hhni}} + \mathcal{X}_{\mathbf{hnhi}}|^2 + |\mathcal{X}_{\mathbf{hnhi}}|^2 \right) & (\text{F4} + \text{Q4}) \\ &+ \sum_{\mathbf{h}} \kappa_{3;3}^{(\mathbf{h})} |\mathcal{X}_{\mathbf{hhhi}}|^2 & (\text{Q6}) \end{aligned} \quad (2.28)$$

which represents a generalization of the master theorem in [41].

In the literature, several efforts have been made to provide analytical expressions for the NLI contributions in Eq. (2.28). In the remainder of this chapter, we will review the theoretical derivation of these expressions consistently with the notation used in this work.

2.2 The Gaussian noise model

Among all the theoretical models proposed in the literature, thanks to its simplicity, the GN model [3] has become the most popular analytical model for the NLI variance estimation. The model's simplicity comes from the key assumption that the signal is a stochastic process with Gaussian statistics. Such an assumption is well satisfied, e.g., by probabilistic shaped modulation formats, or by any signal after accumulating a sufficient dispersion, as in typical dispersion-uncompensated links. Assuming Gaussian distributed symbols yields $\kappa_{2,2} = 0$ and $\kappa_{3,3} = 0$ thus nullifying the FON and HON contributions and significantly reducing the complexity of Eq. (2.18). For this reason, the SON contribution is often called the GN term in the literature, being the only non-zero contribution in a Gaussian framework.

In the absence of polarization dependent effects, such as polarization-dependent loss or polarization mode dispersion, the SON/GN contribution to the NLI variance on polarization $i_3 \in (x, y)$ can be written as:

$$\begin{aligned} \sigma_{\text{GN}i_3}^2 &= \kappa_{1,1}^3 \sum_{\mathbf{h}, \mathbf{m}, \mathbf{n}} \mathcal{X}_{\mathbf{h}\mathbf{m}\mathbf{n}i} (\mathcal{X}_{\mathbf{h}\mathbf{m}\mathbf{n}i}^* + \mathcal{X}_{\mathbf{h}\mathbf{m}\mathbf{n}i}^*) \\ &= \frac{1}{8} \kappa_{1,1}^3 \sum_{h_1, m_1, n_1} \sum_{h_2, m_2, n_2} \left(\sum_{h_3, m_3, n_3} \delta_{h_3 m_3} \delta_{i_3 n_3} + \sum_{h_3, m_3, n_3} \delta_{h_3 m_3} \delta_{i_3 n_3} \delta_{h_3 n_3} \delta_{i_3 m_3} \right) |\mathcal{S}_{\mathbf{h}\mathbf{m}\mathbf{n}i}|^2 \end{aligned} \quad (2.29)$$

where we simplified the notation exploiting $\kappa_{1,1}^{(\mathbf{h})} \kappa_{1,1}^{(\mathbf{n})} \kappa_{1,1}^{(\mathbf{m})} = \kappa_{1,1}^3$, and we expressed the variance in terms of the function $\mathcal{S}_{\mathbf{h}\mathbf{m}\mathbf{n}i}$ weighting the FWM process for a single-polarization propagation as reported in Eq. (2.10). This allows to simplify the overall expression by exploiting the symmetries of $\mathcal{S}_{\mathbf{h}\mathbf{m}\mathbf{n}i}$. The verbatim application of Eq. (2.10) yields the combination of Kronecker's delta in Eq. (2.29). Consistently with the derivation in [41], such weight is equal to 3. Combining this observation with the definition in Eq. (2.10), the variance can be expressed as:

$$\begin{aligned}
\sigma_{\text{GN}i_3}^2 &= 3 \frac{1}{8} \kappa_{1;1}^3 \sum_{h_1, m_1, n_1} \sum_{h_2, m_2, n_2} |\mathcal{S}_{\text{hmni}}|^2 \\
&= 3 \left(\frac{8}{9} \right)^2 \frac{1}{8} \gamma^2 \kappa_{1;1}^3 \sum_{\substack{h_1, m_1, n_1 \\ h_2, m_2, n_2}} \iiint \iiint_{-\infty}^{\infty} \eta(\omega, \omega_1, \omega_2) \eta^*(\mu, \mu_1, \mu_2) \\
&\quad \times \tilde{G}_{h_1 h_2}^*(0, \omega + \omega_1 + \omega_2) \tilde{G}_{m_1 m_2}(0, \omega + \omega_2) \tilde{G}_{i_1 i_2}^*(0, \omega) \\
&\quad \times \tilde{G}_{n_1 n_2}(0, \omega + \omega_1) \tilde{G}_{h_1 h_2}(0, \mu + \mu_1 + \mu_2) \tilde{G}_{m_1 m_2}^*(0, \mu + \mu_2) \\
&\quad \times \tilde{G}_{i_1 i_2}(0, \mu) \tilde{G}_{n_1 n_2}^*(0, \mu + \mu_1) \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} \frac{d\omega}{2\pi} \frac{d\mu_1}{2\pi} \frac{d\mu_2}{2\pi} \frac{d\mu}{2\pi}.
\end{aligned} \tag{2.30}$$

Although selecting only the valid index combination as per Eq. (2.25) reduces the number of summations, the final result in Eq. (2.30) still depends on an infinite summation over the discrete-time index. However, in the special case of sinc pulses, such a summation can be dropped, with significant simplifications, as already observed in [21, 46], thanks to the Poisson summation formula:

$$\sum_{k=-\infty}^{\infty} e^{jk\omega T} = \frac{2\pi}{T} \sum_{k=-\infty}^{\infty} \delta\left(\omega - \frac{2\pi k}{T}\right) \tag{2.31}$$

where $\delta(\omega)$ indicates a Dirac's delta, while the exponential functions are within the definition of the basis functions as per Eq. (2.5). For the sake of clearness, we now explicitly substitute Eq. (2.5) in Eq. (2.30), namely:

$$\begin{aligned}
\sigma_{\text{GN}i_3}^2 &= 3 \left(\frac{8}{81} \right) \gamma^2 \kappa_{1;1}^3 \sum_{\substack{h_1, m_1, n_1 \\ h_2, m_2, n_2}} \iiint \iiint_{-\infty}^{\infty} \eta(\omega, \omega_1, \omega_2) \eta^*(\mu, \mu_1, \mu_2) \\
&\quad \times \tilde{p}_{h_2}^*(\omega + \omega_1 + \omega_2 - \omega_{h_2}) \tilde{p}_{m_2}(\omega + \omega_2 - \omega_{m_2}) \tilde{p}_{i_2}^*(\omega - \omega_{i_2}) \\
&\quad \times \tilde{p}_{n_2}(\omega + \omega_1 - \omega_{n_2}) \tilde{p}_{h_2}(\mu + \mu_1 + \mu_2 - \omega_{h_2}) \tilde{p}_{m_2}^*(\mu + \mu_2 - \omega_{m_2}) \\
&\quad \times \tilde{p}_{i_2}(\mu - \omega_{i_2}) \tilde{p}_{n_2}^*(\mu + \mu_1 - \omega_{n_2}) \\
&\quad \times e^{j(\omega + \omega_1 + \omega_2 - \mu - \mu_1 - \mu_2)h_1 T_{h_2}} e^{j(\mu + \mu_2 - \omega - \omega_2)m_1 T_{m_2}} \\
&\quad \times e^{j(\omega - \mu)i_1 T_{i_2}} e^{j(\mu + \mu_1 - \omega - \omega_1)n_1 T_{n_2}} \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} \frac{d\omega}{2\pi} \frac{d\mu_1}{2\pi} \frac{d\mu_2}{2\pi} \frac{d\mu}{2\pi}.
\end{aligned} \tag{2.32}$$

Thanks to the Poisson formula, all summations in Eq. (2.30) over temporal indexes can be dropped, imposing to zero the frequency combination at the exponent. This coincides with imposing equality between the frequency combination appearing as the argument of basis functions in Eq. (2.30) having the same index. Equating the arguments yields the following linear systems:

$$\begin{cases} \mathbf{h} & \rightarrow \omega + \omega_1 + \omega_2 = \mu_1 + \mu_2 + \mu \\ \mathbf{m} & \rightarrow \omega + \omega_2 = \mu + \mu_2 \\ \mathbf{n} & \rightarrow \omega + \omega_1 = \mu + \mu_1 \end{cases} \quad (2.33)$$

whose solution is $\omega_1 = \mu_1$, $\omega_2 = \mu_2$, $\omega = \mu$. We can thus drop three integrals, for instance the ones with (μ, μ_1, μ_2) , by using the previous substitution. As a consequence, the NLI variance of the SON/GN contribution can be expressed in the following simplified integral:

$$\begin{aligned} \sigma_{\text{GNi}_3}^2 &= 3 \left(\frac{8}{81} \right) \frac{1}{T^3} \gamma^2 \kappa_{1,1}^3 \sum_{h,m,n} \iiint_{-\infty}^{\infty} |\eta(\omega, \omega_1, \omega_2)|^2 \\ &\quad \times |\tilde{p}_h(\omega + \omega_1 + \omega_2 - \omega_h)|^2 |\tilde{p}_m(\omega + \omega_2 - \omega_m)|^2 \\ &\quad \times |\tilde{p}_i(\omega - \omega_i)|^2 |\tilde{p}_n(\omega + \omega_1 - \omega_n)|^2 \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} \frac{d\omega}{2\pi}. \end{aligned} \quad (2.34)$$

Thanks to the performed simplifications, only the summations over channel indexes h_2, m_2 , and n_2 survive. For this reason, in Eq. (2.34) we dropped the subscript 2 in favor of a simpler notation $h_2, m_2, n_2 \rightarrow h, m, n$.

The variance in Eq. (2.34) is computed on one polarization. The result can be extended to the total variance $(x+y)$. In the absence of polarization effects, the $(x+y)$ variance of the GN term can be simply found as [41]:

$$\sigma_{\text{GN}}^2 = 2\sigma_{\text{GNi}_3}^2 \quad (2.35)$$

while different expressions hold with extra polarization effects not accounted by the Manakov equation, e.g., see Chapter 5 and 6. For the remainder of this work, all

the variances expressed without putting in evidence the polarization index, e.g., σ_{GN}^2 , must be read as $(x + y)$ variances.

It is worth mentioning that, in some cases, it is important to have access to the correlation between the NLI generated at different spans. In such cases, it is more convenient to express the link kernel in Eq. (2.34) in terms of summation of contributions, instead of the compact phased-array notation. Namely:

$$\sigma_{\text{GN}i_3}^2 = 3 \sum_{k,s=1}^N \mathcal{C}_{\text{GN}}(k, s) \quad (2.36)$$

$$\begin{aligned} \mathcal{C}_{\text{GN}}(k, s) \triangleq & \left(\frac{8}{81} \right) \frac{1}{T^3} \gamma^2 \kappa_{1;1}^3 \sum_{h,m,n} \iiint_{-\infty}^{\infty} \eta^{(k)}(\omega, \omega_1, \omega_2) \left(\eta^{(s)}(\omega, \omega_1, \omega_2) \right)^* \\ & \times |\tilde{p}_h(\omega + \omega_1 + \omega_2 - \omega_h)|^2 |\tilde{p}_m(\omega + \omega_2 - \omega_m)|^2 \\ & \times |\tilde{p}_i(\omega - \omega_i)|^2 |\tilde{p}_n(\omega + \omega_1 - \omega_n)|^2 \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} \frac{d\omega}{2\pi} \end{aligned} \quad (2.37)$$

where $\eta^{(k)}$ is the fiber kernel in span k , and $\mathcal{C}_{\text{GN}}(k, s)$ is the cross-correlation between span k and s . Such a notation will be particularly useful in the next chapters.

Although the NLI variance expression in Eq. (2.34) yields significant simplifications with respect to the starting expression in Eq. (2.30), it still involves the computation of three summations over the channel indexes h, m, n and three frequency integrals. The computation of the NLI variance can be further simplified by i) exploiting the physical insight on the valid FWM processes thus limiting the combinations of channel indexes and limiting the integration regions, and ii) introducing approximations to calculate the integrals in closed-form expressions. The following two sections are devoted to reviewing these topics.

2.2.1 Physical insight

The nonlinear Kerr effect is generally classified with different names depending on the involved channels, as reviewed in Chapter 1. Historically, the classification was

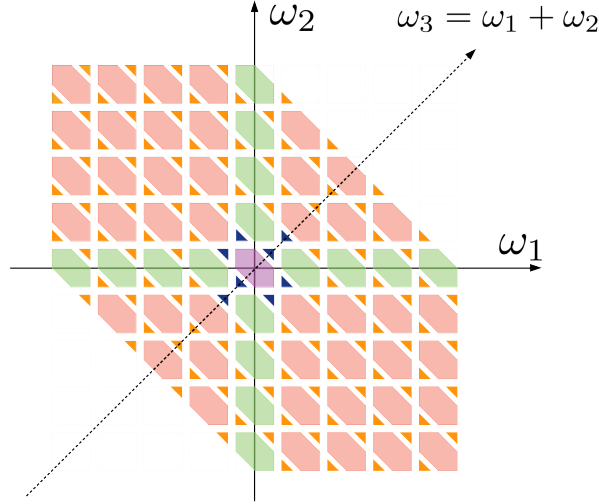


Figure 2.1: Frequencies ω_1 , ω_2 involved in the Kerr effect at $\omega = 0$, where the fourth frequency of the FWM-related process is $\omega_3 = \omega_1 + \omega_2$ for the law of energy conservation.

conceived for quasi constant-wave (CW) WDM channels spaced far apart, thus focusing only on the carriers involved in the process [7]. Referring to Eq. (2.34) where the indexes h, m, n label the WDM channels, in the standard classification the contribution in the nonlinear term summation with indexes $h = m = n = i$ corresponds to SPM, those with indexes $(h = m) \neq (n = i)$ to XPM, and all remaining contributions to FWM [7].

However, in modern transmissions with large bandwidth efficiency, it was first showed in [3] that channels may interact through the Kerr effect even if their carrier frequencies do not satisfy the law of conservation of energy. In this framework, it is more appropriate to interpret the Kerr nonlinearity as an FWM among the Fourier frequencies of the WDM spectrum satisfying $\omega = \omega_1 + \omega_2 - \omega_3$. For the sake of example, if the frequency under test is $\omega = 0$, the remaining frequencies are constrained by $\omega_3 = \omega_1 + \omega_2$. As a consequence, only sub-regions of the plane (ω_1, ω_2) will play a role in the NLI variance.

Fig. 2.1 sketches an example of the frequency domains in the (ω_1, ω_2) plane

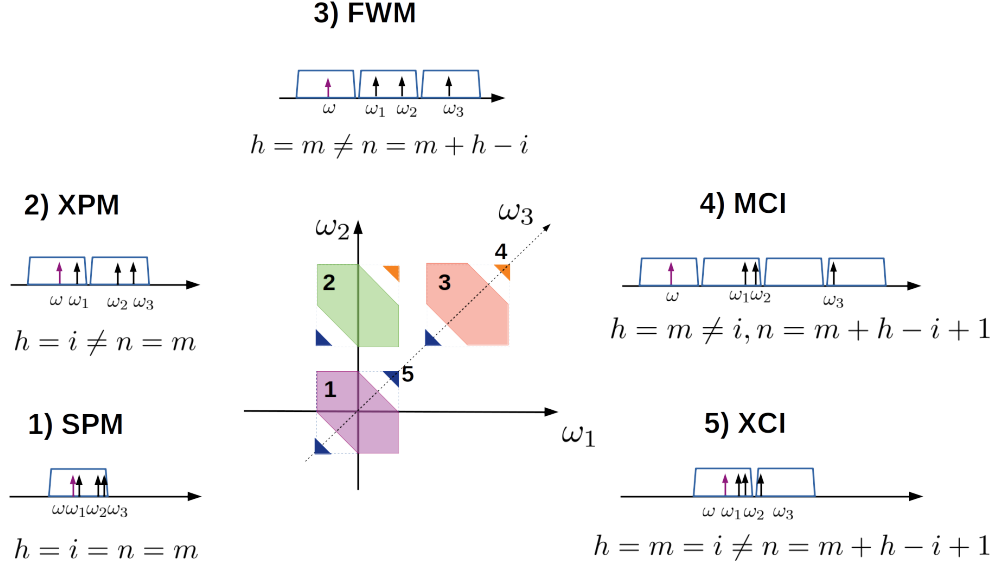


Figure 2.2: Examples of NLI classification, referred to a sub-region of Fig. 2.1.

involved in such FWM, for a 9 channel transmission. The diagonal dotted axis represents the constrained frequency $\omega_3 = \omega_1 + \omega_2$. We refer without loss of generality to a single polarization since the additional contributions coming from cross-polarization interactions can be treated similarly. In Fig. 2.1 we note that the traditional SPM, XPM, and FWM terms, which would correspond in the quasi-CW regime to dots on a regular grid having the carrier frequency spacing, show-up now as hexagons (inscribed in a square of edge equal to the channel bandwidth B) centered at grid dots. In this thesis, we will keep calling the contributions of such hexagons as SPM (purple), XPM (green), and classic FWM (red).

However, when the channel spacing is smaller than $3/2$ the channel bandwidth, new important triangular domains appear in addition to classical hexagon regions, as first noticed in [3]. Such domains correspond to channel indexes combinations governed by the relation $n = h + m - i \pm 1$.

It is worth noting that in [3, 40] all contributions involving two channels were called cross-channel interference (XCI), while all the contributions involving three

or four channels were called multi-channel interference (MCI). We will call the contributions of the blue triangles in Fig. 2.1 as XCI (this is consistent with the XCI naming in [40]), and those of the orange triangles as MCI as better specified in Fig. 2.2, where we highlight by examples the position of the four spectral frequencies generating the Kerr nonlinearity (the purple arrow is the frequency under test ω), our labeling, and their channel index relations.

The efficiency of a generic FWM process is weighted by the kernel, which depends on the phase-matching coefficient $\Delta\beta$, being maximum for $\Delta\beta = 0$. This can occur only when the four frequencies are degenerate or pair-wise degenerate, e.g., the carrier frequencies involved in the SPM and XPM effects. In Fig. 2.2, the contributions of SPM and XPM are thus highly efficient. Among the other effects, XCI is generally efficient since it involves closer frequencies coming from neighboring channels. In particular, XCI contributions are characterized by three spectral components in the channel under test (CUT) while the fourth component belongs to one of the two neighboring channels, or vice-versa (three frequencies in the interfering channel, and one in the CUT). Classic FWM and MCI are generally much less efficient since they involve far-away frequencies.

2.2.2 Closed-form expressions

Even by tailoring the integration domain to compute only valid contributions to the FWM process in Eq. (2.34), the frequency integrals cannot be solved analytically and call for a numerical computation, without extra approximations. Such a computation can be kept in the order of seconds, as will be discussed later in this chapter, which is relevantly smaller than the time required by the SSFM, but still too long for deciding the routing in a live network. The ultimate complexity reduction is enabled by approximated closed-form analytical expressions of Eq. (2.34). In this section, we present the closed-form expression for the SON/GN term variance derived in [3].

As a first approximation, the NLI is assumed to be uniformly distributed over the channel bandwidth, i.e., having a flat power spectral density equal to its value at the channel center frequency. To further simplify the analytical manipulation of Eq. (2.34), the matched filter is assumed to have a rectangular transfer function, as per

the noise equivalent bandwidth definition.

As a second key approximation, it can be assumed that the only non-negligible inter-channel nonlinear effect is XPM [3, 42]. Hence, FWM, XCI and MCI contributions are neglected. Under these assumptions, the GN term variance at the central frequency of the i th channel for a single-span link can be approximated as:

$$\sigma_{\text{GN},1}^2(i) \approx \sigma_{\text{SPM},1}^2(i) + \sum_{\ell \neq i} \sigma_{\text{XPM},1}^2(\ell) \quad (2.38)$$

where $\sigma_{\text{SPM},1}^2(i)$ is the single-span SPM variance at the central frequency of channel i , while $\sigma_{\text{XPM},1}^2(\ell)$ is the single-span XPM variance due to the interfering channel ℓ . Such single-span variances were derived in a closed-form manner in [3] by approximating the hexagons in Fig. 2.1 with squares, and are here reported:

$$\sigma_{\text{SPM},1}^2(i) \approx 3 \frac{16}{81} \gamma^2 \frac{L_{\text{eff}}^2 \alpha P_i^3}{2\pi^2 B^2 \beta_2} j \left[\text{Li}_2 \left(-j\pi^2 \frac{\beta_2}{\alpha} B^2 \right) - \text{Li}_2 \left(j\pi^2 \frac{\beta_2}{\alpha} B^2 \right) \right] \quad (2.39)$$

$$\begin{aligned} \sigma_{\text{XPM},1}^2(\ell) \approx & 6 \frac{16}{81} \gamma^2 \frac{L_{\text{eff}}^2 \alpha P_\ell^2 P_i}{2\pi^2 B^2 \beta_2} j \left[\text{Li}_2 \left(-j \frac{2\pi^2 \beta_2}{\alpha} \left(\ell \Delta f + \frac{B}{2} \right) B \right) \right. \\ & - \text{Li}_2 \left(j \frac{2\pi^2 \beta_2}{\alpha} \left(\ell \Delta f + \frac{B}{2} \right) B \right) - \text{Li}_2 \left(-j \frac{2\pi^2 \beta_2}{\alpha} \left(\ell \Delta f - \frac{B}{2} \right) B \right) \\ & \left. + \text{Li}_2 \left(j \frac{2\pi^2 \beta_2}{\alpha} \left(\ell \Delta f - \frac{B}{2} \right) B \right) \right] \end{aligned} \quad (2.40)$$

where Li_2 is the dilogarithm special function¹. Consistently with [3] these expressions are $(x+y)$ variances, related to the per-polarization variance by Eq. (2.35).

As shown in [4], the NLI variance generated in each span cannot be summed as for uncorrelated random variables. In other words, the NLI variance at the end of a N span link does not coincide with the sum of the individual per-span NLI variances, meaning that the correlations among the spans play an important role. The NLI thus

¹The special function $\text{Li}_2(x)$ is implemented as *dilog(1-x)* in MATLAB.

accumulates *coherently* along the link. Such a coherent accumulation is captured by the phased-array in Eq. (2.17) and is included in the closed-form expressions using the simple scaling rule [3, 47, 48]

$$\sigma_{\text{GN}}^2(i) \approx N^{1+\varepsilon} \sigma_{\text{GN},1}^2(i) \quad (2.41)$$

where ε plays the role of a coherence correction. An expression for this factor was derived in [3] as:

$$\varepsilon \approx \frac{\ln\left(1 + \frac{2}{\alpha NL} \frac{1-N+N\text{HarNum}(N-1)}{\text{asinh}\left(\frac{\pi^2}{2\alpha} \beta_2 B_{\text{WDM}}^2\right)}\right)}{\ln(N)} \quad (2.42)$$

with B_{WDM} the WDM bandwidth and $\text{HarNum}(n)$ the n th harmonic number. This correction factor can take values between 0 and 1. The case of $\varepsilon = 0$ means incoherent accumulation, hence the NLI variance grows with the number of spans. DU links are typically associated with a smaller value of ε compared to DM links ($\varepsilon \approx 0.22$ vs. $\varepsilon \approx 0.66$ in [36]), due to the decorrelation provided by the accumulated dispersion. Moreover, it was shown in [3] that the coherence factor decreases with the WDM bandwidth. For SMF fibers, an $\varepsilon \approx 0.3$ was measured by experiments [48]. The factor reaches values around 0.05 for wideband systems, where XPM is the dominating effect [49]. The accumulation of NLI can thus be safely considered incoherent in such systems, thanks to the high walk-off.

2.3 The Enhanced Gaussian noise model

The signal Gaussianity assumption at the core of the GN model allows significant simplifications since the FON and HON contributions to the overall NLI variance are zero in this framework. However, it was first shown in [42] that neglecting the FON contribution results in an over-estimation of the XPM variance, with a modulation-format dependent error. Such a result is justified by the fact that the FON contribution is weighted by the fourth-order cumulants of the transmitted symbols, which are generally negative for the most popular modulation formats [50]. As an example,

Modulation format	Factor Φ	Factor Ψ
64QAM	-0.619	1.79
16QAM	-0.68	2.08
QPSK	-1	4

Table 2.1: Value of Φ and Ψ in Eq. (2.22) weighting FON and HON contributions.

the values of the factors Φ and Ψ in Eq. (2.22) involved in the fourth and sixth-order moments are listed in Tab. 2.1. In particular, it is worth noting that the term Φ weighting the FON contribution has a negative sign for non-Gaussian distributed symbols, which is higher in absolute value for low cardinality modulation formats.

In [21] it was first highlighted that an accurate estimate of the NLI variance required a modulation-format aware model. A further step was taken in [40] and [41], where also the HON was included in the so-called EGN model. In addition, in [40] the FON contribution was included for all nonlinear effects (SPM, XPM, FWM, XCI, and MCI according to the nomenclature in Fig. 2.2).

In this section, we will review the expressions of the F4, Q4, and Q6 contributions to the NLI variance. These terms, combined with the SON/GN contribution expression, represent all the pieces required for evaluating Eq. (2.18).

Fourth-order noise

The FON contribution is the most important modulation-format correction to the GN model [21]. The general NLI variance expression in Eq. (2.18) exhibits two kinds of FON contributions, labeled as F4 and Q4, which differ in the indexes configuration.

We first focus on the F4 term, which yields the most significant contribution as shown in [41]. Such a term is associated with index configurations in the form **hhni** and **hnhi**. Contrary to the GN contribution, due to the constraints on the index com-

binations, the F4 part does not contribute to all the nonlinear effects. In other words, only sub-regions of the integration domains shown in Fig. 2.1 will participate in the computation of the F4 term variance. As an example, we depicted in Fig. 2.3 the same 9 channels scenario of the previous figure highlighting in red the valid integration domains. In particular, such a FON term contributes to SPM, XPM, and partially to XCI, while all the FWM and MCI regions do not satisfy the index constraint.

Similarly to the GN contribution, we now write the F4 term variance as:

$$\begin{aligned}
\sigma_{F4i_3}^2 &= \sum_{\mathbf{h}, \mathbf{n}} \kappa_{2;2}^{(\mathbf{h})} \kappa_{1;1}^{(\mathbf{n})} |\mathcal{X}_{\mathbf{hhni}} + \mathcal{X}_{\mathbf{hnh i}}|^2 \\
&= \frac{1}{8} \sum_{h_1, n_1} \sum_{h_2, n_2} \kappa_{2;2}^{(\mathbf{h})} \kappa_{1;1}^{(\mathbf{n})} \left(\sum_{h_3, n_3} \delta_{h_3 h_3} \delta_{i_3 n_3} + \sum_{h_3, n_3} \delta_{h_3 h_3} \delta_{i_3 n_3} \delta_{h_3 n_3} \delta_{i_3 h_3} \right. \\
&\quad \left. + \sum_{h_3, n_3} \delta_{h_3 n_3} \delta_{i_3 h_3} + \sum_{h_3, n_3} \delta_{h_3 n_3} \delta_{i_3 h_3} \delta_{h_3 h_3} \delta_{i_3 n_3} \right) |\mathcal{S}_{\mathbf{hhni}}|^2 \\
&= 5 \left(\frac{8}{9} \right)^2 \frac{1}{8} \gamma^2 \sum_{h_1, n_1, h_2, n_2} \kappa_{2;2}^{(\mathbf{h})} \kappa_{1;1}^{(\mathbf{n})} \iiint \iiint_{-\infty}^{\infty} \eta(\omega, \omega_1, \omega_2) \eta^*(\mu, \mu_1, \mu_2) \\
&\quad \times \tilde{G}_{h_1 h_2}^*(0, \omega + \omega_1 + \omega_2) \tilde{G}_{h_1 h_2}(0, \omega + \omega_2) \tilde{G}_{i_1 i_2}^*(0, \omega) \\
&\quad \times \tilde{G}_{n_1 n_2}(0, \omega + \omega_1) \tilde{G}_{h_1 h_2}(0, \mu + \mu_1 + \mu_2) \tilde{G}_{i_1 h_2}^*(0, \mu + \mu_2) \\
&\quad \times \tilde{G}_{i_1 i_2}(0, \mu) \tilde{G}_{n_1 n_2}^*(0, \mu + \mu_1) \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} \frac{d\omega}{2\pi} \frac{d\mu_1}{2\pi} \frac{d\mu_2}{2\pi} \frac{d\mu}{2\pi}
\end{aligned} \tag{2.43}$$

where, once again, we expressed the variance in terms of the scalar function $\mathcal{S}_{\mathbf{hhni}}$ exploiting its symmetries, and highlighting the combination of Kronecker's delta resulting from expressing $\mathcal{X}_{\mathbf{hhni}}$ in terms of $\mathcal{S}_{\mathbf{hhni}}$, as per Eq. (2.12). Such combination results in a weight 5 for the F4 term [41].

Following the same steps of the GN part, we can apply the Poisson formula and hence drop the time summations. As a consequence, also some integrals in frequency can be dropped as well. Applying similar ideas as in Eq. (2.33) we build the following system to identify the integrals that can be dropped:

$$\begin{cases} \mathbf{h} & \rightarrow \omega + \omega_1 + \omega_2 + \mu + \mu_2 = \omega + \omega_2 + \mu + \mu_1 + \mu_2 \\ \mathbf{n} & \rightarrow \omega + \omega_1 = \mu + \mu_1 \end{cases} \tag{2.44}$$

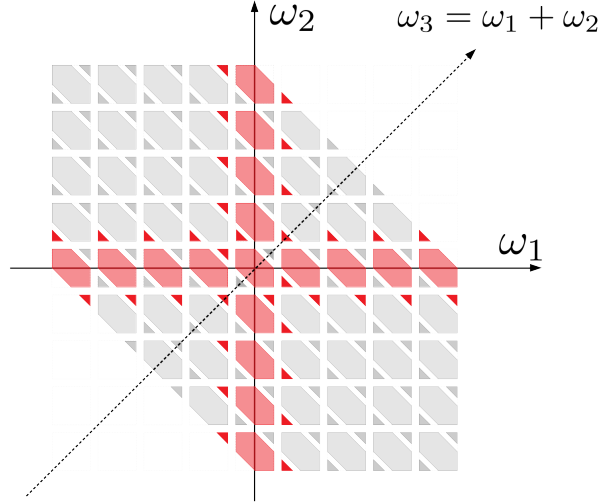


Figure 2.3: Integration domains with non-zero F4 contributions highlighted in red, with reference to Fig. 2.1.

whose solution is $\omega_1 = \mu_1$, $\omega = \mu$. We can thus drop two integrals, for instance the ones with (μ, μ_1) , by using the previous substitution. As a consequence, the NLI variance of the F4 contribution can be expressed as follows:

$$\begin{aligned}
 \sigma_{F4i_3}^2 &= 5 \left(\frac{8}{81} \right) \frac{1}{T^2} \gamma^2 \sum_{h,n} \kappa_{2;2}^{(h)} \kappa_{1;1}^{(n)} \iiint \int_{-\infty}^{\infty} \eta(\omega, \omega_1, \omega_2) \eta^*(\omega, \omega_1, \mu_2) \\
 &\quad \times \tilde{p}_h(\omega + \omega_1 + \omega_2 - \omega_h) \tilde{p}_h^*(\omega + \omega_1 + \mu_2 - \omega_h) \tilde{p}_h(\omega + \omega_2 - \omega_h) \\
 &\quad \times \tilde{p}_h^*(\omega + \mu_2 - \omega_h) |\tilde{p}_i(\omega - \omega_i)|^2 |\tilde{p}_n(\omega + \omega_1 - \omega_n)|^2 \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} \frac{d\omega}{2\pi} \frac{d\mu_2}{2\pi} \quad (2.45) \\
 &= 5 \sum_{k,s=1}^N \mathcal{C}_{F4}(k, s)
 \end{aligned}$$

where in the last equality we expressed the variance of the F4 term as a function of the span cross-correlations \mathcal{C}_{F4} , with similar motivations to get Eq. (2.37) for the GN term. We do not explicitly provide the expression of the cross-correlation since it can be simply obtained from the above equation by substituting the link kernel η with the fiber kernel up to span k and s .

We note that, regardless of the simplifications provided by the Poisson formula, the F4 variance still requires the computation of four integrals in frequency, contrary to three integrals in the GN contribution in Eq. (2.34).

Similarly to the F4 term, in Fig. 2.4 we show the integration domains satisfying the index combinations **hnni** associated with the Q4 term. In particular, we note that such an index configuration reduces the valid integration domains to only the hexagons and triangles lying along the ω_3 axis. As a main consequence, there is no Q4 contribution to XPM. For this reason, the F4 term typically dominates over the Q4 one, hence being the main FON contribution. Another difference with respect to F4 is that Q4 contributes partially to FWM and XCI, as highlighted in Fig. 2.4.

The variance of Q4 is given by the following:

$$\begin{aligned}
\sigma_{Q4i3}^2 &= \sum_{\mathbf{h}, \mathbf{n}} \kappa_{2,2}^{(\mathbf{h})} \kappa_{1,1}^{(\mathbf{n})} |\mathcal{X}_{\mathbf{hnni}}|^2 \\
&= \frac{1}{8} \sum_{h_1, n_1} \sum_{h_2, n_2} \kappa_{2,2}^{(\mathbf{h})} \kappa_{1,1}^{(\mathbf{n})} \left(\sum_{h_3, n_3} \delta_{h_3 n_3} \delta_{i_3 n_3} \right) |\mathcal{S}_{\mathbf{hnni}}|^2 \\
&= \left(\frac{8}{9} \right)^2 \frac{1}{8} \gamma^2 \sum_{h_1, n_1, h_2, n_2} \kappa_{2,2}^{(\mathbf{h})} \kappa_{1,1}^{(\mathbf{n})} \iiint \iiint_{-\infty}^{\infty} \eta(\omega, \omega_1, \omega_2) \eta^*(\mu, \mu_1, \mu_2) \quad (2.46) \\
&\quad \times \tilde{G}_{h_1 h_2}^*(0, \omega + \omega_1 + \omega_2) \tilde{G}_{n_1 n_2}(0, \omega + \omega_2) \tilde{G}_{i_1 i_2}^*(0, \omega) \\
&\quad \times \tilde{G}_{n_1 n_2}(0, \omega + \omega_1) \tilde{G}_{h_1 h_2}(0, \mu + \mu_1 + \mu_2) \tilde{G}_{n_1 n_2}^*(0, \mu + \mu_2) \\
&\quad \times \tilde{G}_{i_1 i_2}(0, \mu) \tilde{G}_{n_1 n_2}^*(0, \mu + \mu_1) \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} \frac{d\omega}{2\pi} \frac{d\mu_1}{2\pi} \frac{d\mu_2}{2\pi} \frac{d\mu}{2\pi}
\end{aligned}$$

where we observe the absence of any weight, differently from the F4 and GN term. This comes from applying Eq. (2.12) and noticing that: $\sum_{h_3, n_3} \delta_{h_3 n_3} \delta_{i_3 n_3} = 1$.

Similarly to the previous cases, the Q4 variance expression can be simplified by dropping integrals according to the linear system:

$$\begin{cases} \mathbf{h} & \rightarrow \omega + \omega_1 + \omega_2 = \mu + \mu_1 + \mu_2 \\ \mathbf{n} & \rightarrow \omega + \omega + \omega_1 + \omega_2 = \mu + \mu + \mu_1 + \mu_2 \end{cases} \quad (2.47)$$

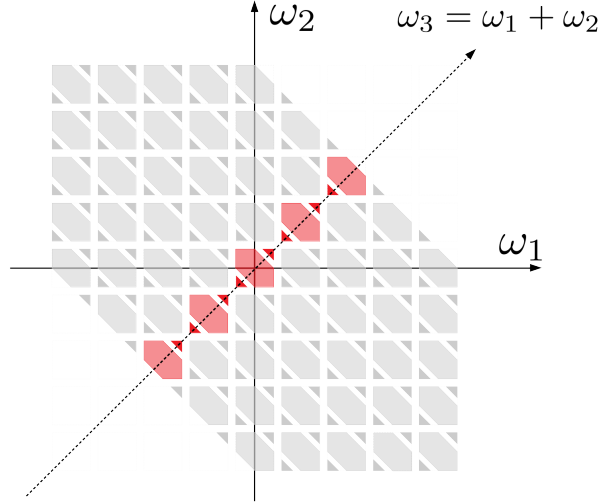


Figure 2.4: Integration domains with non-zero Q4 contributions highlighted in red, with reference to Fig. 2.1.

whose solution is $\mu_2 = \omega_1 + \omega_2 - \mu_1$, $\omega = \mu$. We can thus drop two integrals, for instance the ones with (μ, μ_2) , by using the previous substitution.

As a consequence, the NLI variance of the Q4 contribution can be expressed in the following simplified integral notation:

$$\begin{aligned}
 \sigma_{Q4i_3}^2 &= \frac{8}{81} \frac{1}{T^2} \gamma^2 \sum_{h,n} \kappa_{2;2}^{(h)} \kappa_{1;1}^{(n)} \iiint_{-\infty}^{\infty} \eta(\omega, \omega_1, \omega_2) \eta^*(\omega, \mu_1, \omega_1 + \omega_2 - \mu_1) \\
 &\quad \times |\tilde{p}_h(\omega + \omega_1 + \omega_2 - \omega_h)|^2 \tilde{p}_n(\omega + \omega_2 - \omega_h) \tilde{p}_n^*(\omega + \omega_1 + \omega_2 - \mu_1 - \omega_h) \\
 &\quad \times |\tilde{p}_i(\omega - \omega_i)|^2 \tilde{p}_n(\omega + \omega_1 - \omega_n) \tilde{p}_n^*(\omega + \mu_1 - \omega_n) \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} \frac{d\omega}{2\pi} \frac{d\mu_1}{2\pi} \\
 &= \sum_{k,s=1}^N \mathcal{C}_{Q4}(k, s).
 \end{aligned} \tag{2.48}$$

Higher-order noise

Finally, we now treat the so-called HON contribution to the NLI variance as expressed in Eq. (2.18). This term involves the sixth-order moment of the transmitted symbols and is associated with the indexes configuration **hhhi**. We thus note that such a HON contribution contributes marginally to the overall NLI variance since only a small subset of nonlinear effects can be associated with such an index configuration. Namely, Q6 contributes only to SPM and to the configurations of XCI having three spectral components in one of the two next-neighboring channels. Overall, we have three regions in the integration domains picture of Fig. 2.1 which are valid, as highlighted in red in Fig. 2.5. As a consequence, the HON variance generally represents the smallest contribution to the overall NLI variance as reported in [41].

Similarly to the GN and FON terms, we now express the Q6 contribution to the NLI variance as:

$$\begin{aligned}
\sigma_{Q6i3}^2 &= \sum_{\mathbf{h}} \kappa_{3;3}^{(\mathbf{h})} |\mathcal{X}_{\mathbf{hhhi}}|^2 \\
&= \frac{1}{8} \sum_{h_1, h_2} \kappa_{3;3}^{(\mathbf{h})} \sum_{h_3} \delta_{h_3 h_3} \delta_{h_3 i_3} |\mathcal{S}_{\mathbf{hhhi}}|^2 \\
&= \left(\frac{8}{9}\right)^2 \frac{1}{8} \gamma^2 \sum_{h_1, h_2} \kappa_{3;3}^{(\mathbf{h})} \iiint \iiint_{-\infty}^{\infty} \eta(\omega, \omega_1, \omega_2) \eta^*(\mu, \mu_1, \mu_2) \\
&\quad \times \tilde{G}_{h_1 h_2}^*(0, \omega + \omega_1 + \omega_2) \tilde{G}_{h_1 h_2}(0, \omega + \omega_2) \tilde{G}_{i_1 i_2}^*(0, \omega) \\
&\quad \times \tilde{G}_{h_1 h_2}(0, \omega + \omega_1) \tilde{G}_{h_1 h_2}(0, \mu + \mu_1 + \mu_2) \tilde{G}_{h_1 h_2}^*(0, \mu + \mu_2) \\
&\quad \times \tilde{G}_{i_1 i_2}(0, \mu) \tilde{G}_{h_1 h_2}^*(0, \mu + \mu_1) \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} \frac{d\omega}{2\pi} \frac{d\mu_1}{2\pi} \frac{d\mu_2}{2\pi} \frac{d\mu}{2\pi}.
\end{aligned} \tag{2.49}$$

We then apply the Poisson formula to simplify the variance expression. Since there is only one index involved, the integral to be dropped comes from the following:

$$\mathbf{h} \rightarrow \omega + \omega_1 + \omega_2 - \omega - \omega_1 - \omega - \omega_2 = \mu + \mu_1 + \mu_2 - \mu - \mu_1 - \mu - \mu_2 \tag{2.50}$$

having solution $\omega = \mu$ and hence allowing to drop only the integral for μ .

The resulting Q6 variance expression is then given by:

$$\begin{aligned}
\sigma_{Q6i_3}^2 &= \left(\frac{8}{9}\right)^2 \left(\frac{1}{2}\right)^3 \frac{1}{T} \gamma^2 \sum_h \kappa_{3;3}^{(h)} \iiint \iiint_{-\infty}^{\infty} \eta(\omega, \omega_1, \omega_2) \eta^*(\omega, \mu_1, \mu_2) \\
&\quad \times \tilde{p}_h(\omega + \omega_1 + \omega_2 - \omega_h) \tilde{p}_h^*(\omega + \mu_1 + \mu_2 - \omega_h) \tilde{p}_h(\omega + \omega_2 - \omega_h) \\
&\quad \times \tilde{p}_h^*(\omega + \mu_2 - \omega_h) |\tilde{p}_i(\omega - \omega_i)|^2 \tilde{p}_h(\omega + \omega_1 - \omega_n) \\
&\quad \times \tilde{p}_h^*(\omega + \mu_1 - \omega_n) \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} \frac{d\omega}{2\pi} \frac{d\mu_1}{2\pi} \frac{d\mu_2}{2\pi} \\
&= \sum_{k,s=1}^N \mathcal{C}_{Q6}(k, s).
\end{aligned} \tag{2.51}$$

Although Q6 represents the smallest contribution among those treated in this section, it is the term requiring the highest number of frequency integrals.

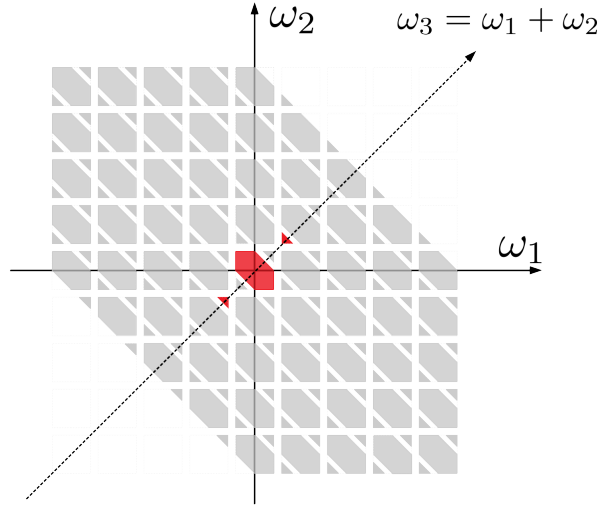


Figure 2.5: Integration domains with non-zero Q6 contributions highlighted in red, with reference to Fig. 2.1.

2.4 Monte Carlo Integration

A simple and effective method to evaluate the triple integrals at the heart of the NLI variance in Eq. (2.34) was proposed by Dar et al. in [46]. The authors proposed to exploit Monte Carlo (MC) integration instead of quadrature rules based on interpolating polynomial.

In brief, Monte Carlo integration relies on a random sampling of a function $f(x)$ to numerically compute an estimate of its integral, namely

$$I = \int_{\mathcal{V}} f(x) dx \approx \frac{V}{N_{\text{MC}}} \sum_{j=1}^{N_{\text{MC}}} f(x_j) \quad (2.52)$$

where N_{MC} is the number of samples, and $f(x_i)$ is the integrand evaluated in the samples drawn with a uniform distribution, and V is the volume of the set \mathcal{V} .

Apart from its simplicity, the MC method represents an efficient approach for multi-dimensional problems, since the standard deviation of the MC error, contrary to quadrature rules, scales with $1/\sqrt{N_{\text{MC}}}$ independently of the number of dimensions. Hence, it better tolerates the so-called “curse of dimensionality” [51].

The MC idea can be fully exploited by randomly sampling also the triple summation $\sum_{h,n,m}$ over the channel indexes in Eq. (2.34). The approach gives some advantages in evaluating FWM contributions because it removes the loops on the channel indexes in $\sum_{h,n,m}$ in favor of a single MC summation, whose vectorized implementation yields some improvements in numerical interpreted languages such as MATLAB. With reference to Fig. 2.1, we note that the FWM integration regions are placed within a hexagon macro-region. This shape results from the combination of two constraints: $h = m + n - i$ for the energy conservation and $(h, m, n) \in (1, \dots, N_{\text{ch}})$, since the channels involved in the FWM process must be within the WDM bandwidth. As a result, for a given channel index n and CUT index i , only the channels m for which the channel h belongs to the WDM comb are valid. It is then clear that an unconstrained sampling of channel indexes in $(1, \dots, N_{\text{ch}})$ would waste samples. Therefore, instead of sampling the channels independently over a square grid, we suggest sampling only in the visible islands in Fig. 2.1, thus by generating one channel index at random (e.g.,

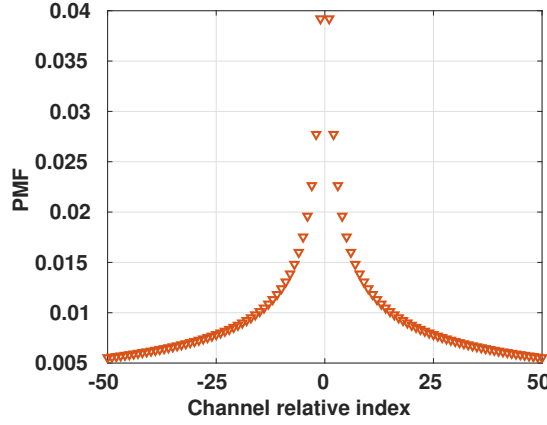


Figure 2.6: Example of PMF in Eq. (2.53) with parameter $q = 0.5$.

n) while conditioning the generation of the remaining one (e.g., m).

We also found some advantages in using importance sampling (IS) [52,53] through different sampling distributions than the simplest uniform distribution adopted by the MC method. A first IS strategy, also called stratified-sampling (SS) [51, 54], is to estimate independently each nonlinear effect according to the terminology of Fig. 2.1. Instead of uniformly selecting the channels joining the NLI process, we propose to estimate XPM due to channel n , $n \neq n_{\text{CUT}}$, by selecting it through a warped probability mass function (PMF) equal to

$$p(n) = \frac{c}{|n - n_{\text{CUT}}|^q} \quad (2.53)$$

with c a normalization constant. Such a sub-optimal PMF pushes more samples in channels closer to the CUT, with a probability governed by q . Intuitively, such channels are expected to dominate the NLI variance. The parameter q can be found by a least-squares fitting of the optimal SS-strategy $\sigma_n / \sum_n \sigma_n$ [51], with σ_n the standard deviation of the integrand within stratum n . The value of σ_n can be found by a short MC pre-run (e.g., a factor 100 shorter of the target number of samples) in the XPM hexagons, with minor implications on the total computational time. The least-squares fitting helps in smoothing the uncertainty of σ_n estimations. The FWM

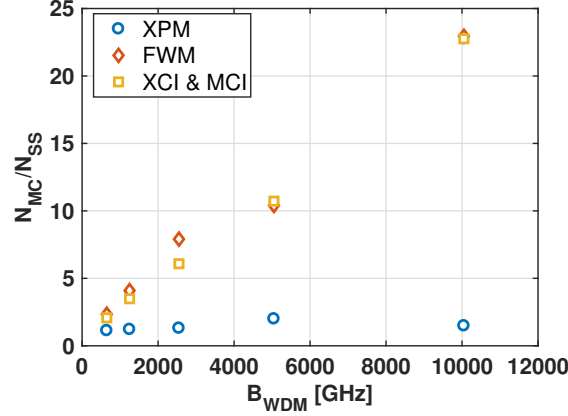


Figure 2.7: Ratio of the number of samples with MC and SS to get a given relative error (-30 dB) vs. WDM bandwidth. 10 span SMF link.

and XCI islands can be sampled in the same way by using the same PMF for each channel involved in the process. The sampling according to the warped PMF can be done by the inverse-transform sampling method [51]. For the sake of completeness, an example of the PMF $p(n)$ for a 101 channels transmission is reported in Fig. 2.6.

The samples generation and the pre-run add a small overhead to the computational time. However, the proposed SS strategy allows saving samples, particularly for the most computationally intensive terms as FWM, XCI, and MCI. Figure 2.7 shows the ratio of the number of samples required by MC and SS for the evaluation of each nonlinear effect at the same relative error of -30 dB. The figure refers to 10 spans. We note significant savings for multi-channel effects, with increasing gains for increasing bandwidths.

The variables $(\omega, \omega_1, \omega_2)$ within each island are sampled randomly by MC. However, for homogeneous point-to-point links, it is possible to apply IS even within the islands of Fig. 2.1, in particular for the GN contribution. The motivation comes from the observation that in such ultra-long links the phased-array term in Eq. (2.17) becomes a peaked function, whose random sampling is very inefficient by standard MC. The problem can be circumvented in the following way. The triple integral in

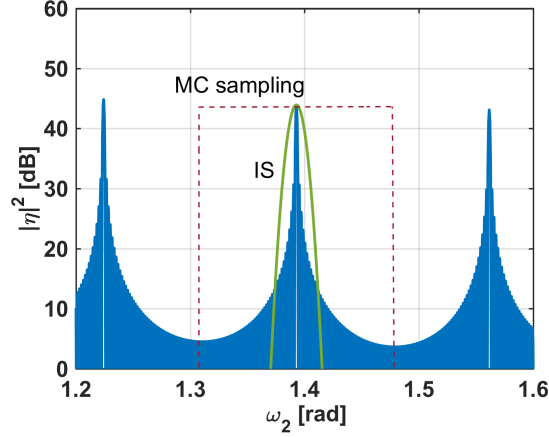


Figure 2.8: Example of kernel and Monte Carlo (MC) or importance sampling (IS) distributions for sampling the frequency ω_2 normalized to the symbol rate. Both sampling strategies randomly select a lobe of the kernel.

Eq. (2.34) is weighted by the kernel function, hence by the phase-matching coefficient $\Delta\beta$ between the four frequencies involved in the FWM process. If, for instance, we sample (h, m, n) in Eq. (2.34) by a discrete uniform distribution and (ω, ω_1) by a uniform distribution as in the MC method, it is convenient to sample the remaining ω_2 over the regions where the kernel takes the largest absolute value. Such regions appear periodically at ω_2 values where the phased array has resonances, that is for $\Delta\beta = 2\pi k$ with k an integer. Figure 2.8 sketches the idea, where we can observe: the kernel function (absolute squared value) versus the variable ω_2 for a given random choice of the remaining variables; the probability density function (PDF) of ω_2 with pure MC sampling (red dashed) and an example of importance sampling by using a Gaussian PDF for ω_2 (solid green).

In practice, pure MC is equivalent to randomly selecting a given lobe of the phased-array, i.e., a value of k , with a discrete uniform distribution, and to randomly sample with a uniform distribution between the two neighboring lobes, as illustrated in the figure. On the other hand, IS concentrates probability around the peak. The best IS PDF replicates the kernel shape. However, it may be better to use simpler PDFs

B_{WDM} [GHz]	length [km]	std(a_{NL} dB)		time [s]	
		MC	IS	MC	IS
50	1000	0.01	0.01	0.1	0.1
50	10000	0.04	0.04	0.1	0.1
5050	1000	0.08	0.07	1.1	1.5
5050	10000	0.24	0.04	1.1	1.5
10050	1000	0.10	0.09	1.1	1.5
10050	10000	0.37	0.06	1.1	1.5

Table 2.2: Computational time per channel and accuracy. Standard Monte Carlo (MC) method vs importance sampling (IS) technique.

in favor of simpler algorithms, motivated by the observation that a PDF sufficiently concentrated around the peak, may give some computational advantages. In particular, we used a Gaussian PDF of variance equal to twice the variance of the absolute squared value of the kernel around a lobe.

As a reference, Tab. 2.2 shows some values of the uncertainty on the normalized variance $a_{\text{NL}} \triangleq \sigma_{\text{GN}}^2/P^3$ obtained by evaluating Eq. (2.34) with the proposed methods, and the corresponding computational time in different scenarios, by using MC or IS with 10^6 samples per effect. All the results are referred to the central WDM channel and are obtained with an INTEL XEON E5-2650 v4 2.20GHz CPU-based architecture. IS was applied only to XPM and FWM hexagons of Fig. 2.1 which experience the largest number of phased-array peaks.

The computational time per channel, reported in Tab. 2.2, is very short. If the

NLI variance of each channel is required, the computation has to be repeated for each of them. However, since the NLI variance is a smooth function over the WDM bandwidth, we suggest estimating it just for a few channels and to use interpolation to find the missing values.

Part I

Ultra-wideband transmissions

Chapter 3

EGN model in the presence of Stimulated Raman Scattering

The imaginary part of the Raman response leads to SRS which amplifies low frequencies at the expense of high frequencies. While SRS among WDM channels can be safely neglected in C band transmissions (≈ 5 THz), as the peak value of the Raman gain is at a frequency shift ≈ 14 THz [7], its effect becomes relevant in ultra-wideband transmissions covering the C+L band (≈ 10 THz bandwidth).

Due to SRS, the signal power profile with respect to frequency becomes tilted. Such an SRS-induced tilt can be removed by means of equalizers. However, the distributed interaction during propagation between the Kerr effect and the Raman scattering cannot be removed by such equalizers. As a consequence, the NLI variance exhibit a residual SRS tilt. In addition, a residual SRS tilt is present even in the ASE noise variance, with a different slope.

Such an undesired SRS-tilt becomes increasingly important for increasing bandwidths. Since the numerical and experimental investigation of ultra-wideband transmissions is extremely complex (see Eq. (1.59)), it is fundamental to find fast analytical models able to capture the SRS impact on the system performance.

Several efforts have been made in the literature to model the impact of SRS on the received signal. An extension of the GN model to include inter-channel SRS in

the estimation of the NLI variance was proposed in integral form [55–57], and in closed-form [5, 58] for the SPM and XPM contributions.

The short computational time of Raman-aware GN models enables addressing complex design problems, such as the optimal channel power allocation at the transmitter side to counteract the SRS effect, as done in [55, 59]. In both works, the authors tackled the problem of optimal power allocation in the presence of uncompensated SRS. These studies were motivated by the fact that in practical systems the dynamic gain equalizers (DGEs) for the compensation of the SRS on the signal power are not placed after each span. For instance, a compensation period value of 4-5 spans can be found in the terrestrial systems literature [55, 59, 60].

In this chapter, we first present an extension of the EGN model to include the SRS gain which allows us to investigate the impact of SRS on the NLI variance and SNR for different modulation formats in ultra-wideband transmissions. Similarly to [61] this model relies on the presence of an ideal DGE at each span end. However, contrary to the aforementioned work, the proposed SRS-EGN model accounts also for XCI, MCI, and FWM (according to the nomenclature in Chapter 2, see Fig. 2.2) in a modulation format-aware manner.

Then, we extend such a model in the DGE-SRS-EGN model which accounts for a sparse DGE position along the optical link for the compensation of the SRS gain on the signal power. We show that such a sparse positioning of DGE has serious implications on the model accuracy. The proposed model includes the interaction between the nonlinear Kerr effect and accumulated SRS in a modulation format-aware manner. We also provide a simple extension to account for a signal-power pre-emphasis at the transmitter side.

Finally, in a GN model framework, we extend the closed-form expressions of [5] to include such sparse DGE positioning along the link. We exploit such closed-form expressions to investigate the optimization of signal-power pre-emphasis in the shape of an opposite SRS gain, to maximize either (i) the minimum SNR in the WDM comb or (ii) the total AIR.

The work presented in this chapter was published in [30, 62–65].

3.1 Stimulated Raman Scattering

The power evolution along distance z for the i th WDM channel, with $i = 1$ the leftmost channel in frequency and $i = N_{\text{ch}}$ the rightmost channel in frequency, is governed by the following [7, 66]:

$$\frac{dP_i}{dz} = -P_i \sum_{j=1}^{i-1} \frac{\omega_j}{\omega_i} \tilde{g}_R(|\omega_i - \omega_j|) P_j + P_i \sum_{j=i+1}^{N_{\text{ch}}} \tilde{g}_R(|\omega_i - \omega_j|) P_j - \alpha(\omega_i) P_i \quad (3.1)$$

with P_i the i th channel power at carrier frequency ω_i , α the fiber attenuation coefficient, and $\tilde{g}_R(|\omega_i - \omega_j|)$ the polarization-averaged Raman gain which is a function of the spacing between channel i and j . The first term on the right-hand side accounts for the depletion of the channel power P_i due to the amplification of the channels $j = 1, \dots, i-1$ centered at lower frequencies while the second term is the gain experienced by the i th channel thanks to the power transfer from channels at higher frequencies. Solving the power evolution in the presence of SRS requires the numerical solution of a system of coupled ordinary differential equations (ODEs), one ODE for each WDM channel.

An example of normalized polarization-averaged Raman gain versus the frequency shift between two channels is depicted in Fig. 3.1. It can be seen that such a gain profile can be approximated as a linear function only up to its peak profile, hence for a frequency shift smaller than ≈ 14 THz. This reasoning was applied in [67] in the derivation of the so-called *triangular approximation* of the Raman gain, where the gain is approximated as a linear function from the origin up to the peak value, then with zero for all the frequency components higher than the peak frequency.

In [69] it was shown that Eq. (3.1) allows an analytical solution under the following assumptions: i) linear dependency of $\tilde{g}_R(|\omega_i - \omega_j|)$ on the frequency shift $\omega_i - \omega_j$, ii) frequency-independent fiber attenuation coefficient, and iii) $\frac{\omega_j}{\omega_i} \approx 1$. The power evolution equation can be thus simplified in the following

$$\frac{dP_i}{dz} = P_i \sum_{j=1}^{N_{\text{ch}}} C_r(\omega_j - \omega_i) P_j - \alpha P_i \quad (3.2)$$

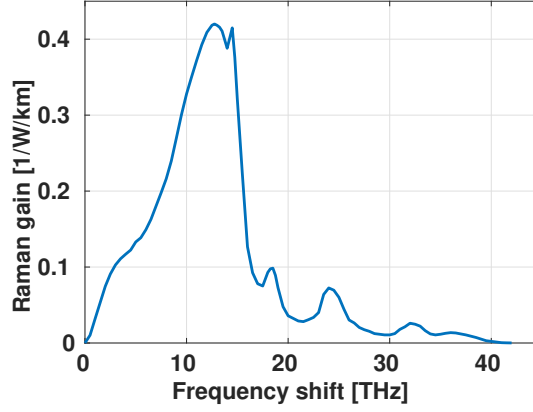


Figure 3.1: Raman gain profile for a SMF interpolated from [7] with a peak value of $0.42 \text{ (W} \cdot \text{km)}^{-1}$ as per [68].

where the first term in the right-hand side accounts for the amplification of the power P_i at the expenses of each channel j in the comb such that $\omega_j > \omega_i$, and depletion to the benefit of a channel having $\omega_j < \omega_i$. The coefficient C_r represents the slope of the linear approximation of the Raman gain and can be found by means of a linear regression on the experimental data.

The system of coupled ODEs can be written as a single equation in terms of WDM power spectral density (PSD) \mathcal{G} instead of channel powers, namely:

$$\frac{d\mathcal{G}(z, \omega)}{dz} = \mathcal{G}(z, \omega) C_r \int_{-\infty}^{\infty} (\nu - \omega) \mathcal{G}(z, \nu) d\nu - \alpha \mathcal{G}(z, \omega) \quad (3.3)$$

where the summation over the channels was replaced by an integration in frequency. Dividing Eq. (3.3) by $\mathcal{G}(z, \omega)$ and taking the derivative with respect to frequency:

$$\frac{d}{d\omega} \left(\frac{\frac{d\mathcal{G}(z, \omega)}{dz}}{\mathcal{G}(z, \omega)} \right) = -C_r \int_{-\infty}^{\infty} \mathcal{G}(z, \nu) d\nu \quad (3.4)$$

where we can recognize in the frequency integral on the right-hand side the total WDM power at coordinate z , which is assumed to decrease exponentially along the fiber length, namely:

$$\int_{-\infty}^{\infty} \mathcal{G}(z, \nu) d\nu = P_t e^{-\alpha z} \quad (3.5)$$

where $P_t = \int_{-\infty}^{\infty} \mathcal{G}(\nu) d\nu$ is the WDM total power at the transmitter side, with $\mathcal{G}(\nu) = \mathcal{G}(0, \nu)$. Equation (3.4) can thus be rewritten in the following compact form:

$$\frac{d}{d\omega} \frac{d}{dz} \ln \mathcal{G}(z, \omega) = -C_r P_t e^{-\alpha z}. \quad (3.6)$$

Following [69], integrating Eq. (3.6) first with respect to distance and then with respect to frequency yields a closed-form solution, by imposing the exponential decay condition for the total power in Eq. (3.5).

With respect to the original problem in Eq. (3.2), hence from the channel power point of view, such a solution can be expressed as follows:

$$P_i(z) = e^{-\alpha z} H(\omega_i, z) P_i(0) \quad (3.7)$$

hence the signal power undergoes fiber attenuation and SRS gain, defined as:

$$H(\omega_i, z) = \frac{P_t e^{-C_r P_t L_{\text{eff}}(z)} \omega_i}{\int_{-\infty}^{\infty} \mathcal{G}(\nu) e^{-C_r P_t L_{\text{eff}}(z)} \nu \frac{d\nu}{2\pi}} \quad (3.8)$$

where $\mathcal{G}(\nu)$ indicates the WDM signal PSD at the fiber input. From Eq. (3.7) it can be seen that the SRS gain induces a frequency-dependent tilt in the signal power profile and it does not affect the total power, which vanishes only through the fiber attenuation.

Figure 3.2 shows the power versus frequency after a single span of 100 km of SMF having attenuation coefficient $\alpha = 0.2$ dB/km for a 10 THz transmission with total WDM power of 23 dBm. Due to the SRS-induced power transfer, the originally frequency-flat power profile exhibits a tilt. The figure compares the numerical solution¹ of the system of coupled differential equations in Eq. (3.1) with the analytical expression in Eq. (3.7), based on the triangular approximation. Note that after the fiber we compensated the fiber losses. It can be seen that the analytical expression

¹Solved with the *ode45* MATLAB function.

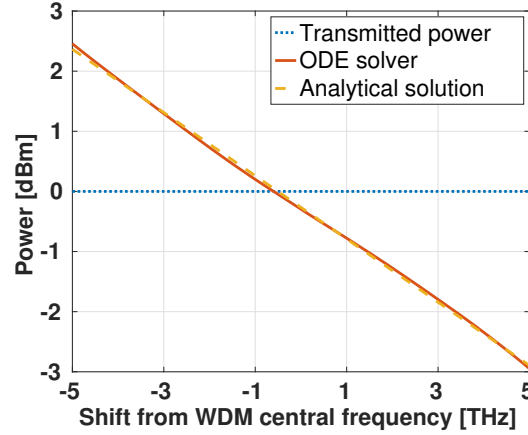


Figure 3.2: Power after one span of SMF for a 10 THz transmission. Solid: numerical solution of Eq. (3.1). Dashed: analytical solution in Eq. (3.7). Dotted: transmitted power.

based on the linear approximation of the Raman gain is in excellent agreement with the numerical solution in the bandwidth under test. The coefficient of the linear approximation of the Raman gain was $C_r = 0.028/(2\pi) \text{ (W} \cdot \text{km} \cdot \text{THz} \cdot \text{rad)}^{-1}$ found by means of a linear regression of the gain in Fig. 3.1 over 10 THz.

The analytical expression of the power evolution in Eq. (3.7) has two main advantages compared to the numerical solution. First, it allows to include the SRS in the GN/EGN model by modifying the power profile during propagation. Second, it highlights the key quantities involved in the SRS gain, such as the total WDM power. For instance, Fig. 3.3 shows the SRS gain versus frequency for the same WDM transmission of Fig. 3.2 with different total power values. It can be seen that the SRS tilt is enhanced by the increase of power.

It is worth noting that Eq. (3.1) is not limited to describe the power evolution of a single-span system. In fact, the system of differential equations can be solved span-by-span by applying at each iteration the correct initial condition, i.e., the signal power at the input of the m th span is the solution of the $(m - 1)$ th iteration. Such a span input power coincides with the power launched into the link only in the presence

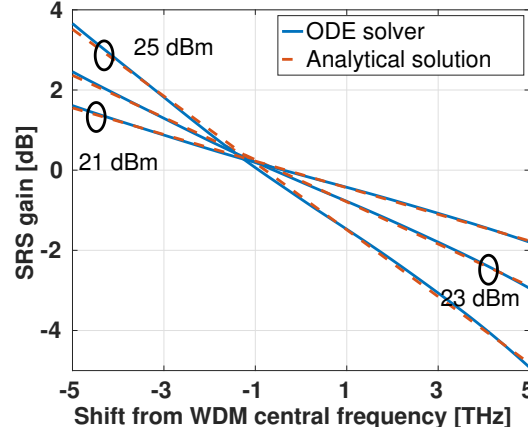


Figure 3.3: SRS gain after one span of SMF for a 10 THz transmission at three different total WDM powers. Solid: numerical solution of Eq. (3.1). Dashed: analytical solution in Eqs. (3.7)-(3.8).

of a DGE at each span-end, which restores the nominal power equalizing the SRS gain. However, in the most general case of arbitrary DGE positioning along the link, there is an uncompensated SRS gain that accumulates from DGE-to-DGE. Figure 3.4(a) sketches an example. The link is composed of N_D link sections, identified by the presence of a DGE at the end of each section. In the most general case, the sections can have a different number of spans N_s each. To highlight the effects of the SRS accumulation across the spans, Fig. 3.4(b) sketches an example of the signal PSD evolution in the first section of a link with $N_s = 3$ spans.

Under the same assumption of Eq.(3.2), the problem of the power evolution in the presence of accumulated SRS can be solved analytically following similar steps used to derive Eq. (3.7). Let us generalize the definition of SRS gain at coordinate z in Eq. (3.8), to the SRS gain experienced by the WDM channel centered at frequency ω_i from the input of link up to coordinate z of the m th span, with $m = 1, \dots, N_s$, as

$$H_m(\omega_i, z) = \frac{P_i e^{-m C_r P_i L_{\text{eff}}(z) \omega_i}}{\int_{-\infty}^{\infty} \mathcal{G}(\nu) e^{-m C_r P_i L_{\text{eff}}(z) \nu} \frac{d\nu}{2\pi}} \quad (3.9)$$

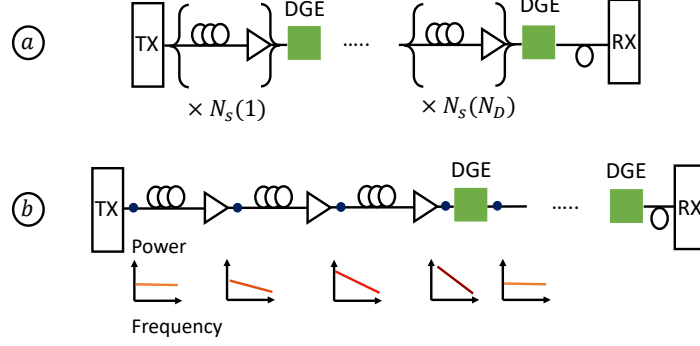


Figure 3.4: Sketch of: a) generic link structure with N_D sections having a variable number of spans $N_s(d)$ each, and b) WDM PSD tilt evolution due to the accumulated SRS in a link section with $N_s = 3$ spans.

such that Eq. (3.8) is a particular case of the more general Eq. (3.9) with $m = 1$. Thus, Eq. (3.7) with the Raman gain in Eq. (3.9) represents the analytical solution of the multi-span power evolution and encompasses all the arbitrary placement of the DGE along the link. As an example, Fig. 3.5 compares the SRS gain after $m = 1, 2, 4$ spans obtained by numerically solving the systems of ODEs and the theoretical gain in Eq. (3.9), showing an excellent agreement.

3.2 SNR generalization

After SRS gain equalization on the desired signal at the receiver, the SNR expression in Eq. (1.64) must be generalized to include the residual SRS gain on the NLI and ASE variance. Such a residual gain yields a frequency-dependent SNR, namely

$$\text{SNR}(\nu) = \frac{P}{\sigma_{\text{NLI}}^2(\nu) + \sigma_{\text{ASE}}^2(\nu)} \quad (3.10)$$

where ν is a generic WDM channel central frequency.

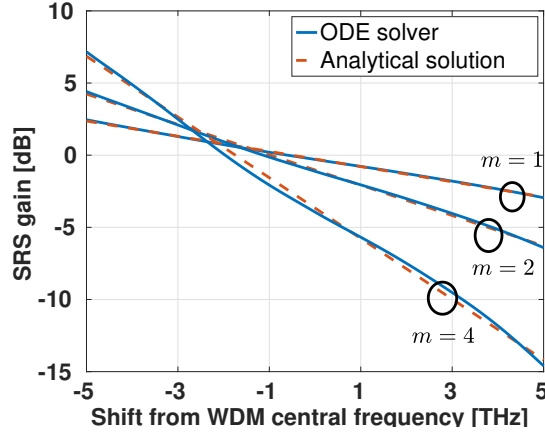


Figure 3.5: Raman gain after m spans of SMF without DGEs for a 10 THz transmission. Solid: numerical solution of Eq. (3.1). Dashed: analytical solution in Eqs. (3.7)-(3.9). Total WDM power: 23 dBm.

The ASE variance in Eq. (1.61) can thus be generalized as follows:

$$\begin{aligned}\sigma_{\text{ASE}}^2(\nu) &= h\nu FGB \sum_{d=1}^{N_D} \sum_{k=1}^{N_s(d)} \frac{H_{N_s(d)-k}(\nu, L)}{H_{N_s(d)}(\nu, L)} \\ &= h\nu FGB \sum_{d=1}^{N_D} \sum_{k=1}^{N_s(d)} H_k^{-1}(\nu, L)\end{aligned}\quad (3.11)$$

where $H_{N_s(d)-k}(\nu, L)$ is the SRS gain accumulated in the subsequent $N_s(d) - k$ spans and $H_{N_s(d)}^{-1}(\nu, L)$ is the gain of an ideal DGE which perfectly compensates the SRS gain experienced by the signal power in the d th link section of $N_s(d)$ spans. It is worth noting that Eq. (3.11) assumes ideal DGEs, i.e., noiseless and lossless devices. An extension to lossy and noisy DGE nodes will be tackled in Section 3.5.

An equivalent block diagram representation of a link section is sketched in Fig. 3.6. For the sake of simplicity, the diagram does not report the NLI variance which will be extensively treated in what follows. Note that we are assuming the signal power to be transparent from section to section since the DGE at the end of each section restores the nominal power level.

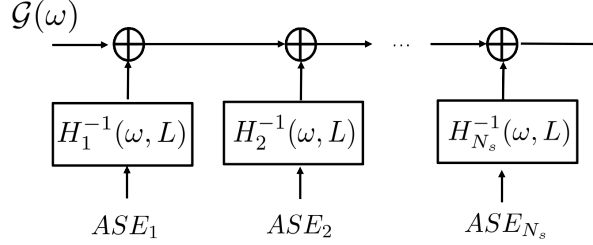


Figure 3.6: Equivalent block diagram representation of a link section, i.e., the cascade of N_s amplified spans and a DGE. ASE_k , $k = 1, \dots, N_s$, is the variance of the ASE noise introduced by the k th amplifier.

In the following sections, we will focus on the generalization of the link kernel presented in Section 2.1.1 to include SRS, which can be used in the GN/EGN model frameworks presented in Chapter 2 to estimate the NLI variance.

3.3 The SRS-EGN model

Including SRS in the NLI model requires a frequency-dependent power profile, due to the nature of the Raman scattering process. The Manakov equation in Eq. (1.44) and its RP1 solution in Eq. (1.49)-(1.53), can thus be extended by substituting the fiber attenuation α with a proper frequency-dependent term including the SRS gain, as discussed in Chapter 1.

With ideal DGEs and amplifiers after each span, both the SRS gain and the fiber loss are perfectly recovered at the span end. Hence, the normalized power profile, with the SRS gain triangular approximation, becomes

$$\rho(\zeta, \omega) = \Upsilon(L_{\text{eff}}(\zeta)) e^{-\alpha \zeta} e^{-P_t C_r L_{\text{eff}}(\zeta) \omega} \quad (3.12)$$

where the factor Υ is a normalization factor ensuring that the total power decreases exponentially along the fiber length, see Eq. (3.5), yielding:

$$\Upsilon(L_{\text{eff}}(\zeta)) \triangleq \frac{P_t}{\int_{-\infty}^{\infty} \mathcal{G}(\nu) e^{-P_t C_r L_{\text{eff}}(\zeta) \nu} \frac{d\nu}{2\pi}}. \quad (3.13)$$

Similarly to the previous analyzed case without SRS, in which $\rho = e^{-\alpha\zeta}$, we note that the normalized power profile is span-independent. The link kernel in Eq. (2.13), that we report here for the reader convenience

$$\eta(\omega, \omega_1, \omega_2) = \sum_{k=1}^N e^{j\Delta\beta(k-1)L} \int_0^L e^{j\Delta\beta\zeta} e^{-\alpha\zeta} d\zeta \quad (3.14)$$

can then be written in terms of a generic power profile as follows:

$$\begin{aligned} \eta(\omega, \omega_1, \omega_2) &= \sum_{k=1}^N e^{j\Delta\beta(k-1)L} \int_0^L e^{j\Delta\beta\zeta} \\ &\times \sqrt{\frac{\rho(\zeta, \omega + \omega_1 + \omega_2) \rho(\zeta, \omega + \omega_1) \rho(\zeta, \omega + \omega_2)}{\rho(\zeta, \omega)}} d\zeta \end{aligned} \quad (3.15)$$

where the only difference with respect to the link kernel treated in Chapter 2 is the presence of $\rho(\zeta, \omega)$ instead of the power loss induced by the fiber attenuation.

After the substitution of the extended power profile in Eq. (3.12), the kernel can still be expressed as the product of two terms:

$$\eta(\omega, \omega_1, \omega_2) = \chi(\omega, \omega_1, \omega_2) \eta_1^R(\omega, \omega_1, \omega_2) \quad (3.16)$$

where χ is the phased-array term defined in Eq. (2.17), accounting for accumulated effects, while η_1^R is a generalization of the single-span kernel η_1 to the case with SRS, equal to [57]

$$\eta_1^R = \int_0^L \Upsilon(L_{\text{eff}}(\zeta)) e^{-\alpha\zeta} e^{-P_{\text{Cr}} L_{\text{eff}}(\zeta)(\omega + \omega_1 + \omega_2)} e^{j\Delta\beta\zeta} d\zeta \quad (3.17)$$

$$\approx \left(1 - s\right) \eta_1 + \frac{1 - e^{-2\alpha L + j\Delta\beta L}}{2\alpha - j\Delta\beta} s \quad (3.18)$$

where $s \triangleq \frac{P_{\text{Cr}}(\omega + \omega_1 + \omega_2)}{\alpha}$, and η_1 is given in Eq. (2.16). The last approximation has been derived in [5] under the assumption that the power is uniformly distributed

across the WDM bandwidth and that SRS is weak. The weak SRS assumption allows to expand the SRS term with a Taylor series truncated to the first order, namely:

$$e^{-P_t C_r L_{\text{eff}}(\zeta)\omega} \approx 1 - P_t C_r L_{\text{eff}}(\zeta)\omega \quad (3.19)$$

for the SRS exponential in Eq. (3.17) and Eq. (3.13). In particular, the first order series expansion in Eq. (3.19) yields $\Upsilon \approx 1$ when the power is uniformly distributed over the WDM bandwidth, which significantly simplifies the analytical expression of the kernel, avoiding the numerical computation of the integral in Eq. (3.15).

Due to the shape of the Raman gain versus frequency shift depicted in Fig. 3.1, it is clear that its linear approximation holds only for bandwidths not exceeding ≈ 15 THz. For larger bandwidths, not only the triangular approximation of the Raman gain does fail, but also the assumption of a frequency-independent fiber attenuation becomes unjustified [70].

The extension of the EGN model in the SRS-EGN model goes straightforward by replacing the new link kernel in the framework presented in Chapter 2 to obtain all the contributions to the NLI variance. However, it is worth noting that, in the presence of SRS, the link kernel depends on the total power P_t . As a consequence, the scaling rule with P^3 for the NLI variance of Eq. (1.65) holds *only* at fixed total power.

3.3.1 Numerical validation

We implemented the new generalized SRS-EGN model by adopting the link kernel expression in Eq. (3.16) in the EGN model. In particular, the frequency integrals involved in the variance computation are performed by means of Monte Carlo integration [46], and all the nonlinear effects (SPM, XPM, FWM, XCI, and MCI) are included.

The links under test were DU links composed of SMFs and based on ideal end-span lumped amplification with frequency-flat gain. The SRS gain on the signal was recovered by ideal DGEs at each span-end. The SMFs parameters are listed in Tab. 3.1. In addition, the slope of the triangular approximation of the Raman gain of the fiber was $C_r = 0.028/(2\pi) \text{ (THz}\cdot\text{rad}\cdot\text{km}\cdot\text{W})^{-1}$.

Fiber parameter	Value
Attenuation α	0.2 dB/km
Length L	100 km
Dispersion D	17 ps/(nm·km)
Dispersion slope S	0.057 ps/(nm ² ·km)
Nonlinear coefficient γ	1.26 1/(W·km)

Table 3.1: Single-mode fiber parameters considered in this work.

While the EGN model relies on a linear approximation of the Raman gain, in the SSFM simulations we used a polynomial interpolation of the experimental gain as in Fig. 3.1, and we accounted for both real and imaginary parts of the Raman spectrum, see Eq. (1.37).

The WDM signal was a comb of PDM signals with channel spacing $\Delta f = 50$ GHz. Each channel was shaped with RRC pulses of 0.01 roll-off and modulated at the symbol rate 49 Gbaud. The number of symbols was 67550, and the SSFM solved the Manakov equation with SRS through a symmetric-step updated according to the CLE criterion with a maximum tolerable FWM phase matching $\Delta\phi_{\text{FWM}} = 10$ rad [28]. At the receiver, ideal chromatic dispersion compensation was followed by a matched filter detection and a least-squares equalizer with 1 tap that recovered the average polarization and phase mismatch. The NLI variance in the SSFM simulations was estimated as the variance of the difference between the received and the transmitted sequence of complex symbols, in the absence of ASE noise.

In a first test, we investigated 64QAM signals with a 0 dBm transmitted power per channel, over a WDM bandwidth of 10 THz. As a reference, such a channel power is close to the NLT value for the central channel in the absence of SRS and with EDFA's

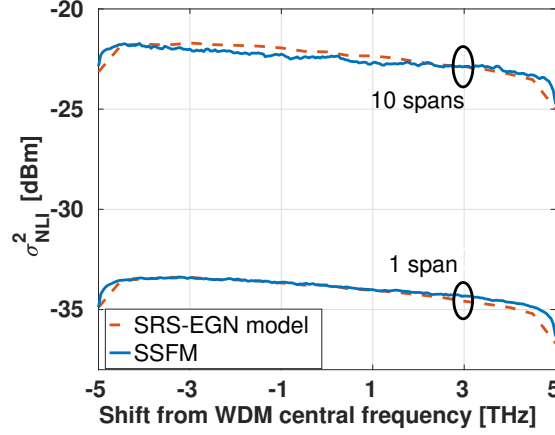


Figure 3.7: NLI variance σ_{NLI}^2 vs. frequency shift. PDM-64QAM, 201 channels, after 100 km and 10×100 km. DGE placed at each span end. Solid line: SSFM simulations. Dashed lines: SRS-EGN model.

noise figure $F = 5$ dB. Figure 3.12 shows the NLI variance versus the frequency shift after $n \times 100$ km with $n = 1, 10$. We used solid lines for SSFM results and dashed lines for the SRS-EGN model estimation. It is worth noting that the NLI variance in Fig. 3.7 exhibits a tilt, regardless of the presence of a DGE at each span-end. This tilt is due to the distributed interaction between SRS and Kerr effects and cannot be removed by the zero-forcing equalization of the DGEs. It can be seen from the figure that the model is in good agreement with the SSFM simulations, confirming the capability of the SRS-EGN model to capture this residual SRS gain on the NLI variance, as well as the validity of the linear approximation of the Raman profile up to a bandwidth of 10 THz.

3.3.2 Impact of the modulation format

To appreciate the impact of the modulation format, Fig. 3.8 shows the NLI variance as a function of the frequency shift with respect to the WDM central frequency, obtained with the SRS-EGN model. We compare PDM signals modulated by 64QAM and QPSK. We considered a 10 THz transmission with a symbol rate of 49 Gbaud

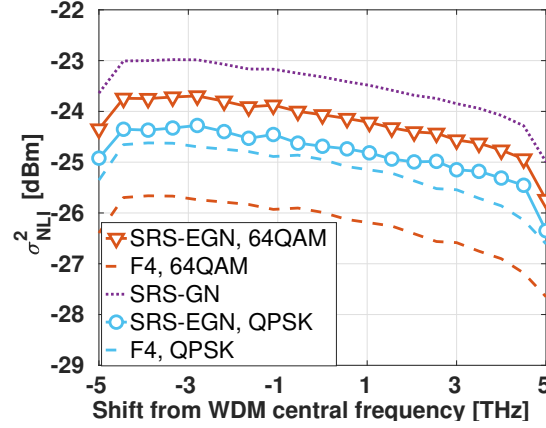


Figure 3.8: NLI variance σ_{NLI}^2 in the presence of SRS vs. frequency shift. 201 channels, 10×100 km SMF link. Solid lines: SRS-EGN model. Circles: 64QAM. Triangles: QPSK. Dashed lines: F4 contribution in absolute value.

and channel spacing of 50 GHz over a 10×100 km DU-link composed of SMFs. The total transmitted power was $P_t = 24$ dBm. As can be seen in the figure, the SRS-GN model provides the same estimated value of σ_{NLI}^2 for both modulation formats. The estimation error introduced by the GN-model is more severe for the QPSK than the 64QAM, due to a higher, in absolute terms, FON contribution. In particular, we reported with dashed lines the F4 term contribution, in absolute value, which is 2 dB higher for QPSK than for 64QAM. However, it is worth noting that the SRS induces an identical tilt in the NLI variance profile regardless of the modulation format. Figure 3.8 suggests that such a tilt might be mitigated by adopting heterogeneous modulation formats. Namely, a possible strategy is to assign the modulation format with smaller σ_{NLI}^2 , due to a more negative F4 contribution, to channels at lower frequencies which are those most amplified by the SRS, to counterbalance the tilt in the NLI variance profile.

The results of a first investigation on the heterogeneous modulation format strategy are reported in Fig. 3.9 for a simple scenario with a 51 channel WDM transmission, for the same link of Fig. 3.8. In this case, we considered a channel power

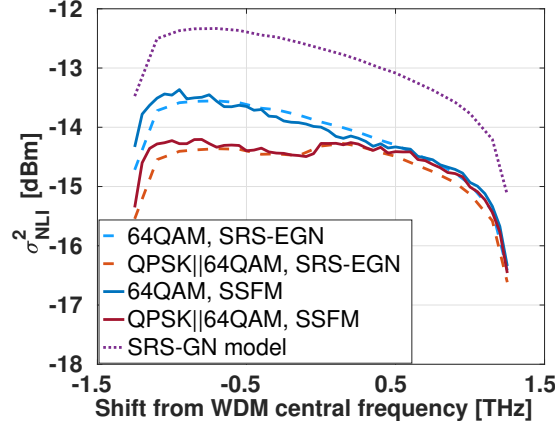


Figure 3.9: NLI variance σ_{NLI}^2 in the presence of SRS vs. frequency shift. 51 channels with enhanced SRS, 10×100 km SMF link. Solid lines: SSFM simulations. Dashed lines: SRS-EGN model. Dotted: SRS-GN model QPSK||64QAM: WDM with left-half QPSK, right-half 64QAM.

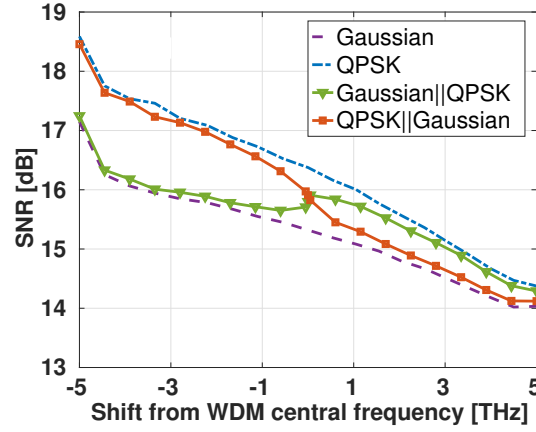
of 3 dBm and we artificially inflated the Raman gain by considering a slope for its linear approximation equal to $C_r = 0.14/(2\pi) (\text{THz} \cdot \text{rad} \cdot \text{km} \cdot \text{W})^{-1}$ to emphasize the SRS over the 2.5 THz bandwidth of the WDM, a value for which reliable SSFM simulations are feasible in a short time. The figure shows the NLI variance obtained with SSFM simulations (solid lines) and the SRS-EGN model (dashed lines) for a homogeneous modulation format scenario with PDM-64QAM and a heterogeneous scenario where the left-half of the spectrum, including the central channel, is PDM-QPSK modulated, while the right-half carries PDM-64QAM channels. It can be seen that the maximum error introduced by the EGN is ≈ 0.3 dB while the SRS-GN model (dotted) overestimates σ_{NLI}^2 by ≈ 1.4 dB. It is interesting to notice that the use of heterogeneous modulation formats affects the Raman-induced tilt in NLI variance.

Exploiting the SRS-EGN model, we analyzed the same link and set-up of Fig. 3.8 over the WDM bandwidth of ≈ 10 THz. Here we also accounted for the Raman tilt on the ASE noise arising from lumped optical amplifiers having noise figure 5 dB. We computed the SNR for both the homogeneous/heterogeneous modulation format se-

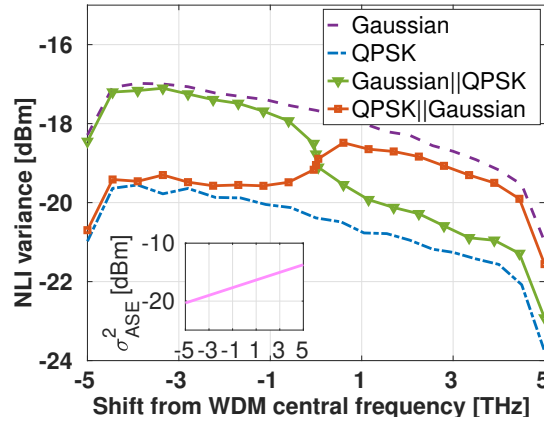
tups. As homogeneous cases, we considered PDM-QPSK or PDM-Gaussian modulation formats (i.e., complex Gaussian distributed symbols). For the heterogeneous scenarios, we considered the 101 frequency left-most channels modulated PDM-QPSK, while the remaining 100 right-most channels were PDM-Gaussian modulated. Figure 3.10(a) shows the SNR as a function of the frequency shift with respect to the WDM central frequency for the four scenarios.

First, it can be seen that the SNR is frequency-dependent although the signal power has been equalized by the DGE. This is a consequence of the fact that the noise variance is frequency-dependent. To better understand the reasons for the SNR behavior, in Fig. 3.10(b) we reported the corresponding variances of the linear and nonlinear noise. The main section of the figure represents the NLI variance while the ASE variance can be found in the inset. While the tilt in the NLI variance is due to distributed SRS and has the same sign of the tilt induced in the signal power before equalization, the non-flat profile of the ASE variance is due to the presence of the DGE and thus exhibits an opposite tilt.

Second, we note that, contrary to the GN model prediction, the homogeneous and the heterogeneous scenarios do not coincide. The SRS-EGN model can account for the impact of the modulation format on the NLI variance, both for homogeneous and heterogeneous cases, thus providing different SNR values for each scenario. Regarding the heterogeneous cases, we note that the two spectrum allocation strategies are not equivalent in terms of Raman-induced tilt in the SNR. While the asymptotic tilt far from the WDM center is of negative slope and follows the homogeneous case, in the center of the spectrum the slope of the tilt depends on the spectrum allocation strategy. In particular, when QPSK is on the left-hand side the tilt is positive and adds up with the tilt of ASE, while in the other case it counteracts the tilt of ASE. Focusing on the homogeneous cases (dashed and dotted lines) in Figs. 3.10(a)-(b), thanks to the Raman-aware EGN-model, we notice that while NLI variance experiences the same tilt for both modulation formats, the tilt in the SNR is different due to the different balance between NLI and ASE.



(a)



(b)

Figure 3.10: (a) SNR and (b) NLI variance, with ASE noise variance in the inset, vs. frequency shift. 201 channels, 10×100 km. QPSK||Gaussian: WDM with left-half QPSK, right-half Gaussian distributed symbols. Gaussian||QPSK: WDM with left-half Gaussian distributed symbols, right-half QPSK.

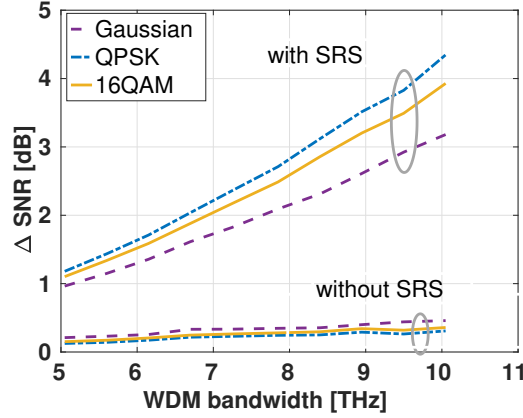


Figure 3.11: SNR gap between the two outer channels vs. the WDM bandwidth.

To better appreciate the SRS impact on the SNR, in Fig. 3.11 we plot the SNR gap ΔSNR between the two outer channels of the spectrum as a function of the WDM bandwidth. We observe that the QPSK scenario is the most affected by the Raman tilt and it represents the worst case in terms of GN model error. In particular, we observe 1 dB more SRS tilt at 10 THz compared to the Gaussian scenario. For the sake of comparison, in Fig. 3.11 we also reported the SNR gap due to the third-order dispersion induced tilt in the absence of Raman. We note that here the modulation format has a minor effect. We thus conclude that the error in the SNR estimation introduced by neglecting the modulation format, as the GN-model does, is not frequency-independent.

3.4 The DGE-SRS-EGN model

The SRS-EGN model presented in Section 3.3 is based on the power profile in Eq. (3.12) derived under the assumption that the SRS gain is compensated at each span-end through a DGE. Such an assumption can be limiting for estimating the performance of a practical terrestrial system, where such dynamic devices are placed more sparsely along the link. For instance, a compensation period of 4-5 spans can be found

in the literature [55, 59, 60].

The structure of a link with sparse DGE positioning is sketched in Fig. 3.4. In this scenario, the Raman gain-tilt is not recovered at every span. As a result, the SRS gain accumulates between two DGEs, and its expression can be generalized as in Eq. (3.9). Such an accumulation undermines the assumptions of weak SRS gain that lead to Eq. (3.17). Moreover, the absence of gain equalization at the end of each span allows other residual gains/losses to accumulate from span to span.

Similar scenarios were considered in [55], where the accumulation of Raman scattering was included by numerically solving the power evolution. Here we propose a simplified expression of the power profile.

3.4.1 Integral model

Let us focus on a given link section between two neighboring DGEs. The PSD of the signal at the input of the k th span, $k \in (1, \dots, N_s)$, should account for the accumulated SRS and any extra gain/loss per span $\Lambda(\omega)$:

$$\mathcal{G}_k(\omega) = \mathcal{G}(\omega) e^{-P_t C_r ((k-1)L_{\text{eff}}(L))\omega} \Lambda(\omega)^{(k-1)} \quad (3.20)$$

where, for the sake of simplicity, we assumed identical $\Lambda(\omega)$ per span. Please note that such an extra gain/loss reduces to 1 when a DGE is placed at the end of each span. The normalization factor should change accordingly:

$$\Upsilon_k(L_{\text{eff}}(\zeta)) \triangleq \frac{P_t}{\int_{-\infty}^{\infty} \mathcal{G}_k(\nu) e^{-P_t C_r L_{\text{eff}}(\zeta)\nu} \frac{d\nu}{2\pi}} \quad (3.21)$$

where ζ is the local coordinate within the k th span. In this framework, the normalized power profile between two neighboring DGEs can be written as

$$\rho_k(\zeta, \omega) = \underbrace{\Upsilon_k(L_{\text{eff}}(\zeta))}_{\text{normalization factor}} \underbrace{\Lambda(\omega)^{(k-1)}}_{\text{acc. loss}} \underbrace{e^{-\alpha\zeta}}_{\text{local loss}} \underbrace{e^{-P_t C_r (k-1)L_{\text{eff}}(L)\omega}}_{\text{acc. SRS}} \underbrace{e^{-P_t C_r L_{\text{eff}}(\zeta)\omega}}_{\text{local SRS}} \quad (3.22)$$

where the different contributions to the extended power profile are highlighted.

Contrary to the previous cases, the power profile in Eq. (3.22) is span-dependent. It is worth noting that Eq. (3.22) does not hold with a frequency-dependent attenuation $\alpha(\omega)$, since the signal power profile has been derived with a frequency-independent loss [69] in the triangular approximation of SRS. As an approximation, Eq. (3.22) can be used as well by forcing a constant $\alpha(\omega) \equiv \alpha$ only in the attenuation profile $e^{-\alpha\zeta}$ at local coordinate ζ , while its full profile $\alpha(\omega)$ can be left in the accumulated gain/loss term $\Lambda(\omega)$.

The substitution of the new power profile Eq. (3.22) in Eq. (3.15) yields the link kernel expression

$$\eta = \sum_{d=1}^{N_D} e^{j\Delta\beta(d-1)LN_s(d)} \sum_{k=1}^{N_s(d)} e^{j\Delta\beta(k-1)L} g(\omega, \omega_1, \omega_2)^{\frac{k-1}{2}} \times e^{-P_l C_r(k-1)L_{\text{eff}}(L)(\omega+\omega_1+\omega_2)} \eta_k^R(\omega, \omega_1, \omega_2) \quad (3.23)$$

where d is the index of the link section between two DGEs, $N_s(d)$ is the number of spans in the d th link section, k is the span index within a section, and the term $g(\omega, \omega_1, \omega_2)$ collecting extra gains/losses is defined as

$$g(\omega, \omega_1, \omega_2) = \frac{\Lambda(\omega + \omega_1)\Lambda(\omega + \omega_2)\Lambda(\omega + \omega_1 + \omega_2)}{\Lambda(\omega)}. \quad (3.24)$$

The single-span kernel generalizes to:

$$\eta_k^R = \int_0^L \Upsilon_k(L_{\text{eff}}(\zeta)) e^{-\alpha\zeta} e^{-P_l C_r L_{\text{eff}}(\zeta)(\omega+\omega_1+\omega_2)} e^{j\Delta\beta\zeta} d\zeta \quad (3.25)$$

with the main difference with Eq. (3.17) that the normalization factor depends on the span index k within the link section, as per Eq. (3.21).

Since the exponential SRS in Υ includes also the accumulated gain, the weak SRS assumption used in [5] yielding $\Upsilon \approx 1$ may not be justified in this more general scenario with sparse DGEs. As a direct consequence, the z -integral in Eq. (3.25) cannot be expressed in closed-form and requires to be evaluated through numerical integration. For this reason, in this work, we refer to this model as the *integral model*.

Contrary to Section 2.1.1 and 3.3 it is not possible to simply identify a phased-array term accounting for the accumulated effects between spans, since the spans are

not all identical due to the absence of a DGE at some spans. The outer summation over d in Eq. (3.23) accounts for the dispersion accumulated in the previous link sections. Note that there is no accumulated SRS from previous link sections since each section ends with an ideal DGE. The inner summation over k , on the other hand, accounts for both dispersion and SRS accumulated within the section.

We note that the proposed model can be easily generalized to account for a non-uniform power allocation. For instance, a power pre-emphasis might be applied to the transmitted signal to counteract the SRS effect [55] [56] [59]. We find it convenient to investigate a pre-emphasis in the form of an SRS gain with an opposite sign, namely

$$P_i = H_{-\bar{k}}(\omega_i, L)P. \quad (3.26)$$

Using this notation, the pre-emphasis is governed by the factor \bar{k} which indicates the amount of SRS that is pre-compensated, expressed in an equivalent number of spans. As a consequence, the link kernel in Eq. (3.23) can be applied verbatim after the following substitution in the Raman exponential:

$$(k-1) \xrightarrow[\text{pre-emphasis}]{\text{with}} (k-\bar{k}-1) \quad (3.27)$$

where $k-\bar{k}$ plays the role of an equivalent span index.

3.4.2 Simplified model

The numerical effort to evaluate the general kernel in Eq. (3.23) with sparse DGEs is particularly heavy since the integrand function is quickly oscillating, thus requiring many function evaluations for an accurate result. For this reason, in this section we introduce approximations of the z -integral in Eq. (3.23), with big savings in computational time. For the sake of simplicity and generality, we search for an approximation yielding a link kernel in the same form as the one without SRS accumulation in Eq. (3.16), i.e.,

$$\eta(\omega, \omega_1, \omega_2) \approx \chi^R(\omega, \omega_1, \omega_2) \eta_1^R(\omega, \omega_1, \omega_2) \quad (3.28)$$

where η_1^R is the single-span kernel in Eq. (3.17) and χ^R is an inter-span term. Note that Eq. (3.28) differs from the model without accumulated SRS only in the term χ^R .

Removing the dependence of Υ on z yields great numerical savings since it results in an integral along the distance that can be evaluated in closed-form. This idea was implicit in [5] where Υ was approximated to 1, as a result of a weak-SRS assumption. As commented before, this approximation does no longer hold with sparse DGE positioning, hence it must be properly adapted to the new scenario. Since Υ is a monotonic function of $L_{\text{eff}}(\zeta)$, it can be bounded by $\Upsilon(L_{\text{eff}}(L)) \leq \Upsilon(L_{\text{eff}}(\zeta)) \leq \Upsilon(0)$. Hence, it seems reasonable to approximate the integral in Eq. (3.25) by a weighted midpoint numerical quadrature [71]. To further increase the accuracy, we perform such an approximation after making the change of variable $L_{\text{eff}}(\zeta) = y$, namely:

$$\eta_k^R = \int_0^{L_{\text{eff}}(L)} \Upsilon_k(y) e^{-P_{\text{Cr}} y (\omega + \omega_1 + \omega_2)} e^{-j \frac{\Delta\beta}{\alpha} \ln(1 - \alpha y)} dy \quad (3.29)$$

$$\triangleq \int_0^{L_{\text{eff}}(L)} \Upsilon_k(y) w(y) dy \quad (3.30)$$

$$\approx \Upsilon_k\left(\frac{L_{\text{eff}}(L)}{2}\right) \int_0^{L_{\text{eff}}(L)} w(y) dy \quad (3.31)$$

where the exponential functions have been used as the weighting function $w(y)$, and in the last step, according to the midpoint rule [71], the integrand function Υ is factored out and sampled at its midpoint. We note that integrating the weighting function corresponds to carrying out the integration of η_k^R as if $\Upsilon = 1$. Therefore, assuming now a local weak-SRS regime for the integral computation is legit, and we get:

$$\eta_k^R \approx \Upsilon_k\left(\frac{L_{\text{eff}}(L)}{2}\right) \eta_1^R(\omega, \omega_1, \omega_2) \quad (3.32)$$

where $\eta_1^R(\omega, \omega_1, \omega_2)$ is the *closed-form* expression of the single-span kernel without accumulated SRS in Eq. (3.17). The midpoint rule has an error scaling with $\mathcal{O}(L_{\text{eff}}(L)^3)$ [71] hence is a good compromise between accuracy and simplicity. As a result of the approximation, we thus propose to approximate the local term $L_{\text{eff}}(\zeta)$ in Eq. (3.21) with a lumped factor $\frac{1}{2}L_{\text{eff}}(L)$, obtaining the following midpoint normalization factor

$$\Upsilon_k \triangleq \frac{P_t}{\int_{-\infty}^{\infty} \mathcal{G}(v) \Lambda(v)^{(k-1)} e^{-P_t C_r (k-\frac{1}{2}) L_{\text{eff}}(L) v} \frac{dv}{2\pi}}. \quad (3.33)$$

Thanks to this approximation, the link kernel in Eq. (3.23) can be written as

$$\begin{aligned} \eta &= \sum_{d=1}^{N_D} e^{j\Delta\beta(d-1)LN_s(d)} \sum_{k=1}^{N_s(d)} e^{j\Delta\beta(k-1)L} g(\omega, \omega_1, \omega_2)^{\frac{k-1}{2}} \\ &\times e^{-P_t C_r (k-1) L_{\text{eff}}(L)(\omega+\omega_1+\omega_2)} \Upsilon_k \eta_1^R(\omega, \omega_1, \omega_2) \end{aligned} \quad (3.34)$$

which can now be easily expressed in the compact form of Eq. (3.28), where we defined

$$\begin{aligned} \chi^R &\triangleq \sum_{d=1}^{N_D} e^{j\Delta\beta(d-1)LN_s(d)} \sum_{k=1}^{N_s(d)} e^{j\Delta\beta(k-1)L} g(\omega, \omega_1, \omega_2)^{\frac{k-1}{2}} \\ &\times e^{-P_t C_r (k-1) L_{\text{eff}}(L)(\omega+\omega_1+\omega_2)} \Upsilon_k. \end{aligned} \quad (3.35)$$

We call the *simplified model* the NLI model based on the link kernel of Eq. (3.34) which relies on the approximated normalization factor defined in Eq. (3.33).

3.4.3 GN-term closed-form expressions

In this section, we generalize the closed-form expressions of the SPM and XPM variance, first derived in [5] in a GN-framework with DGE at every span, to the case of accumulated SRS, i.e., sparse DGE positioning along the link.

For the reader's convenience, we report here the single-span variance expressions of [5]. The SPM variance for channel i can be written as

$$\begin{aligned} \sigma_{\text{SPM},1}^2(i) &\approx \frac{4}{9} \frac{P_i^3}{B_i^2} \frac{\gamma^2 \pi}{\phi_i 3\alpha^2} \\ &\times \left[\frac{T_i - \alpha^2}{\alpha} \text{asinh}\left(\frac{\phi_i B_i^2}{\pi \alpha}\right) + \frac{4\alpha^2 - T_i}{2\alpha} \text{asinh}\left(\frac{\phi_i B_i^2}{2\pi \alpha}\right) \right] \end{aligned} \quad (3.36)$$

with $\phi_i = \frac{3}{2}\pi^2(\beta_2 + \beta_3\omega_i)$, $T_i = (2\alpha - P_t C_r \omega_i)^2$, and $\omega_i \triangleq 2\pi f_i$, with f_i the low-pass carrier frequency of channel i , referred, without loss of generality, to the reference

system centered at the WDM central frequency [5]. The XPM contribution of the ℓ th interfering channel to channel i is

$$\sigma_{\text{XPM},1}^2(i,\ell) \approx \frac{32}{27} \frac{P_\ell^2 P_i}{B_\ell} \frac{\gamma^2}{\phi_{i,\ell} 3\alpha^2} \times \left[\frac{T_\ell - \alpha^2}{\alpha} \text{atan}\left(\frac{\phi_{i,\ell} B_i}{\alpha}\right) + \frac{4\alpha^2 - T_\ell}{2\alpha} \text{atan}\left(\frac{\phi_{i,\ell} B_i}{2\alpha}\right) \right] \quad (3.37)$$

with B_i the bandwidth of channel i , and $\phi_{i,\ell} = \pi(\omega_\ell - \omega_i)[\beta_2 + \frac{1}{2}\beta_3(\omega_\ell + \omega_i)]$. The approximation is reliable if $|\omega_i - \omega_\ell| \gg 2\pi B_i$ [5].

As in the case of the absence of SRS, the NLI accumulation of channel i along the link can be approximated through the simple scaling rules

$$\sigma_{\text{SPM},N}^2(i) = \mathcal{N}_i^{1+\varepsilon} \sigma_{\text{SPM},1}^2(i) \quad (3.38)$$

$$\sigma_{\text{XPM},N}^2(i) = \sum_{\ell \neq i} \mathcal{N}_\ell \sigma_{\text{XPM},1}^2(i,\ell) \quad (3.39)$$

where the scaling factor is defined as $\mathcal{N}_\ell \triangleq |\chi(0,0,\omega_\ell)|_{\text{inc}}^2$ and the subscript inc indicates that only the incoherent contribution is taken into account, i.e., correlations between different spans are neglected.

In the absence of accumulated SRS, the coherent and incoherent contributions [3] to the modulus square of χ can be derived as

$$\begin{aligned} |\chi|^2 &= \sum_{k=1}^N e^{j\Delta\beta L(k-1)} \sum_{s=1}^N e^{-j\Delta\beta L(s-1)} = \sum_{k=0}^{N-1} e^{j\Delta\beta Lk} \sum_{s=0}^{N-1} e^{-j\Delta\beta Ls} \\ &= 1 + \sum_{k \neq 0} e^{j\Delta\beta Lk} + \dots + 1 + \sum_{k \neq N-1} e^{j\Delta\beta Lk} e^{-(N-1)j\Delta\beta L} \\ &= N + \sum_{k \neq 0} e^{j\Delta\beta Lk} + \dots + \sum_{k \neq N-1} e^{j\Delta\beta L(k-(N-1))} \\ &= \underbrace{N}_{\text{incoherent}} + 2 \underbrace{\sum_{n=1}^{N-1} (N-n) \cos(\Delta\beta L \times n)}_{\text{coherent}} \end{aligned} \quad (3.40)$$

hence the incoherent contribution amounts to $\mathcal{N} = N$, consistently with Eq. (2.41).

As done in Chapter 2, the coherent NLI accumulation along spans has been included through the coherence factor ε . Contrary to the previous case, we now included the coherence factor only in the SPM accumulation in Eq. (3.38), as done in [5]. We thus assume incoherent XPM motivated by the large walk-off experienced in a typical DU link. Hence, the coherence factor in Eq. (2.42) is now modified replacing the WDM bandwidth B_{WDM} with the channel bandwidth B .

Note that the scaling rules in Eq. (3.38) and (3.39) require the accumulated effects to be factored out by the term χ . Unfortunately, the proposed integral model cannot be cast in this form. On the other hand, the simplified model presented in this section can be expressed in such a form as in Eq. (3.28). Therefore, the simplified model opens the path to closed-form expressions in the presence of accumulated SRS.

For a single-span, η_1^{R} in the simplified model coincides with the closed-form expression of the single-span z -integral derived in [5], therefore the single-span SPM and XPM variance expressions in Eq. (3.36) and Eq. (3.37) still hold in this framework. On the other hand, in a multi-span link the term \mathcal{N} must be modified to account for sparse DGE positioning along the distance.

For the sake of simplicity, let us focus on a link section and assume that there are no extra losses ($\Lambda(\omega)=1$). After sampling the factor χ^{R} at central channel frequency ω_ℓ of the interfering channel and defining $b \triangleq P_{\text{t}} C_{\text{t}} L_{\text{eff}} \omega_\ell$, Eq. (3.40) can be generalized in

$$\begin{aligned}
 |\chi^{\text{R}}|^2 &= \sum_{k=1}^{N_s} \Upsilon_k e^{-b(k-1)} e^{j\Delta\beta L(k-1)} \sum_{s=1}^{N_s} \Upsilon_s e^{-b(s-1)} e^{-j\Delta\beta L(s-1)} \\
 &= \Upsilon_1^2 + \sum_{k \neq 1} \Upsilon_k \Upsilon_1 e^{-b(k-1)} e^{j\Delta\beta L(k-1)} + \dots + \Upsilon_{N_s}^2 e^{-2(N_s-1)b} \\
 &\quad + \sum_{k \neq N_s} \Upsilon_k \Upsilon_{N_s} e^{j\Delta\beta L(k-N_s)} e^{-b(k+N_s-2)} \\
 &= \underbrace{\sum_{r=1}^{N_s} \Upsilon_r^2 e^{-2b(r-1)}}_{\text{incoherent}} + \underbrace{\sum_{s=1}^{N_s} \sum_{k \neq s} \Upsilon_k \Upsilon_s e^{-b(k+s-2)} e^{j\Delta\beta L(k-s)}}_{\text{coherent}}.
 \end{aligned} \tag{3.41}$$

Case	Scaling factor \mathcal{N}_ℓ
$N_s = 1$	N
$N_s > 1$, Simplified model	$\sum_{d=1}^{N_D} \sum_{k=1}^{N_s(d)} \Upsilon_k^2 \Lambda(\omega_\ell)^{2(k-1)} e^{-2P_t C_t L_{\text{eff}}(L) \omega_\ell (k-1)}$
$N_s > 1$, Integral model	not available

Table 3.2: Scaling factor \mathcal{N}_ℓ for NLI accumulation with spans. $N_s = 1$: case analyzed in [5]. $N_s > 1$: novel extension with sparse DGE positioning along the link.

It is worth noting that sampling the factor χ^R at the central channel frequencies implies the assumption of frequency-flat SRS gain within a channel bandwidth, which is consistent with the assumptions underlying the single-span variance expressions in Eq.(3.36) and Eq. (3.37).

The values of the scaling factor \mathcal{N} are summarized in Tab. 3.2, where we included also the additional gain/loss term $\Lambda(\omega_\ell)$ and the external summation on the link sections. Note that the case $N_s = 1$, for which the scaling factor simply reduces to the number of spans N , encompasses both the case of absence of local SRS, as in [3], and the absence of accumulated SRS, as in [5]. With $N_s > 1$ and the integral model, the scaling factor \mathcal{N} does not exist, hence the single-span contribution cannot be scaled for the given number of spans. In such a case, a full numerical spatial integration over the entire link length is necessary.

3.4.4 Numerical validation

We implemented both the integral and the approximated model by adopting the link kernel expressions in Eq. (3.23) and Eq. (3.34), respectively, in the EGN model. Throughout this section, we will refer to the respective EGN models with the labels "integral DGE-SRS-EGN" and "simplified DGE-SRS-EGN" model. In particular, in both models, the frequency integrals involved in the variance computation are per-

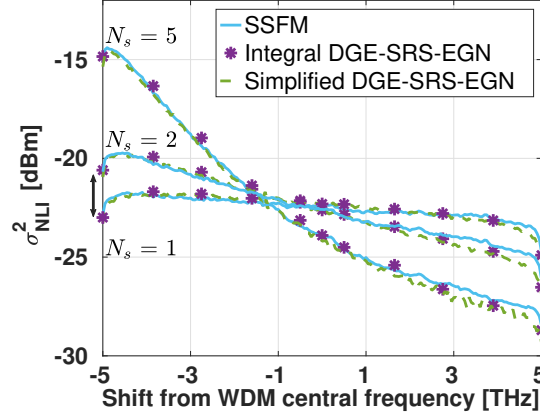


Figure 3.12: NLI variance σ_{NLI}^2 vs. frequency shift. PDM-64QAM, 201 channels, 10×100 km with DGE period $N_s = 1, 2$ or 5 spans. Solid line: SSFM simulations. Markers: Integral DGE-SRS-EGN. Dashed line: simplified DGE-SRS-EGN model.

formed through Monte Carlo integration, and all the nonlinear effects (SPM, XPM, FWM, XCI, and MCI) are included.

Due to the oscillatory nature of the fiber kernel, the z -integration of the integral model is implemented through the Filon's method [72]. Such a method is a quadrature technique for highly oscillating function integrals in the form $\int_a^b f(x)e^{j\omega g(x)}dx$, hence suitable for Eq. (3.23). On the other hand, the simplified DGE-SRS-EGN model relied on the closed-form expression of the single-span kernel, as well as the GN-term closed-form expressions.

Figure 3.12 shows the NLI variance versus the frequency shift after 10 spans with DGE placed every 1, 2, or 5 spans. All the transmission and link parameters, except for the DGE period, are the same used for Fig. 3.7. We used solid lines for SSFM results, markers for the integral DGE-SRS-EGN model, and dashed lines for the simplified DGE-SRS-EGN model. It can be seen that both the integral and the simplified EGN models proposed in this work correctly estimate the impact of accumulated SRS on the NLI. Both models exhibit an average error across the WDM bandwidth smaller than 0.3 dB for each DGE placement configuration, with the largest value

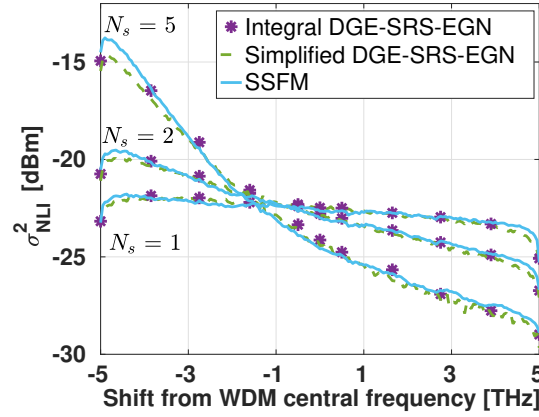


Figure 3.13: NLI variance σ_{NLI}^2 vs. frequency shift. PDM-16QAM, 201 channels, 10×100 km with DGE period $N_s=1,2$ or 5 spans. Solid line: SSFM simulations. Markers: Integral DGE-SRS-EGN. Dashed line: simplified DGE-SRS-EGN model.

concentrated at the bandwidth edges. In particular, we verified that the error of the integral DGE-SRS-EGN model is mainly related to the underlying triangular approximation of the Raman gain.

Most important, it can be seen that the NLI tilt is emphasized by the accumulation of SRS between DGEs, yielding curves with $N_s > 1$ far apart from the benchmark $N_s = 1$ case usually analyzed in the literature. For instance, even with a DGE every two spans, i.e., $N_s = 2$, the gap with the curve having a DGE at every span is ≈ 2 dB.

Then, we verified the accuracy of the proposed models for a 16QAM modulation, for the same set-up of Fig. 3.12. Figure 3.13 shows a good agreement between theory and SSFM simulations, with similar accuracy as that of the 64QAM case.

It is worth noting that simulating the transmission of a WDM bandwidth of 10 THz over 1000 km via SSFM had a nearly prohibitive cost in terms of computational time, requiring at least 21 days for each curve in Fig. 3.12 and 3.13 using graphical process units (GPUs).

As an example of the complexity reduction enabled by the simplified model, we report the computational times required to compute the NLI variance with the two mod-

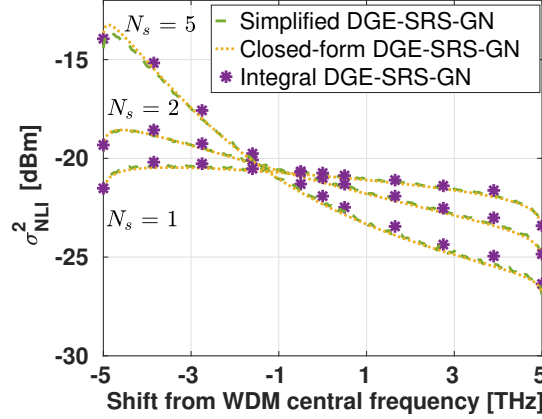


Figure 3.14: NLI variance σ_{NLI}^2 vs. frequency shift. Gaussian distributed symbols, 201 channels, 10×100 km with DGE period $N_s=1,2$ or 5 spans. Markers: Integral DGE-SRS-GN. Dashed line: simplified DGE-SRS-GN. Dotted: extended-closed forms.

els. Using a server-grade architecture, the integral model required ≈ 50 min/channel to compute the NLI variance of Fig. 3.12. The simplified model required less than 1 min/channel by using the same number of Monte Carlo samples for frequency integration.

The computational time can be further reduced by replacing the Monte Carlo integrations in the frequency domain with the closed-form expressions introduced in Section 3.4.3. Although such expressions can be computed in a few seconds, it is worth noting that they i) account only for SPM and XPM, ii) postulate a scaling of SPM with the coherence factor ε , iii) use the same ε as the case without SRS, and iv) assume Gaussian distributed symbols. We thus investigated the reliability of such expressions with Gaussian distributed symbols, while all the other link and transmission parameters were those of Fig. 3.12. The corresponding NLI variance, as well as its estimation by the integral and the simplified model, are reported in Fig. 3.14. The dotted lines represent the closed-form results, the solid lines indicate the simplified Monte Carlo model while the markers represent the σ_{NLI}^2 obtained with

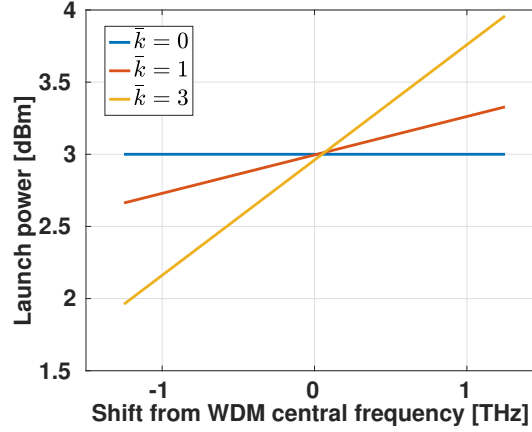


Figure 3.15: Launch power for variable pre-emphasis factor \bar{k} , see Eq. (3.26), vs. frequency shift.

the integral model. The figure shows that the average gap between the closed-form and its Monte Carlo counterpart is less than 0.1 dB for all the DGE periods, making the closed-form expression a fast reliable alternative for estimations with Gaussian distributed symbols. The closed-form expression accuracy is expected to decrease when low-dispersion fibers or low symbol rates are taken into account, due to the higher relative contribution of FWM.

Pre-emphasis

We checked the validity of the model with transmitted signal power pre-emphasis as per Eq. (3.26). Since SSFM simulations of 10 THz over 1000 km are extremely time-consuming, we focused on a faster set-up by reducing the bandwidth and the link length. We thus transmitted 51 channels, with the same frequency spacing and symbol rate as in the previous figures, with channel power $P = 3$ dBm.

To test the impact of the pre-emphasis only, we estimated the NLI variance after 300 km, without intermediate DGEs. We considered both SMFs and NZDSFs, with dispersion parameter $D = 17$ ps/(nm·km) and $D = 4.5$ ps/(nm·km), respectively. For the sake of comparison, the NZDSF under test differed from the SMF considered so

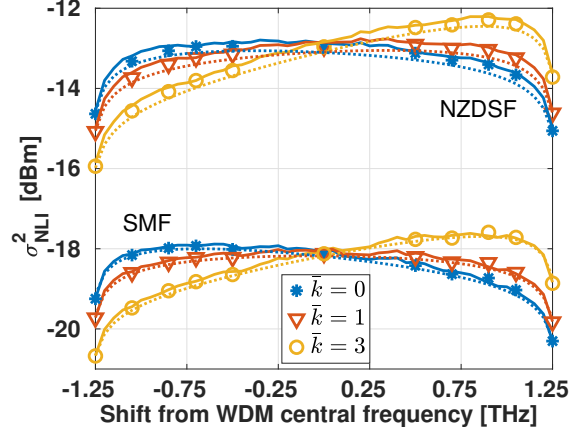


Figure 3.16: NLI variance σ_{NLI}^2 vs. frequency shift. Gaussian distributed symbols, 51 channels, 3×100 km of SMFs or NZDSFs, without inline DGEs. Variable pre-emphasis factor \bar{k} , see Eq. (3.26). Solid: SSFM. Markers: Integral DGE-SRS-GN model. Dotted: extended-closed forms by using Eq. (3.26).

far only in the dispersion coefficient. For both cases, we report the NLI variance in the absence of SRS pre-compensation ($\bar{k} = 0$) and for two pre-emphasis choices $\bar{k} = 1, 3$. The launch power in Eq. (3.26) for the three cases is reported in Fig. 3.15. We recall that, according to Eq. (3.27), a pre-emphasis factor of \bar{k} yields a frequency-flat signal power after \bar{k} spans.

Figure 3.16 shows the NLI variance estimated with SSFM simulations, the integral DGE-SRS-EGN model and the closed-form expressions modified according to Eq. (3.27). It can be seen that the simulations are in good agreement with the theory. In particular, the modified closed-forms exhibit a maximum error of ≈ 0.25 dB with SMFs, and an average error across the WDM bandwidth smaller than 0.1 dB. The maximum gap slightly increases to 0.3 dB with NZDSFs, while on average it is 0.2 dB. This can be mainly attributed to the absence of FWM, XCI, and MCI terms in the closed-form expressions, whose relative contribution to the overall NLI increase at lower dispersion.

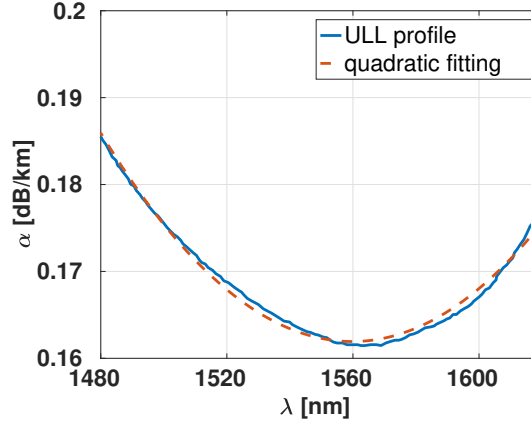


Figure 3.17: Attenuation coefficient of a Corning ULL fiber vs. wavelength. Solid line: extracted from [70]. Dashed: quadratic polynomial interpolation.

Frequency-dependent attenuation

In this section, we check the validity of Eq. (3.22) with a frequency-dependent fiber attenuation. To this aim, we included a frequency-dependent attenuation profile in the SSFM simulations. Here we changed the fiber, a Corning ULL, for which an experimental attenuation profile is available in [70]. We approximated such a profile with a quadratic fitting polynomial in a dB scale:

$$\alpha \text{ [dB/km]} \approx \alpha_2(\lambda - \lambda_0)^2 + \alpha_1(\lambda - \lambda_0) + \alpha_0 \quad (3.42)$$

with $\lambda_0 = 1550$ nm and coefficients: $\alpha_2 = 3.7685 \times 10^{-6}$, $\alpha_1 = -7.3764 \times 10^{-5}$, and $\alpha_0 = 0.162$. The attenuation profile and its quadratic fitting are shown in Fig. 3.17.

The link under test was composed of $N = 5$ spans with DGE repetition every 1 and 5 spans. Each amplifier had a frequency-independent gain restoring the nominal total power, and the frequency-dependent residual loss was equalized by the DGE.

Figure 3.18 shows the NLI variance estimated with SSFM simulations, the integral and the simplified DGE-SRS-EGN model with local fiber attenuation coefficient

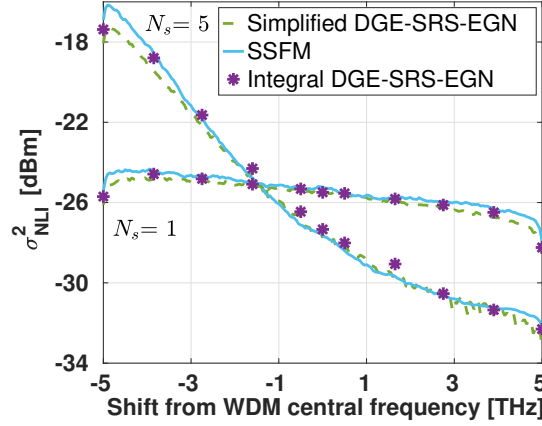


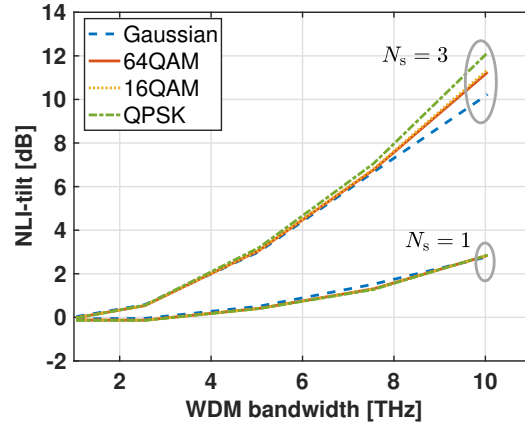
Figure 3.18: NLI variance σ_{NLI}^2 vs. frequency shift. PDM-64QAM, 201 channels, 5×100 km of ULL fiber with DGE period $N_s=1,5$ spans. Solid line: SSFM simulations with $\alpha(\omega)$. Dashed line: simplified DGE-SRS-EGN. Markers: integral DGE-SRS-EGN.

in Eq. (3.22) equal to $\alpha = \alpha_0$ dB/km. The average error across the WDM bandwidth between simulations and theory is less than 0.3 dB for $N_s = 1$, meaning that neglecting the local effects of the frequency-dependent fiber attenuation as in Eq. (3.22) has a minor impact on the NLI, consistently with the discussion in [57]. When the DGE period is increased to $N_s = 5$ spans, the average error is ≈ 0.3 dB.

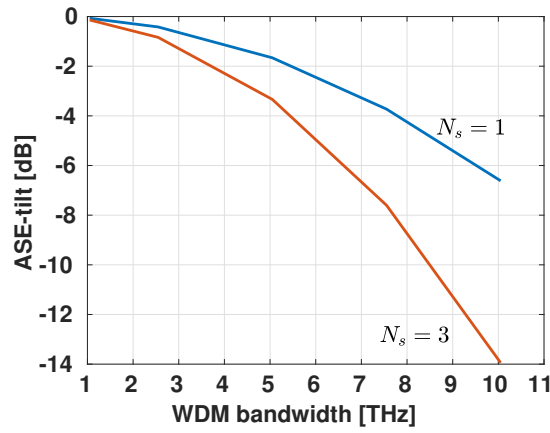
In particular, we note that the impact of a sparse DGE positioning on the NLI variance is barely affected by the frequency-dependent fiber attenuation. Figure 3.18 shows a maximum gap of ≈ 8 dB between the curves with $N_s = 1$ and 5 spans, which is comparable to the gap estimated with a frequency-flat fiber attenuation.

3.4.5 Modulation format

We next investigate the joint impact of the modulation format and the accumulated SRS on the NLI variance using the simplified model. We considered the transmission of a $B_t = 0.05 \times N_{\text{ch}}$ THz WDM comb with variable bandwidth $B_t = (1, 2.5, 5, 7.5, 10)$ THz and $N_{\text{ch}} = (21, 51, 101, 151, 201)$ channels, respectively. The channel power was



(a)



(b)

Figure 3.19: For a 30×100 km link with DGE repetition every $N_s=1$ and $N_s=3$ spans: (a) NLI variance tilt in dB due to SRS (between the two edge WDM channels) for QPSK (dash-dotted), 16QAM (dotted), 64QAM (solid) and Gaussian (dashed) transmissions vs. WDM bandwidth; (b) ASE variance tilt in dB versus WDM bandwidth.

fixed to 1 dBm.

We considered different modulation formats: Gaussian, 64QAM, 16QAM, and QPSK. The link under test was composed of 30 spans of SMFs with DGE placed every 1 or 3 spans. For each scenario, we measured the tilt of the variance as the difference between the value in dB over the first and the last WDM channel in frequency.

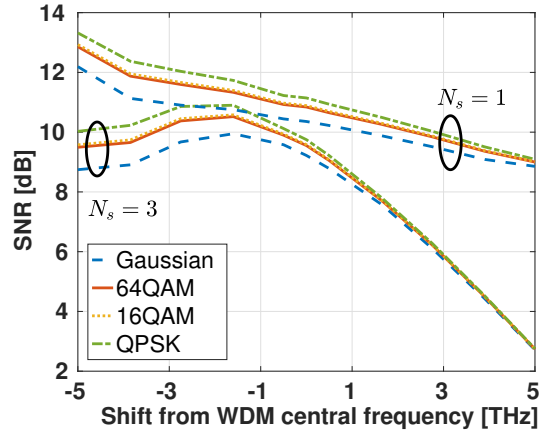
Figure 3.19(a) shows the SRS-induced NLI variance tilt as a function of the WDM bandwidth. It can be seen that with a DGE after each span, the NLI-tilt is almost independent of the modulation format. This observation is consistent with the results reported in the literature [61] and in Fig. 3.8. On the other hand, when the DGE period is increased to 3 spans, the tilt increases significantly as the bandwidth increases. For the extreme case of $B_t = 10$ THz, we estimated ≈ 2 dB of difference between the tilt of PDM-QPSK and PDM-Gaussian transmission.

For the sake of completeness, we report in Figure 3.19(b) the ASE variance tilt for the same link, with a frequency-flat EDFA noise figure of 5 dB. Note that the ASE variance undergoes an SRS-induced tilt with an opposite sign compared to the NLI variance, yet not identical in absolute value.

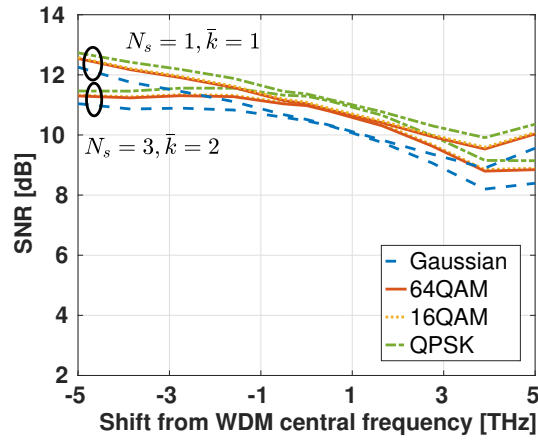
We then estimated the received SNR with both ASE and NLI. Figure 3.20(a) shows the SNR versus the frequency offset from the central channel at $B_t = 10$ THz, with DGE placed every 1 and 3 spans. It can be seen that the SRS-induced tilt on the SNR with $N_s = 1$ does depend on the modulation format although the NLI-tilt is almost format-independent, due to the different balance between the NLI and ASE variances. Therefore, the error introduced in the SNR estimation by neglecting the modulation format is frequency-dependent.

Figure 3.20(a) also shows that the SNR non-flatness in frequency is enhanced by the presence of accumulated SRS. In particular, the SNR deviation across the WDM bandwidth increases by ≈ 4 dB when the number of spans between the equalizers increases from 1 to 3.

It is worth noting that such tilts on the ASE and NLI have different implications on the system design. For instance, a 1 dB variation in the NLI variance, ASE variance, or SNR, corresponds approximately to a deviation of 1/3, 2/3, and 1 dB,



(a)



(b)

Figure 3.20: SNR vs. frequency offset from central channel at WDM bandwidth 10 THz for: (a) same setup of Fig. 3.19; (b) with power pre-emphasis factor \bar{k} in Eq. (3.26) optimized to reduce the SNR variations in frequency.

respectively, in the system reach [9, p. 324]. The same scaling laws apply even for the errors introduced by the models.

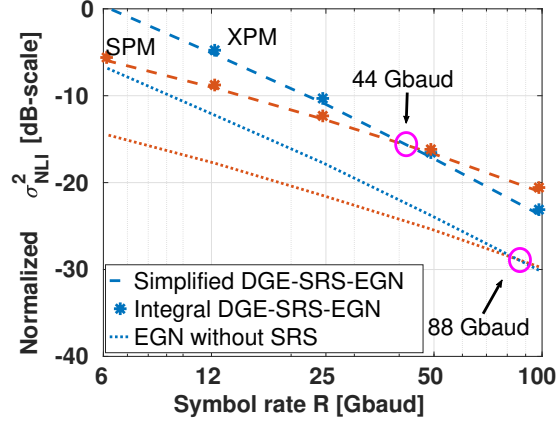
To counteract this undesired imbalance in the SNR and achieve similar performance on all the WDM channels, a power pre-emphasis can be applied at the transmitter side [55, 56, 59, 73, 74]. Figure 3.20(b) shows the SNR estimation versus frequency, for the same setup of Fig. 3.20(a), with signal power pre-emphasis as per Eq. (3.26). The pre-emphasis factor \bar{k} was optimized to minimize the maximum SNR difference between two arbitrary channels at fixed total power. Namely, the pre-emphasis factor was $\bar{k} = 1$ for $N_s = 1$ and $\bar{k} = 2$ for $N_s = 3$. Please note that a fast optimization of the pre-emphasis factor is possible thanks to the proposed simplified DGE-SRS-EGN model.

In the absence of signal power pre-emphasis, in Fig. 3.20(a) we measured up to 8 dB of SNR imbalance for the case having $N_s = 3$. On the other hand, Fig. 3.20(a) shows that the maximum SNR deviation in frequency can be reduced to ≈ 1.5 dB with optimized pre-emphasis factor \bar{k} .

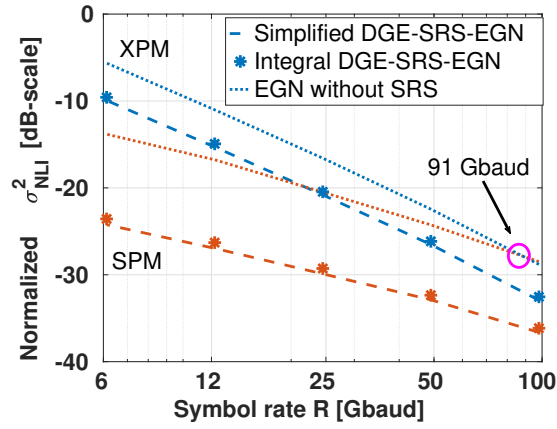
Although the pre-emphasis through \bar{k} is sub-optimal and counteracts only the SRS tilt, the simplicity of the generalized simplified model and its approximated closed-form formulas is remarkable. Such models can be used for fast predictions of system performance closer to reality than existing models unaware of pre-emphasis. More accurate predictions with arbitrary pre-emphasis can be obtained by using the integral model.

3.4.6 Self- and cross-channel effects

Finally, we investigate the accuracy of the simplified DGE-SRS-EGN model on the estimation of SPM and XPM, using the integral DGE-SRS-EGN model as a benchmark. We perform the exploration at variable symbol rates, thus extending the main findings of [49] to SRS. To this aim, we considered the transmission of a 16QAM WDM signal having $B_t = 10$ THz and $N_{\text{ch}} = (1601, 801, 401, 201, 101)$ channels with symbol rate $R = (6.125, 12.25, 24.5, 49, 98)$ Gbaud. The relative channel spacing $\Delta f/R$ was fixed to 1.02 and the total WDM power was $P_t = 23$ dBm. The link under test was composed of 20 spans of SMFs with DGE placed every 4 spans.



(a)



(b)

Figure 3.21: NLI variance σ_{NLI}^2 of SPM and XPM, normalized to the cube of the channel power, vs. the symbol rate, estimated for the (a) first, and (b) last WDM channel. PDM-16QAM, $B_t = 10$ THz, $P_t = 23$ dBm, 20×100 km of SMFs with DGE period $N_s = 4$ spans.

The model predictions are depicted in Fig. 3.21. Figure 3.21(a) and (b) show the variance of the first and last WDM channel in frequency, respectively, normalized to the cube of the channel power. We focused on the first and last WDM channels since they experience the largest SRS, thus representing a worst-case scenario for the underlying weak-SRS assumption adopted by the simplified model. FWM is not reported since its variance was estimated to be 6 dB smaller than XPM at 6 Gbaud, with an increasing gap for increasing R . In both figures, it can be seen that the difference between the simplified (dashed) and the integral model (markers) is negligible over the entire symbol rate axis for both nonlinear effects.

In the same figure, we also show with dotted lines the same curves in the absence of SRS. The figure shows that the presence of SRS changes the balance between the nonlinear effects in a frequency-dependent manner. For instance, the SRS-unaware model predicts that the crossing point between SPM and XPM on the first WDM channel occurs at $R = 88$ Gbaud. SRS moves such a point to 44 Gbaud by inflating more SPM than XPM.

An opposite behavior occurs for the last WDM channel, as shown in Fig. 3.21(b). In the absence of SRS, the crossing point is $R = 91$ Gbaud, slightly different from 3.21(a) because of third-order dispersion. On the contrary, when SRS is taken into account, XPM remains the main nonlinear impairment, being the crossing point with SPM at $R = 510$ Gbaud.

3.5 Optimization of signal power pre-emphasis

In this section, we exploit the DGE-SRS-GN model closed-form expression to optimize the signal power allocation to maximize either (i) the minimum SNR in the WDM comb or (ii) the achievable information rate. Previous results from the literature [55, 59] confirmed the intuition that more power should be allocated to those channels that are most depleted by the SRS. In particular, optimal power levels vs frequency exhibits an almost-linear behavior in the dB scale over the C+L band. However, the optimization problem over ultra-wide bandwidths without any constraints on the channel power levels can have high complexity. For instance, in [59] the authors

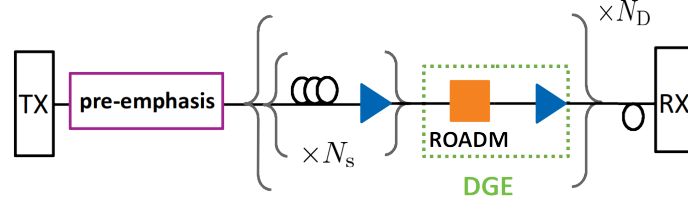


Figure 3.22: Sketch of a generic link structure with N_D sections of N_s spans each. The DGE node is the cascade of a lossy ROADM and an amplifier.

relied on the steepest descent algorithm combined with particle swarm optimization and the GN model. We are thus motivated to reduce the complexity of the problem by introducing a constraint. We limit the search space to power allocations in the form of an opposite SRS gain, as for the pre-emphasis in Eq. (3.26), which is a linear function in a dB scale.

We aim at estimating the SNR of a link having sparse DGE positioning for the SRS gain equalization on the signal power. Figure 3.22 shows an example of the link under test, with a total of N spans subdivided into N_D sections of N_s spans each, with end-section DGEs. All line amplifiers have flat gain equal to the span loss and identical noise figure.

Contrary to the previous sections of this chapter, we model the DGE as a lossy filter followed by a noisy flat amplifier restoring the TX total power P_t . Such a model emulates the equalization usually performed in a reconfigurable optical add-drop multiplexer (ROADM) node through wavelength selective switches (WSS).

Under the assumption of identical spans and N_D link sections, the variance of the ASE noise introduced by the line amplifiers in Eq. (3.11) generalizes to

$$\sigma_{\text{ASE}}^2(\nu) = h\nu FBN_D \left(G \sum_{k=1}^{N_s} H_k^{-1}(\nu, L) + G_D \right) \quad (3.43)$$

with F the amplifier noise figure, G the gain equal to the span loss, and G_D the gain of the amplifier within the DGE node that recovers its losses. The factor $H_k^{-1}(\nu, L)$ accounts for the net SRS gain experienced by the ASE noise introduced by the k th amplifier, including the DGE gain, as in Eq. (3.11). The equivalent block diagram

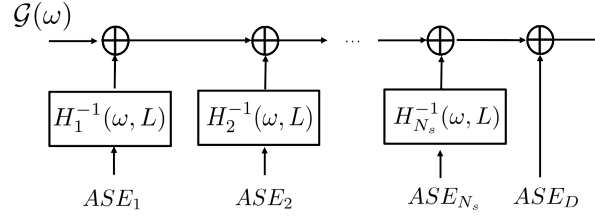


Figure 3.23: Equivalent block diagram representation of a link section with noisy DGE node. ASE_k , $k = 1, \dots, N_s$, is the variance of the ASE noise introduced by the k th amplifier without the effects of Raman scattering.

representation of a link section in Fig. 3.6 thus generalizes the one sketched in Fig. 3.23 including the ASE noise introduced by the amplifier within the DGE node.

As a sanity check of the SNR estimation based on the DGE-SRS-GN model, Fig. 3.24 shows a comparison with SSFM simulations for a 4×100 km link of SMF with parameters listed in Tab 3.1. We focused on a simple scenario of 51 WDM channels with Gaussian-distributed symbols, symbol rate $R = 49$ Gbaud, and channel spacing $\Delta f = 50$ GHz. The pre-emphasis was applied with a factor $\bar{k} = 4$ at a total WDM

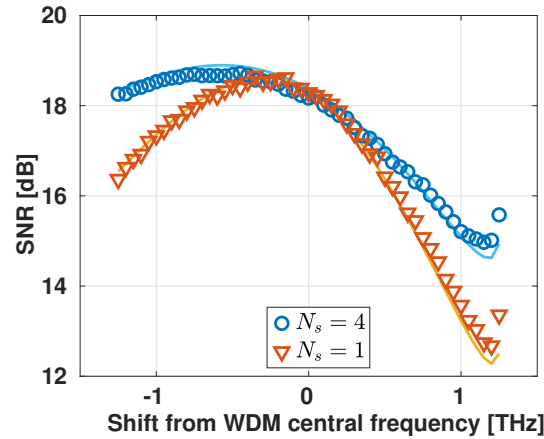


Figure 3.24: SNR vs. frequency with a DGE at each span-end ($N_s = 1$) and without inline DGEs ($N_s = 4$). Markers: SSFM simulations. Lines: DGE-SRS-GN model.

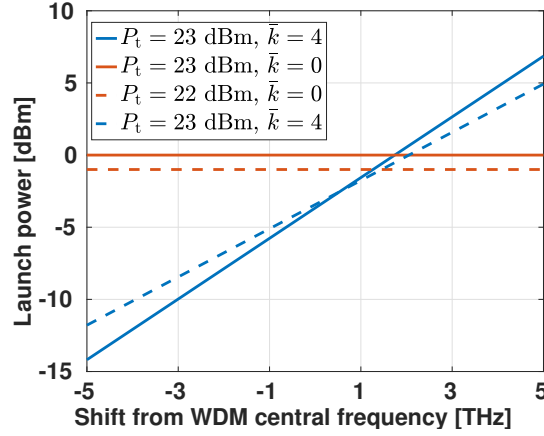


Figure 3.25: Launch power for pre-emphasis factor $\bar{k} = 0$ or $\bar{k} = 4$ and total power $P_t = 22, 23$ dBm, vs. frequency shift.

power of 20 dBm. To stress the model, we artificially inflated the Raman gain slope to $C_r = 5 \times 0.028 / (2\pi)$ (THz·rad·km·W) $^{-1}$ to emphasize the SRS over the 2.5 THz bandwidth of the signal, a value for which reliable SSFM double checks were feasible. The amplifiers' noise figure was $F = 5$ dB. Each DGE node introduced both SRS-dependent losses, due to the passive equalization, and additional 11 dB of SRS-independent losses due to the WSSs cascade in the ROADM node. At the receiver side, the pre-emphasis is removed from the signal. As a consequence, the ASE variance in Eq. (3.43) is further subject to the pre-emphasis undo, i.e., a multiplication by $1/H_{-\bar{k}}(\nu, L)$, while the NLI variance expression obtained adopting Eq. (3.26) is already in the reference system that equalizes such an operation. Figure 3.24 shows the RX SNR obtained with the placement of a DGE at each span-end ($N_s = 1$) and without inline DGEs ($N_s = 4$). For both cases, the theoretical estimation is in good agreement with the simulations.

We next optimize the power allocation in terms of signal power pre-emphasis for link topologies with different numbers of DGEs. For a fixed total power P_t , we first proceed with an uniform power allocation among all the WDM channels, i.e., $P = P_t/N_{\text{ch}}$. Then, we impose an opposite-sign SRS gain, as per Eq. (3.26), governed

by the pre-emphasis factor \bar{k} . An example is sketched in Fig. 3.25 for two total power values, and a 10 THz WDM signal. The dashed lines correspond to the homogeneous power allocation before pre-emphasis, hence with $\bar{k} = 0$, while the solid lines report the channel power values after the pre-emphasis filtering with a factor $\bar{k} = 4$.

We considered two different optimization strategies. The first strategy aims at maximizing the performance of the worst channel in the WDM comb, thus avoiding penalizing some users due to SRS-induced unbalances, namely

$$(P_t, \bar{k})_{\text{opt}} = \underset{P_t, \bar{k}}{\operatorname{argmax}} (\min \text{SNR}(f_i)) \quad (3.44)$$

where the min is taken among the WDM channels. We call it the *max-min* strategy.

The second strategy aims at maximizing the link AIR when treating NLI as a Gaussian noise:

$$(P_t, \bar{k})_{\text{opt}} = \underset{P_t, \bar{k}}{\operatorname{argmax}} \left(2B \sum_{i=1}^{N_{\text{ch}}} \log_2(1 + \text{SNR}(f_i)) \right) \quad (3.45)$$

and we call this the *max-AIR* strategy.

For these optimizations, the link under test was composed of 12×100 km of SMF with parameters as in Tab. 3.1 and $C_r = 0.028/(2\pi)$ (THz·rad·km·W)⁻¹. The transmitted signal was composed of 201 WDM channels carrying Gaussian-distributed symbols, with R and Δf as in Fig. 3.24, for a total bandwidth of 10 THz. We considered four link topologies with DGEs placed every $N_s = 1, 2, 3$ or 4 spans. For each link topology, we varied the power P_t from 20 to 24 dBm, by steps of 0.5 dB, and \bar{k} from 0 to 4, by steps of 0.1, and tested all the combinations of (P_t, \bar{k}) .

The optimized pair of total power and pre-emphasis factor $(P_t, \bar{k})_{\text{opt}}$ values for the two strategies are reported in Fig. 3.26 for the different link topologies. Solid lines refers to the max-min strategy, while dashed lines to the max-AIR strategy.

The figure shows that each DGE placement along the link calls for a different power allocation. In particular, as the DGE period increases more pre-emphasis is needed to counteract the accumulation of the SRS gain. Moreover, such an optimal

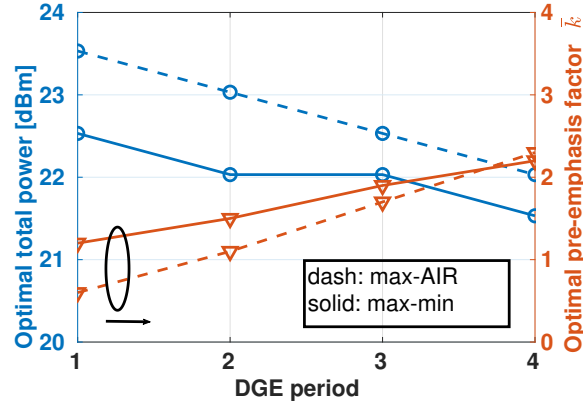


Figure 3.26: Optimal total power (left axis) and optimal pre-emphasis factor (right axis) vs. the DGE period N_s , i.e., number of spans between two equalizers, for the two strategies.

power allocation differs for the two strategies. We note that the max-AIR strategy accepts 0.5 – 1 dB more total power, hence more SRS gain, compared to the max-min strategy which needs to limit the amount of SRS-tilt to increase the performance of the worst channel in the comb.

For the sake of completeness, the total AIR and minimum SNR as a function of the total power and pre-emphasis factor are reported in Fig. 3.27 and 3.28, respectively. It can be seen that these surfaces are concave and hence an isolated maximum exists. In both figures, the magenta circle highlights the optimal solution according to the max-AIR strategy while the green triangle indicates the optimal solution for the max-min strategy. It is worth noting that, thanks to the GN model closed-form expression simplicity, the exhaustive search results reported in Fig. 3.27 and 3.28 can be computed in a couple of minutes, without the need for complex algorithmic optimizations.

The estimated total AIR and minimum SNR in the WDM comb are represented in Fig. 3.30 for max-AIR (top) and max-min (bottom). Hence, the top figure collects the values relative to the magenta circles in Fig. 3.27 while the bottom figure collects the

value the green triangles values in Fig. 3.28. Figure 3.29 also reports with markers the optimization results obtained without pre-emphasis ($\bar{k} = 0$), hence with an optimized frequency-flat power. We note that the latter results are below the solid lines obtained with optimized pre-emphasis for all the values of the DGE period. In particular, the benefit of the \bar{k} -optimized pre-emphasis increases with the number of spans between neighboring DGEs, for both strategies.

Figure 3.29 shows that, in both cases, the best link topology is the one having

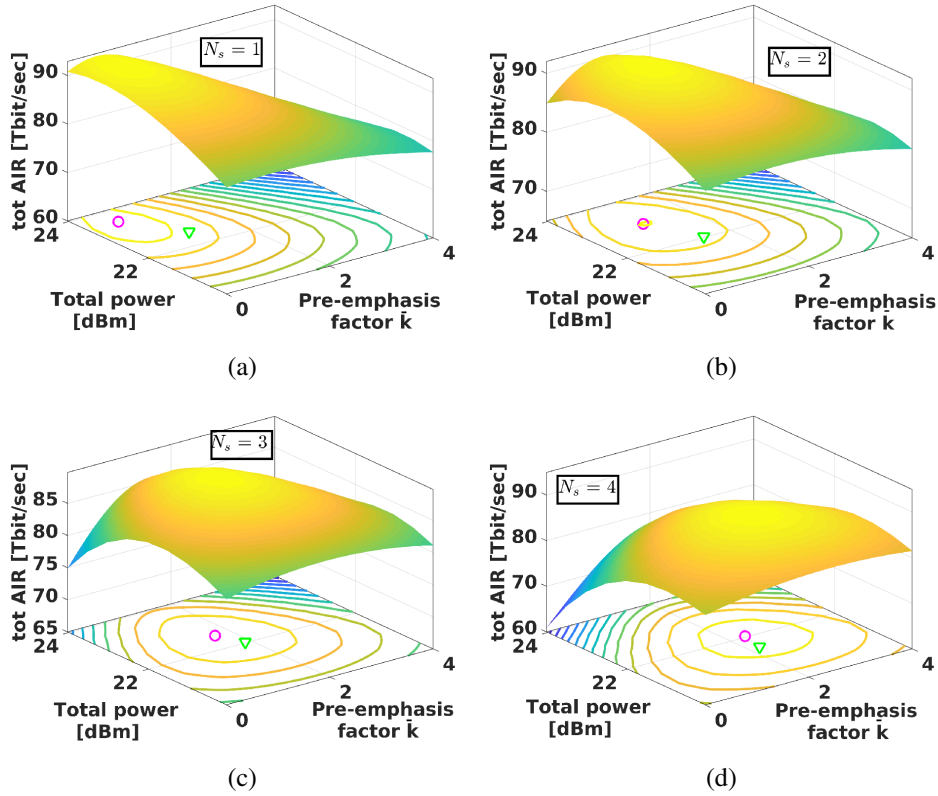


Figure 3.27: Total AIR vs. total power and pre-emphasis factor for $N_s = 1, 2, 3$, and 4. Magenta circles and green triangles: $(P_t, \bar{k})_{\text{opt}}$ according to the max-AIR and the max-min strategy, respectively.

a DGE at each span-end. Increasing the DGE period, hence moving towards less expensive link structures with fewer equalizers, introduces a penalty in the objective function. For the max-min strategy, such a penalty is up to 3 dB for $N_s = 4$ when no pre-emphasis is applied. However, we note that such a penalty reduces to less than 1 dB by applying an optimized pre-emphasis. Similarly, for the max-AIR case, such a cost-saving link structure implies less than 10% of capacity reduction with the optimal power allocation.

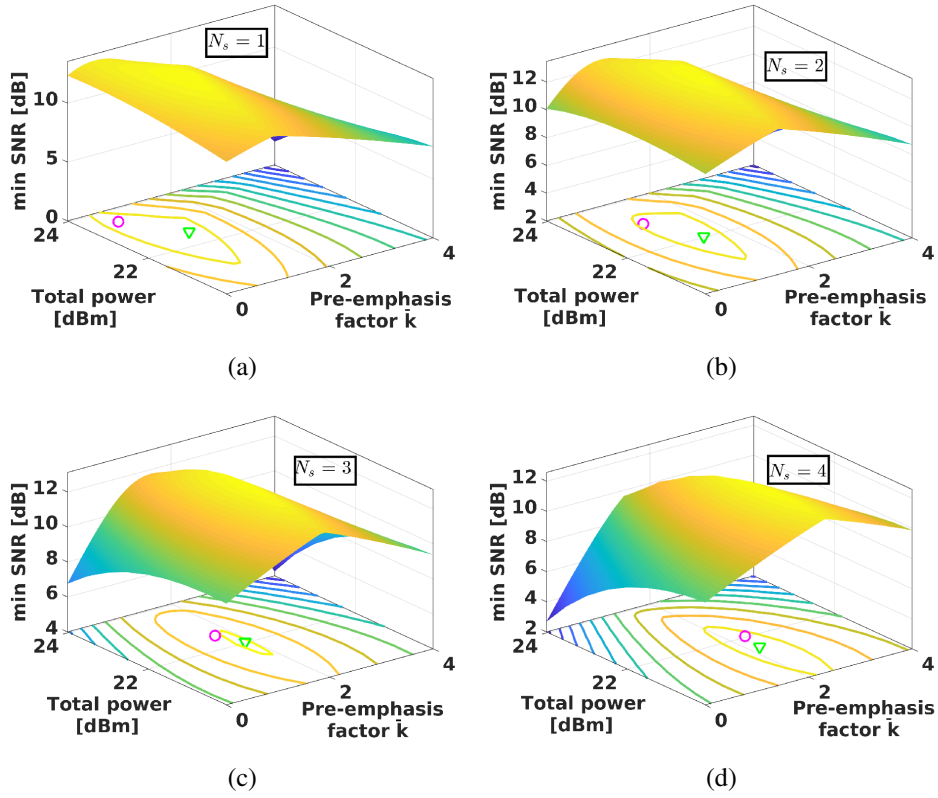


Figure 3.28: Minimum SNR across the WDM comb vs. total power and pre-emphasis factor for $N_s = 1, 2, 3$, and 4. Magenta circles and green triangles: $(P_t, \bar{k})_{\text{opt}}$ according to the max-AIR and the max-min strategy, respectively.

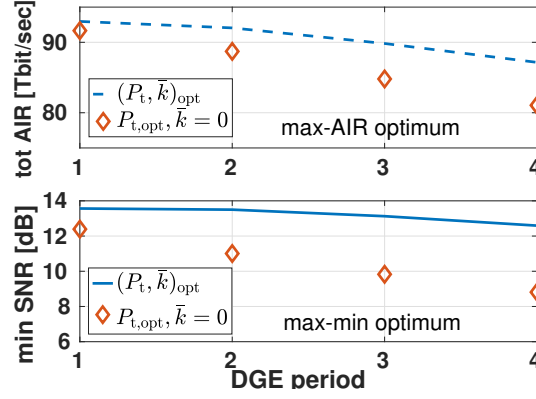


Figure 3.29: Optimal values of total AIR (top) and minimum SNR in the WDM comb (bottom) vs. the DGE period N_s . Lines: with optimal pre-emphasis. Markers: without pre-emphasis ($\bar{k} = 0$).

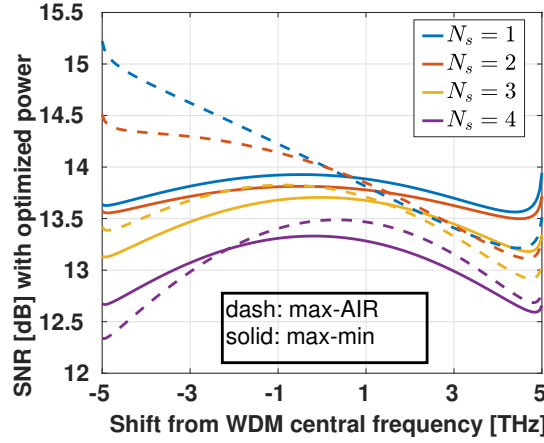


Figure 3.30: SNR with optimized powers for the two strategies vs. frequency, for a DGE period of $N_s=1,2,3,4$ spans. $N=12$ spans. Dashed: max-AIR strategy. Solid: max-min strategy.

The SNR at the optimal power allocation is reported in Fig. 3.30, versus the frequency shift with respect to the central WDM frequency. Such an SNR is almost frequency-flat for the max-min strategy (solid lines), hence yielding similar performance for all channels. On the other hand, the SNR curves maximizing the total capacity exhibit up to 2 dB of tilt across the bandwidth for $N_s = 1$. It can be seen that such an SNR tilt decreases as the spans between neighboring DGEs increase. Therefore, the cost-saving solution of $N_s = 4$, which introduces a penalty in the total achievable information rate, has the advantage to mitigate the SNR variations in frequency.

Chapter 4

Modeling of semiconductor optical amplifiers

Ultra-wideband transmissions for high capacity systems face many challenges. From the technological point of view, an important issue that must be addressed is the need to broaden the amplification bandwidth. The most promising strategies investigated in the literature are multi-band EDFA amplification [75], Raman amplification [76–79], hybrid EDFA-Raman schemes [80, 81], parametric amplifiers [82], and SOAs.

Regardless of their amplification bandwidth beyond 12 THz, in the past years, SOAs were not considered suitable for WDM transmission systems, due to their noise figure, typically higher than EDFAs, and their poor nonlinear gain dynamics [12, 83]. Motivated by the advent of transmissions beyond the C band, the feasibility of WDM transmission along an SOA-amplified link was investigated numerically in [84] and demonstrated experimentally in [85] thus stimulating further studies [86, 87] and design of novel SOAs. In particular, in [88] the GN model for the estimation of the NLI variance was extended to an SOA-amplified optical link.

Both the numerical (SSFM) and the analytical (GN model) estimation of the performance of an SOA-amplified system requires a reliable field propagation model along the amplifier. In the literature, several SOA models were proposed, spacing from very accurate models [83, 89] based on tens of parameters related to the phys-

ical device characteristic, to simpler parametric models well suited to be fitted to experimental results [90–94]. Thanks to its simplicity, the latter category of models, typically called *integrated-gain* or *reservoir* models, gained particular interest in long-haul transmission systems to relax the complexity of the system. In this scenario, it is desirable to find a trade-off between the model accuracy and the numerical effort required to simulate each amplifier.

In this chapter, we first review the Connelly model in [89] for a bulk InP–InGaAsP SOA. Then, to simplify such a cumbersome model, we identify its key features and investigate their role in SOA behavior. As another simplification technique, we propose a parametric integrated-gain model whose parameters are extracted by fitting to the reference model [89].

Finally, we investigate the performance of an SOA-amplified. In this Chapter, we relied on SSFM simulations for the estimation of the system performance to test the integrated-gain model in such a multi-span context. Such a numerical investigation might serve as a first step towards the extension of the GN-model to SOA-amplified links based on the proposed integrated-gain model for the amplifier, as it was done in [88] with the simple model in [90].

The work presented in this chapter was carried out during a six-month internship at Nokia Bell Labs, Paris-Saclay, France.

4.1 Connelly model overview

SOAs are composed of an active material in which light amplification is provided by stimulated emission of photons resulting from carrier recombination across the conduction band (CB) and valence band (VB). For amplification to be possible, population inversion is required which is provided by an electrical bias current I .

The carrier density n is depleted, and hence the material gain is reduced, not only by recombination resulting in frequency-dependent signal amplification but also because of spontaneous emission resulting in ASE noise. In addition, the propagation of both signal and noise along the SOA length is bidirectional due to the presence of reflective facets at the SOA edges.

For small input signal power, the SOA amplification is a linear process with the signal power providing the maximum available gain, also called *small-signal gain*, which is independent of input power variations. As the input signal power increase, the carrier density gets depleted by signal amplification triggering a *nonlinear regime* in which the available gain decreases with the growth of the input power. From a dynamic point of view, the SOA gain fluctuates in time in response to the temporal variations of the input signal.

Mathematically, the propagation of the signal field $E_k(z, t)$ at frequency ω_k along the amplifier length, for the forward (+) and backward (−) direction, is governed by the following equation [89]

$$\frac{\partial E_k^\pm(z, t)}{\partial z} = \left\{ \mp j\beta_k \pm \frac{1}{2} \left[\Gamma(1 + j\alpha_H)g_k(n) - \alpha_s(n) \right] \right\} E_k^\pm(z, t) \quad (4.1)$$

where: β_k is the propagation coefficient, Γ is the optical confinement factor, n is the carrier density, α_s is the SOA loss coefficient, α_H is the linewidth enhancement factor, and $g_k(n)$ is the material gain, which depends both on the wavelength and the carrier density. The corresponding signal power is defined as $P_k^\pm(z, t) = |E_k^\pm(z, t)|^2$. The expression of the material gain coefficient and the definition of all the involved parameters and physical quantities can be found in Appendix A. The attenuation coefficient is linearly related to the carrier density by $\alpha_s(n) \triangleq K_0 + \Gamma K_1 n$.

Similarly, the ASE noise field $W_j(z, t)$ at frequency ω_j propagates along the z direction according to

$$\frac{\partial W_\ell^\pm(z, t)}{\partial z} = \left\{ \mp j\beta_\ell \pm \frac{1}{2} \left[\Gamma(1 + j\alpha_H)g_\ell(n) - \alpha_s(n) \right] \right\} W_\ell^\pm(z, t) + v_\ell^\pm(z, t) \quad (4.2)$$

where $v_\ell^\pm(z, t)$ is the spontaneously emitted noise coupled into $W_\ell(z, t)$. As for the signal, the noise power is given by $N_\ell^\pm(z, t) = |W_\ell^\pm(z, t)|^2$.

The signal and noise evolution depend on the carrier density $n(z, t)$ in the active region, which obeys the following rate equation [89]:

$$\begin{aligned}
\frac{\partial n(z,t)}{\partial t} = & \frac{I}{qV_a} - r(n) \\
& - \frac{\Gamma}{A_a} \sum_{k=1}^{N_{\text{ch}}} g_k(n) \frac{P_k^+(z,t) + P_k^-(z,t)}{\hbar\omega_k} \\
& - \frac{2\Gamma}{A_a} \sum_{\ell=1}^{N_{\text{ase}}} K_\ell g_\ell(n) \frac{N_\ell^+(z,t) + N_\ell^-(z,t)}{\hbar\omega_\ell}
\end{aligned} \tag{4.3}$$

where: N_{ch} and N_{ase} are the number of signal and ASE channels, I is the bias current, q is the elementary charge, A_a and V_a are the SOA's active region area and volume, \hbar is the Planck constant divided by 2π , and K_ℓ is a normalization factor. $r(n)$ is the recombination rate, defined as:

$$r(n) \triangleq A_{\text{rad}}n + B_{\text{rad}}n^2 + A_{\text{nrad}}n + B_{\text{nrad}}n^2 + C_{\text{aug}}n^3 + D_{\text{leak}}n^{3.5}. \tag{4.4}$$

For the sake of simplicity, the values of the coefficients of the recombination rate along with all the other SOA parameters considered in this work are collected in Tab. A.1 in Appendix A.

4.1.1 Numerical solution

With a constant-wave signal and in the absence of ASE, it is $\partial n / \partial t = 0$. A steady-state carrier density \bar{n} thus exists, such that:

$$\begin{aligned}
F(\bar{n}) \triangleq & \frac{I}{qV_a} - r(\bar{n}) \\
& - \frac{\Gamma}{A_a} \sum_{k=1}^{N_{\text{ch}}} g_k(\bar{n}) \frac{P_k^+(z) + P_k^-(z)}{\hbar\omega_k} \\
& - \frac{2\Gamma}{A_a} \sum_{\ell=1}^{N_{\text{ase}}} K_\ell g_\ell(\bar{n}) \frac{N_\ell^+(z) + N_\ell^-(z)}{\hbar\omega_\ell} = 0.
\end{aligned} \tag{4.5}$$

As there is no analytical solution for the steady-state model, we adopted the numerical solution proposed in [89], here summarized for the sake of completeness.

The numerical algorithm consists in dividing the amplifier into N_{sec} sections of length L_{sec} . For each section j , the carrier density is assumed constant and it is approximated with its value in the center of the section, namely:

$$\bar{n}(z) \approx \bar{n}\left(\frac{z_j + z_{j+1}}{2}\right) \quad z_j < z < z_{j+1}. \quad (4.6)$$

This method allows to split the signal and noise propagation along z into a concatenation of N_{sec} propagation equations. Since the parameters of each section are constant, the input-output fields of the section k are related by:

$$E_k^\pm(z_j) = \exp\left\{\left[-j\beta_k + \frac{1}{2}\left(\Gamma(1 + j\alpha_H)\bar{g}_k(z_{j\mp 1}) - \bar{\alpha}_s(z_{j\mp 1})\right)\right]L_{\text{sec}}\right\}E_k^\pm(z_{j\mp 1}). \quad (4.7)$$

Similarly, with constant parameters per section Eq. (4.2) can be solved with zero initial condition at the SOA input. The ASE input power for a generic section j can then be expressed as:

$$\begin{aligned} N_\ell^\pm(z_j) = & \exp\left\{[\Gamma\bar{g}_\ell(z_{j\mp 1}) - \bar{\alpha}_s(z_{j\mp 1})]L_{\text{sec}}\right\}N_\ell^\pm(z_{j\mp 1}) \\ & + \frac{\exp\left\{[\Gamma\bar{g}_\ell(z_{j\mp 1}) - \bar{\alpha}_s(z_{j\mp 1})]L_{\text{sec}}\right\} - 1}{\Gamma\bar{g}_\ell(z_{j\mp 1}) - \bar{\alpha}_s(z_{j\mp 1})}|\mathbf{v}_\ell^\pm(z_{j\mp 1})|^2. \end{aligned} \quad (4.8)$$

Accounting for input-output coupling losses α_c and SOA facets reflectivity coefficients r_1 and r_2 , the signal at the input of the first section and at the SOA output are respectively:

$$E_k^+(z_1) = \sqrt{\alpha_c}(1 - \sqrt{r_1})E_{k,in}^+ \quad (4.9)$$

$$E_{k,out}^+ = \sqrt{\alpha_c}(1 - \sqrt{r_2})E_k^+(L_a) \quad (4.10)$$

where $E_{k,in}^+$ is the signal entering the SOA and L_a is the SOA length. The signal field traveling in the backward direction at the boundary is $E_k^-(L_a) = \sqrt{r_2}E_k^+(L_a)$. Regarding the noise, the boundary conditions are $N_\ell^+(z_1) = 0$, and $N_\ell^-(L_a) = r_2N_\ell^+(L_a)$. A sketch of the section-wise SOA discretization is shown in Fig. 4.1.

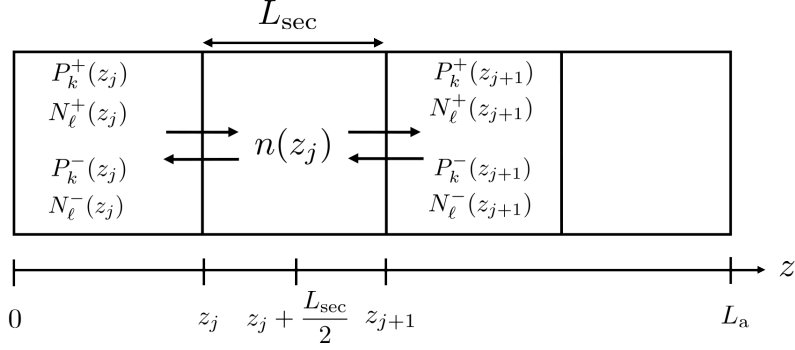


Figure 4.1: Sketch of the SOA amplifier divided into sections.

The key steps of the steady-state algorithm are summarized in Fig. 4.2. The algorithm iteratively adjusts the carrier density values of each section j by means of a weight $w(j)$ such that $F(\bar{n}(z_j))$, defined in Eq. (4.5), approaches zero. For each section, we used a weight initial value equal to $w(j) = 0.1$, which gets halved every time that the sign of $F(j)$ changes with respect to the previous iteration. The algorithm stops when the maximum relative change in the carrier density with respect to the previous iteration is below a threshold value $\vartheta = 0.001$. The threshold value was found by iteratively adapting its value until the maximum percentage variation in the exponent of Eq. (4.7) was below 1%.

The carrier density \bar{n}_0 in the absence of signal and noise, which satisfies the following:

$$F(\bar{n}_0) = \frac{I}{qV_a} - r(\bar{n}_0) = 0 \quad (4.11)$$

is used as initial value for the steady-state model. This value is found by means of the Newton's method [71], which iteratively updates the carrier density according to:

$$\bar{n}_{\text{new}} = \bar{n}_{\text{old}} - \frac{F(\bar{n}_{\text{old}})}{\dot{F}(\bar{n}_{\text{old}})} \quad (4.12)$$

where we adopted the following compact notation $\dot{F}(\bar{n}) = \frac{dF(\bar{n})}{d\bar{n}}$. The algorithm begins with a carrier density initial value of 10^{24} m^{-3} and terminates when the percent-

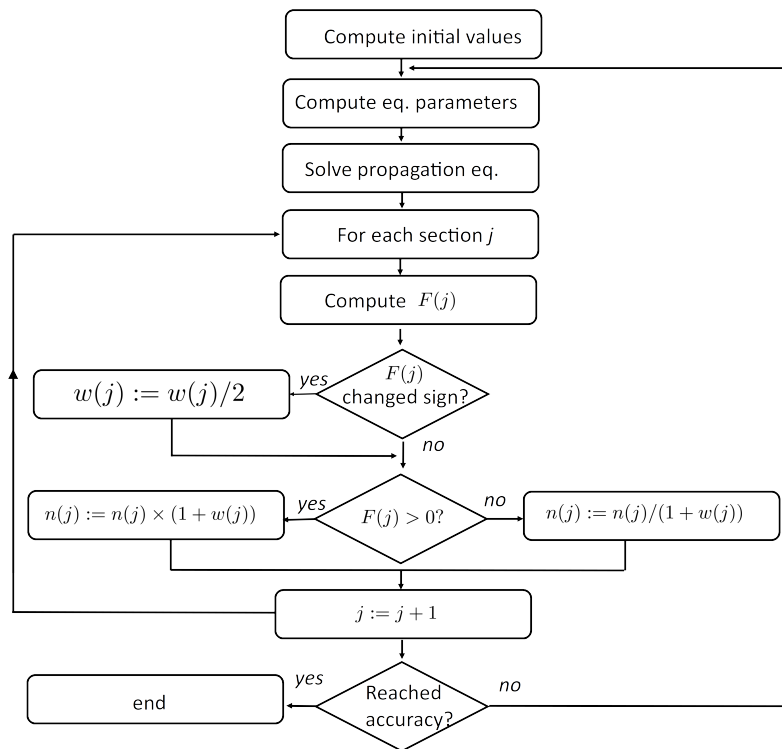


Figure 4.2: Steady-state algorithm [89].

age variation in \bar{n} , with respect to the previous iteration, is below 0.1%.

Having found the steady-state solution of the carrier density, we use it to solve the time evolution of the carrier density rate equation, Eq. (4.3), through the fourth-order Runge-Kutta method (RK4) [71]. The carrier density initial value for the time evolution is a vector having N_{sec} elements:

$$\mathbf{n}(t_0) \triangleq \{n(z_1, t_0), \dots, n(z_{N_{\text{sec}}}, t_0)\}. \quad (4.13)$$

The vector of initial values can be found as the solution of a steady-state pre-run. Time-dependent quantities are discretized assumed to be constant within a time step. Hence, at each time t the quantities dependent on the carrier density value are computed, and the z -propagation equations in Eq. (4.7)-(4.8) are solved with the current value of $\mathbf{n}(t)$.

4.2 SOA model simplifications

The model reviewed in Sec. 4.1 is based on a system of coupled differential equations in space and time, and relies on tens of parameters, as reported in Tab. A.1. From a system point of view, it is desirable to rely on a simpler model that captures the SOA key features.

To simplify the amplifier model, we identify three main features characterizing the SOA behavior: spatial resolution (z -dependent parameters), presence of ASE noise, and bidirectionality (presence of reflective facets). We consider all the possible combinations of these SOA key features, obtaining eight different amplifier configurations, and we investigate the impact of these features on the overall SOA behavior. This amplifier model decomposition is sketched by the binary tree in Fig. 4.3.

On top of the tree, we start by choosing between the presence or absence of reflective facets, hence unidirectional or bidirectional field propagation. Then, for each case, we consider both a noisy and noiseless amplifier model. At this level of the tree, we have four leaves each representing a possible SOA model. For instance, UN stands for unidirectional and noisy amplifier. As a last step, we choose whether to include the spatial resolution in the model. This feature is labeled as multi-section (M) in Fig.

4.3, with reference to the discretization approach along the SOA length required to solve such a model. In the most simple case, we reduce the SOA to be unidirectional, without ASE noise, and having carrier density constant along its length. This last assumption goes by the name of single-section (S), since it represents a special case of the space-resolved model having spatial resolution equal to the whole amplifier length. Following this logic, this simple SOA configuration is represented by the leftmost leaf of the binary tree and it is labeled U-S. On the contrary, the rightmost leaf represents the most rigorous SOA characterizations considered in this work, which is comprehensive of all the key features. This BNM model coincides with the Connelly model in Sec. 4.1.

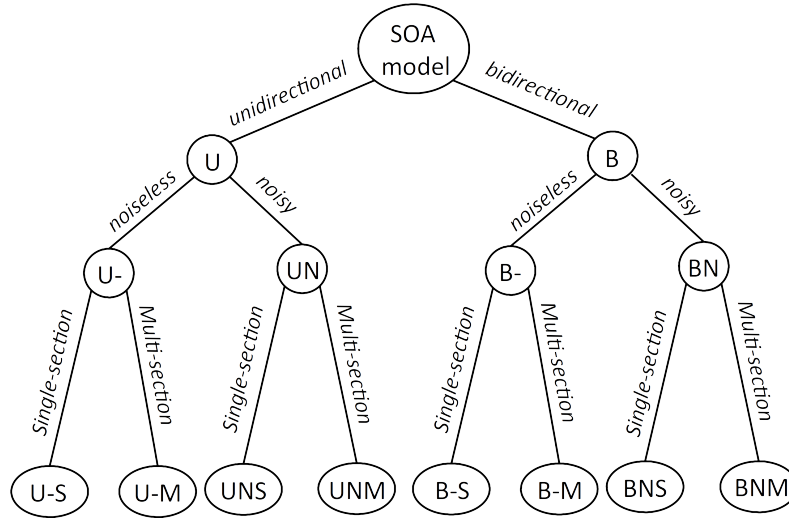


Figure 4.3: A binary tree of the possible SOA modeling choices. We adopted the following nomenclature: "U" for unidirectional SOA, "B" for bidirectional SOA, "N" for noisy SOA, "-" for noiseless SOA, "S" (single-section) for the absence of spatial resolution, and "M" (multi-section) for inclusion of spatial resolution. The leftmost leaf represents the simplest SOA according to the adopted categories, i.e., U-S, while the rightmost leaf BNM represents the Connelly model in Sec. 4.1, which is the most complex model considered in this work.

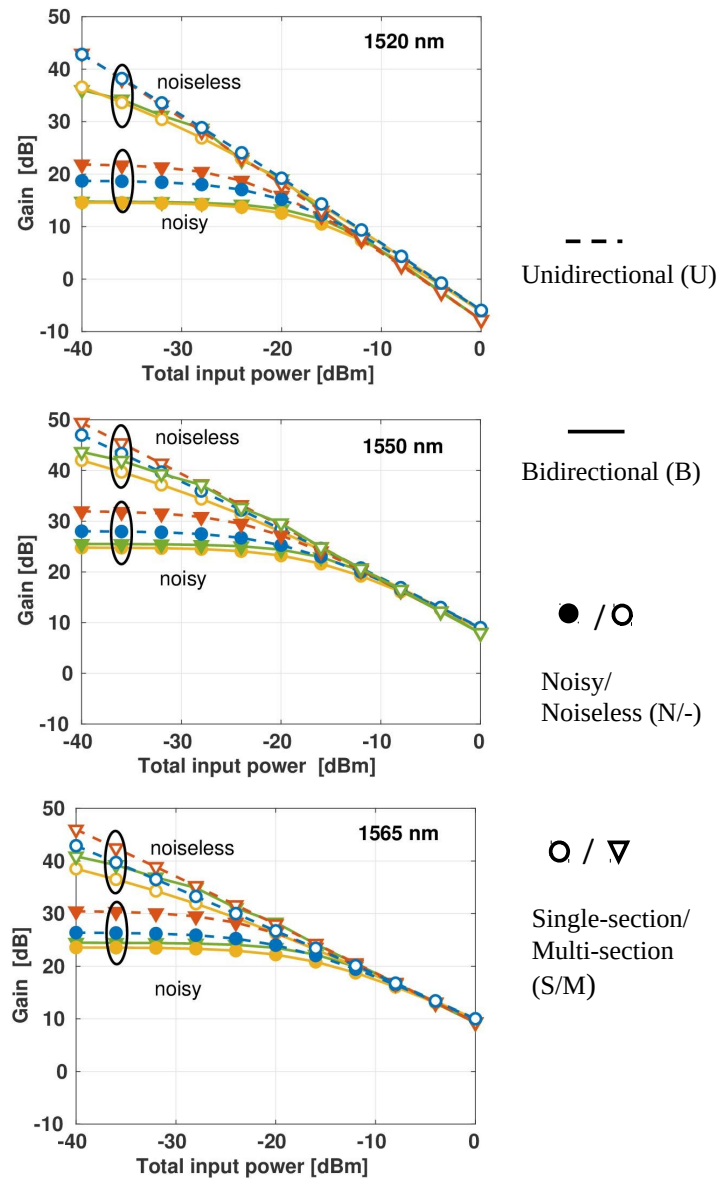


Figure 4.4: Steady-state gain vs. SOA total input power at 1520 nm (top), 1550 nm (center), and 1565 nm (bottom), computed with the eight SOA models in Fig. 4.3.

At first, we focus on the static behavior of a generic single SOA in an optical link. We simulate the field propagation along the SOA for the eight amplifier configurations represented by the eight final leaves of Fig 4.3. The numerical implementation of the amplifier is discussed in Sec. 4.1.1, and all values used in this work are reported in Tab. A.1. We considered a C-band transmission of 3 CW signals at wavelengths $\lambda_1 = 1520$ nm, $\lambda_2 = 1550$ nm, and $\lambda_3 = 1565$ nm. Figure 4.4 shows the input-output SOA gain at steady-state, measured as the difference (in dB) between the output and input power at each wavelength, as a function of the total signal power at SOA input. The input power values were chosen considering that, in an optical link, the signal power suffers from ≈ 20 dB of fiber losses over 100 km of span length. The power allocation was homogeneous, such that each channel carried 1/3 of the total power.

Figure 4.4 collects the results obtained with the eight SOA models. The dashed curves are associated with the four unidirectional SOA models, while the solid lines represent bidirectional models. We used empty markers for noiseless models and full markers for noisy SOAs. Finally, circles are associated with single-sections implementations while triangles identify multi-section models. For this latter case, we found that a spatial resolution of 50 section was enough to provide accurate results for the SOA under test.

Regarding the noisy models, all the gain curves exhibit a small-signal gain regime at small input power values. In this regime, the signal power is outweighed by the ASE power, and spontaneous emission represents the major contribution to carrier depletion. As a consequence, the gain is independent of small variations in such low signal power. The nonlinear regime is then triggered at different input power according to the wavelength and the considered amplifier model. In particular, at 1520 nm the nonlinear regime is triggered at smaller input powers compared to the higher wavelengths under test. On the contrary, noiseless amplifiers operate in the nonlinear regime for the whole range of considered power values. In fact, for these amplifiers, the small-signal gain regime requires input powers that are orders of magnitude smaller than the considered values.

Figure 4.4 also shows that unidirectional models overestimate the SOA small-signal gain, meaning that the presence of reflective facets particularly inflates the

ASE noise power by introducing the backward propagating noise. In addition, it can be seen that, for each wavelength, the absence of spatial resolution in the modeling of the SOA results in a smaller steady-state gain in the small-signal gain regime. The discrepancy with respect to space-resolved models is more significant for unidirectional amplifiers. On the other hand, as the nonlinear regime is triggered, the slope of the gain curves is steeper for those models that account for the z -dependency of the carrier density.

As a next investigation, we considered only two SOA models, the UNM and BNM models, and we focused on the impact of the number of WDM channels on the SOA gain. At first, we simulated the transmission of a single CW signal for three different wavelengths $\lambda_1 = 1520$ nm, $\lambda_2 = 1550$ nm, and $\lambda_3 = 1565$ nm to cover the $N_{\text{ch}} = 1$ scenario.

Then, we considered the $N_{\text{ch}} = 3$ case, that is the one in Fig. 4.4. Finally, we increased the channel number by keeping fixed the three wavelengths of interest (WOI) $\lambda_1, \lambda_2, \lambda_3$ and including equally spaced channels, namely $N_{\text{ch}} = 7$ with 7.5 nm spacing, and $N_{\text{ch}} = 13$ with 3.75 nm spacing. The results are reported in Fig. 4.5. The figure shows that single channel transmissions are not enough to characterize the wavelength-dependent gain for a full C-band transmission. On the other hand, we measured a negligible discrepancy between the 7 and 13 channel cases, for all the wavelengths and for SOA models, as highlighted in the insets of Fig. 4.5. Moreover, it can be noticed that the transmission of the three WOIs yields a gain discrepancy of less than 1 dB compared to the 13 channel cases. This small gap at low channel count can be attributed to the fact that the WOIs are chosen to include the C-band edges, i.e., 1520 nm and 1565 nm, in addition to the wavelength 1550 nm for which the material gain is close to its maximum value. Hence, adding further channels in these ranges has a minor impact. Therefore, Fig. 4.5 suggests that transmission of 7 wavelengths with 7.5 nm of spacing can be sufficient to characterize the steady-state gain of a WDM full C-band transmission.

Figure 4.6 reports the results of Fig. 4.5 for the 7 channel case and BNM model. It can be noticed that not only the small-signal gain is wavelength dependent, but also that the gain curve slopes in the nonlinear regime are different. In particular, the

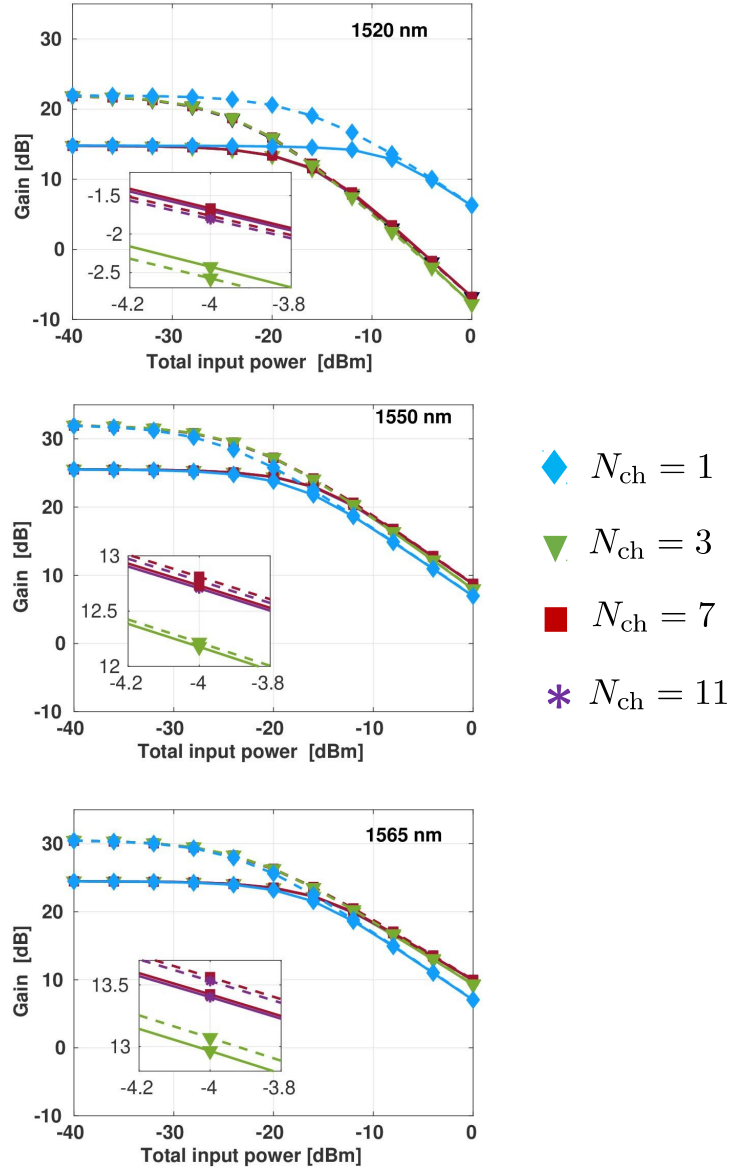


Figure 4.5: Steady-state gain vs. SOA total input power, measured at the WOI 1520 nm (top), 1550 nm (center), and 1565 nm (bottom), for $N_{ch} = 1, 3, 7, 13$ channels. Dashed lines: UNM. Solid lines: BNM.

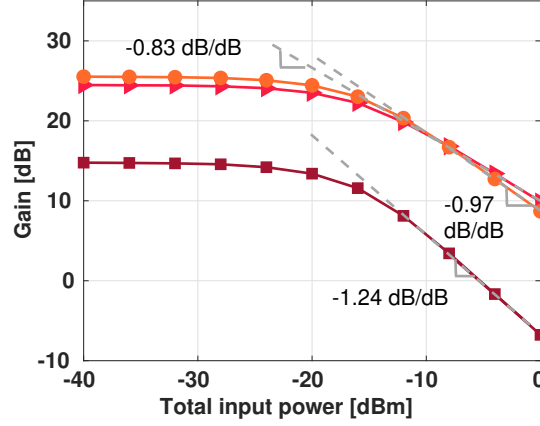


Figure 4.6: Steady-state gain vs. SOA total input power, measured at the WOI 1520 nm, 1550 nm, and 1565 nm for a 7 WDM channel transmission with channel spacing 7.5 nm. The SOA model is BNM. Dashed lines highlight the different slopes of the gain curves in the nonlinear regime.

smallest wavelength 1520 nm exhibits the smallest linear gain and the steepest slope in the nonlinear regime. The slope decreases as the wavelength increase, while the small-signal gain reaches its maximum at 1550 nm and then decreases.

For the sake of completeness, we reported in Fig. 4.7 the steady-state gain as a function of wavelength for different total input power values. The setup and the SOA model are the same as Fig. 4.6. The figure clearly shows the gain wavelength dependency as well as the nonlinear gain compression phenomenon. It can be seen that as the input power increases the wavelength associated with the gain peak increases. Therefore, the wavelength dependency represents a key feature for the SOA steady-state gain characterization, on top of the other classified features.

4.3 The integrated-gain model

In the previous section, we investigated the possibility of simplifying the cumbersome SOA model in [89] by removing some basic amplifier features at the modeling stage,

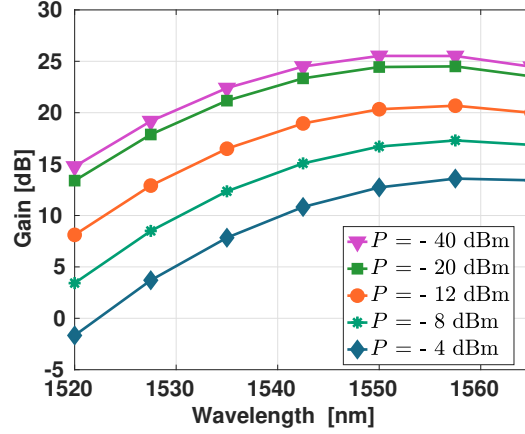


Figure 4.7: Steady-state gain vs. wavelength for a 7 WDM channel transmission equally spaced over the wavelength axis, for different input power. The SOA model is BNM.

still relying on the same framework for describing the material gain, which can be found in Appendix A.

Another approach to simplify the SOA modeling is represented by parametric models [90–92, 94]. The key idea shared by all these model is to work with a linearized and z -integrated material gain. The former approximation consists in modeling the material gain as a linear function of the carrier density. Such an approach drastically simplifies the problem by removing many parameters and functions involved in the material gain description. In addition, the carrier density dependency on distance is removed by means of an integration along z .

In this section, we propose a parametric model that we called the integrated-gain model. We start from the rate equation without ASE noise:

$$\frac{\partial n(z, t)}{\partial t} = \frac{I}{qV_a} - r(n) - \frac{\Gamma}{A_a} \sum_{r=1}^{N_{ch}} g_r(n) \frac{P_r(z, t)}{\hbar\omega_k}$$

which is coupled to the following simplified expression of the evolution of the signal power along distance:

$$\frac{\partial P_r(z, t)}{\partial z} = \Gamma g_r(n) P_r(z, t) \quad (4.14)$$

where SOA absorption is neglected.

As a first simplification, we approximate the recombination rate to be a linear function of the carrier density, namely $r(n) \approx \frac{n}{\tau_k}$ [90] where we let the carrier lifetime τ be a function of the wavelength. We then introduce the definitions $n_{ss} \triangleq \frac{I}{qV_a} \tau_k$ and $P_{sat,k} \triangleq \hbar \omega_k A_a / \tau_k$, where the latter quantity is called the saturation power [90]. Hence we obtain a differential equation for each wavelength λ_k in the form:

$$\tau_k \frac{\partial n(z, t)}{\partial t} = n_{ss} - n - \Gamma \sum_{r=1}^{N_{ch}} g_r(n) \frac{P_r(z, t)}{P_{sat,r}}. \quad (4.15)$$

A key step towards a simpler model is to approximate the material gain as a linear function of the carrier density [14, 94] as:

$$g_r(n) \approx \frac{1}{\Gamma} a_r (n - n_{0r}). \quad (4.16)$$

We thus have to solve:

$$\tau_k \frac{\partial n(z, t)}{\partial t} = n_{ss} - n - \sum_{r=1}^{N_{ch}} a_r (n - n_{0r}) \frac{P_r(z, t)}{P_{sat,r}} \quad (4.17)$$

coupled with:

$$\frac{\partial P_r(z, t)}{\partial z} = a_r (n - n_{0r}) P_r(z, t). \quad (4.18)$$

We found convenient introducing a z -independent aggregate carrier density:

$$\underline{n} \triangleq \int_0^{L_a} n(z, t) dz. \quad (4.19)$$

To let appear this factor in the rate equation, we first integrate the right-hand side of Eq. (4.17) over the distance:

$$\int_0^{L_a} a_r (n - n_{0r}) P_r(z, t) dz = P_r(0, t) (e^{\int_0^{L_a} a_r (n - n_{0r}) dz} - 1) \quad (4.20)$$

where the input-output signal powers are related by Eq. (4.18), whose solution is:

$$P_r(z, t) = e^{\int_0^z a_r(n - n_{0r}) d\zeta} P_r(0, t). \quad (4.21)$$

Then, we define the z -integrated gain (after the linear approximation) as:

$$h_r \triangleq \int_0^{L_a} a_r(n - n_{0r}) dz = a_r(\underline{n} - \underline{n}_{0r}). \quad (4.22)$$

The rate equation thus simplifies in the following expression:

$$\tau_k \frac{dn(t)}{dt} = \underline{n}_{ss} - \underline{n} - \sum_{r=1}^{N_{ch}} (e^{h_r} - 1) \frac{P_r(0, t)}{P_{sat, r}} \quad (4.23)$$

which can be rearranged in terms of h_k , by exploiting Eq. (4.22), as:

$$\tau_k \frac{dh_k(t)}{dt} = h_{0k} - h_k - a_k \sum_{r=1}^{N_{ch}} \varepsilon_r (e^{h_r} - 1) P_r(0, t) \quad (4.24)$$

with $h_{0k} \triangleq a_k(\underline{n}_{ss} - \underline{n}_{0k})$ and $\varepsilon_r \triangleq 1/P_{sat, r}$. Such a simple *integrated-gain* (IG) model is based on four wavelength-dependent parameters $\tau_k, h_{0k}, a_k, \varepsilon_k$ that can be used as fitting parameters to be properly tuned. The proposed model can be seen as an extended version of the model in [90], based on the integrated gain, where we included the wavelength dependency of the material gain. The proposed model also differs from the multi-stage reservoir model proposed in [94], where the rate equation is manipulated obtaining a wavelength-independent differential equation for the reservoir in each stage.

Similarly to [92, 94] we now include the spatial resolution via a multi-stage model. The idea of the multi-stage model is to use the integrated-gain model along a section, or stage, of length L_{st} and model the evolution along distance as a concatenation of stages. For the generic stage $s = 1, \dots, N_{st}$ we can write a rate equation in the form:

$$\tau_k \frac{dh_{ks}(t)}{dt} = h_{0ks} - h_{ks} - a_{ks} \sum_{r=1}^{N_{ch}} \varepsilon_{rs} (e^{h_{rs}} - 1) P_r(z_s, t) \quad (4.25)$$

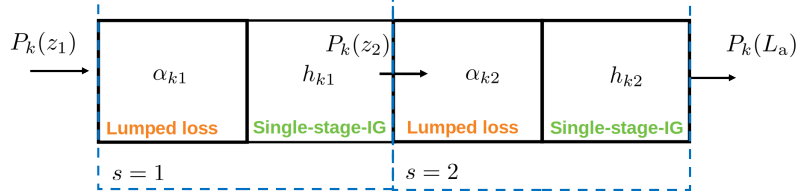


Figure 4.8: Sketch of the SOA modeled through a two-stage integrated gain h_k for wavelength λ_k with lumped losses.

being $P_r(z_s, t)$ the signal power at the input of the s th section at coordinate z_s , with $z_1 = 0$. We also found a benefit in including the SOA losses by means of a lumped attenuation α_{rs} , similarly to [92]:

$$\tau_k \frac{dh_{ks}(t)}{dt} = h_{0ks} - h_{ks} - a_{ks} \sum_{r=1}^{N_{ch}} \epsilon_{rs} e^{-\alpha_{rs} L_{st}} (e^{h_{rs}} - 1) P_r(z_s, t) \quad (4.26)$$

where the loss α_{rs} is stage and wavelength dependent, and is treated as another model fitting parameter. The idea is sketched in Fig. 4.8 for two stages.

4.3.1 Numerical implementation

The right-hand side of the rate equation in Eq. (4.26) at steady-state can be written as

$$D(\bar{h}_{ks}) = h_{0ks} - \bar{h}_{ks} - a_{ks} \sum_{r=1}^{N_{ch}} \epsilon_{rs} e^{-\alpha_{rs} L_{st}} (e^{\bar{h}_{rs}} - 1) P_r(z_s) \quad (4.27)$$

for each stage $s = 1, \dots, N_{st}$, with \bar{h} the time-averaged integrated gain. Solving the steady-state model means searching for time averaged and space integrated gain \bar{h}_{ks} such that $D(\bar{h}_{ks}) = 0$. The problem can be solved by iteratively adjusting the value of \bar{h}_{ks} as in the algorithm described in Sec. 4.1.1.

The steady-state wavelength and stage-dependent model fitting parameters $(\epsilon_{ks}, h_{0ks}, a_{ks}, \alpha_{ks})$ are found by minimizing the least square error, over a set of N_{pow} different input signal power, between the per-stage gain of the integrated-gain (IG) model G_{ks}^{IG} and the gain G_{ks}^C computed with the Connelly model in the same SOA chunk.

We performed the fitting step by means of the MATLAB function *lsqcurvefit*, which iteratively adjusts the values of $(\varepsilon_{ks}, h_{0ks}, a_{ks}, \alpha_{ks})$ until the condition $e_{ks} < \theta_e$ is met, being e_{ks} the least-squares error and θ_e a threshold. The flowchart of the adopted algorithm is shown in Fig. 4.9. The numerical solution for the single-stage integrated model follows the same flowchart with $N_{st} = 1$.

We now search for the best τ in a least-squares sense. For each stage s the time evolution in Eq. (4.26) is solved by means of the RK4 method, as in Section 4.1.1, starting from the steady-state integrated gain. The algorithm stops when the error, in least-squares sense, between the SOA output signal power obtained with the IG model and the Connelly model, is minimized.

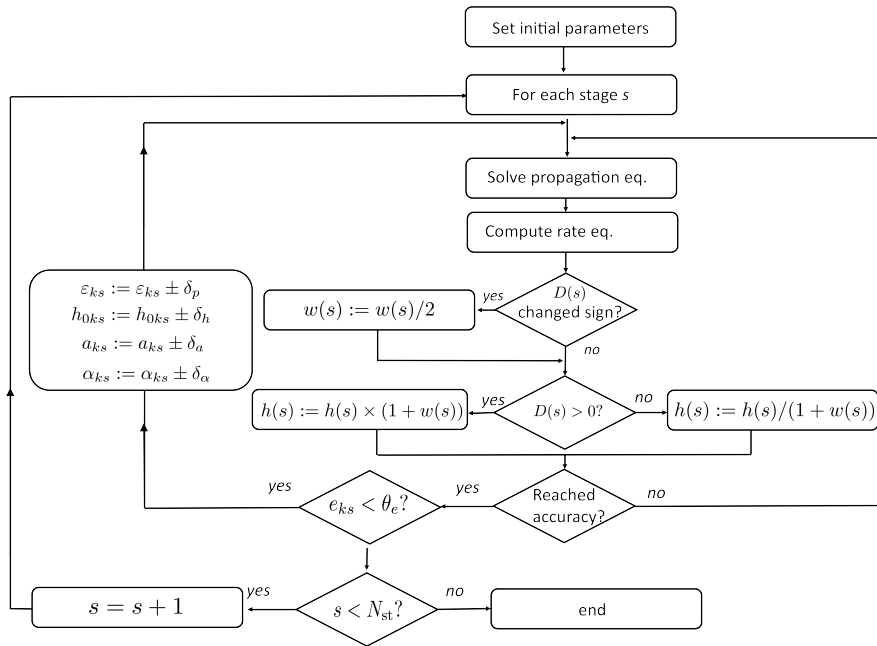


Figure 4.9: Steady-state algorithm for best parameter settings of the proposed integrated-gain model.

4.3.2 Numerical results

In this section, we test the accuracy of the proposed integrated-gain model, first in its simplest single-stage form, and then in the multi-stage version with the inclusion of lumped losses.

Single-stage IG model without losses

At first, we tested the single-stage IG model without lumped losses, which is governed by Eq. (4.24). The model parameters are found by least-squares error fitting as per Sec. 4.3.1. Figure 4.10 shows the static gain obtained with the BNM model versus the total input power, measured at three wavelengths of interest out of 7 transmitted wavelengths, for the same setup of Fig. 4.6. Dashed lines represent the gain computed with the IG model. It can be seen that the matching between the BNM and the IG model is excellent in the linear regime while it becomes poor in the highly nonlinear regime.

Such a poor performance of the IG model in the nonlinear regime results also in

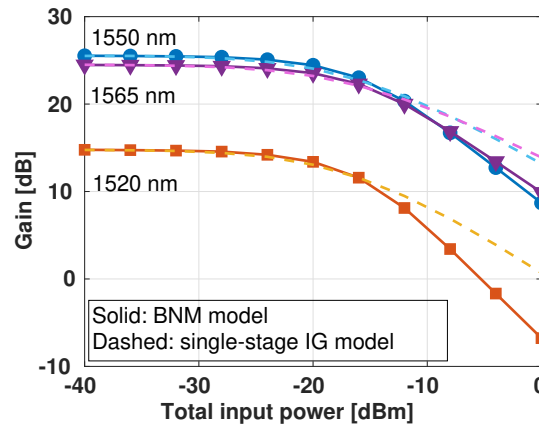
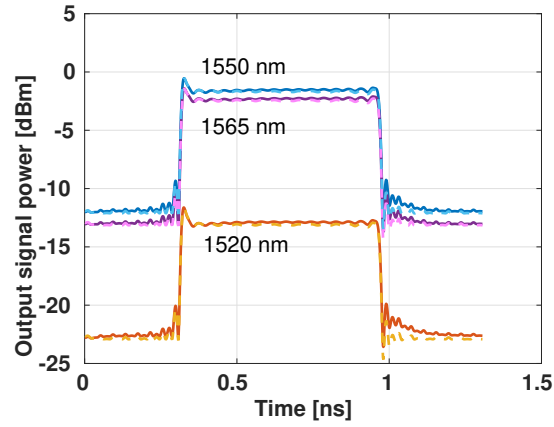
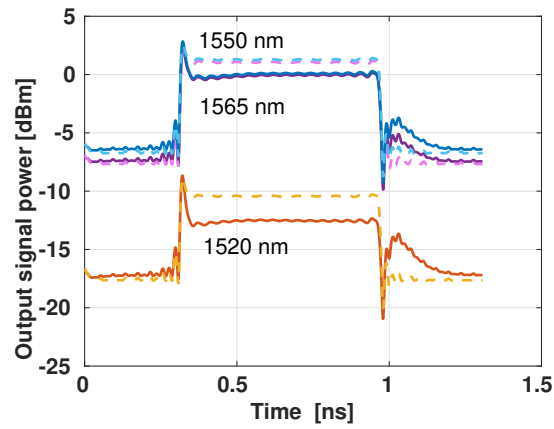


Figure 4.10: Steady-state gain vs. total input power measured at 1520 nm, 1550 nm, and 1565 nm, for the 7-channel case of Fig. 4.6. Solid: BNM SOA model. Dashed: single-stage IG model without losses fitted to BNM model.



(a)



(b)

Figure 4.11: Output signal power vs. time, for the 7-channel case of Fig. 4.6, measured at 1520 nm, 1550 nm, and 1565 nm. Solid: BNM SOA model. Dashed: lossless single-stage IG model without losses fitted to BNM model. Total input power (a) -16 dBm and (b) -10 dBm.

an inaccurate description of the signal evolution in time, at high input power. As an example, we considered the transmission of OOK modulated signals, with rectangular supporting pulses of 0.64 ns. The evolution in time and space is solved by means of the BNM model and the single-stage IG model without losses.

Figure 4.11 shows the signal power versus time at the output of the amplifier for the wavelengths of interest at total input power (a) -16 dBm and (b) -10 dBm. From the figure, a characteristic overshoot at the rising edge can be observed. This behavior is due to the fact that temporal variations in the input signal result in temporal fluctuations in the SOA gain. In particular, when the SOA operates in the nonlinear regime, the gain of the amplifier is a function of the input power, as previously discussed. Hence, temporal oscillations in the signal power yield a transient in the temporal dynamics of the gain. As a consequence, at the rising edge, the gain transient manifests an overshoot, which can be observed for all the wavelengths, and is more severe for high signal power values, as shown in the bottom part of the figure.

In Fig. 4.11 we also observe an excellent match between the lossless IG-model and the BNM model in the low power regime, where the static fitting captured the full SOA behavior. This is not true in the nonlinear regime with an error up to 2 dB. Therefore, the single-stage IG model cannot be considered reliable in the SOA nonlinear regime.

Unidirectional multi-stage IG model with lumped losses

The investigations carried out in Sec. 4.2 showed that the slope of the static gain curve is affected by the spatial resolution. We are thus motivated to investigate the multi-stage IG model. Such a model includes the spatial resolution by means of stage concatenation, as discussed in Sec. 4.3, as if several short amplifiers were concatenated to build the SOA, provided that the signal in the backward propagation direction is negligible. To account for the SOA distance-dependent absorption, the multi-stage IG model also includes a stage-dependent lumped loss.

We found that a stage-dependent fitting is essential to capture the correct space evolution of the gain along the SOA length. In addition, we accounted for the carrier density-dependent material absorption by including a stage-dependent lumped loss

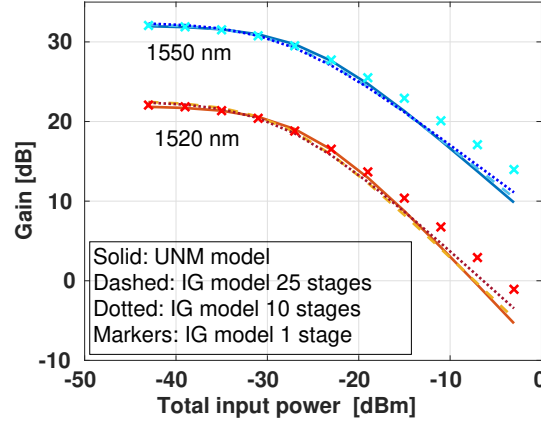


Figure 4.12: Steady-state gain vs. total input power. 2 channels WDM transmission at $\lambda_1 = 1520$ nm and $\lambda_2 = 1550$ nm. UNM model (solid), and multi-stage IG model with lumped losses and 25 (dashed), 10 (dotted), and 1 (markers) stages.

in the set of static fitting parameters. It is worth noting that a stage-dependent fitting approach requires the knowledge of the true amplifier gain at different coordinates. In this work, these gain values are provided by numerical simulations of the reference SOA model (here the UNM model). Investigations on efficient stage-dependent parameters extraction from experimental measurements are left as future work.

Figure 4.12 shows the static gain curve as a function of the total input power measured at $\lambda_1 = 1520$ nm and $\lambda_2 = 1550$ nm. Since we assumed zero reflectivity, we now consider the UNM model as the reference model, whose results are reported with solid lines. Dotted and dashed lines refer to a 10-stage and 25-stage IG model, respectively. For completeness, markers represent the result of the single-stage IG model. Each stage of the IG model had the same length of m UNM model sections, with $m = 5$ for the 10-stage IG model, and $m = 2$ in the 25-stage case. It can be seen from the figure that the multi-stage model well captures the amplifier behavior at the two wavelengths under investigation, even in the high nonlinear regime, with better accuracy when a finer spatial resolution is used by means of 25 stages.

Regarding the dynamic behavior, Fig. 4.13 shows the signal output power at the

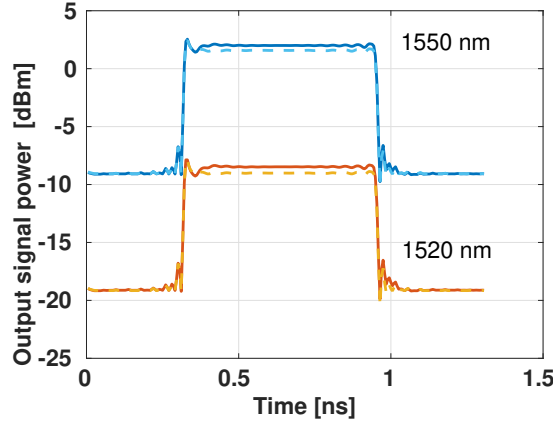


Figure 4.13: Output signal power vs. time for a 2 channel WDM transmission at $\lambda_1 = 1520$ nm and $\lambda_2 = 1550$ nm. Solid: UNM SOA model. Dashed: 25-stages IG model. Total input power -25 dBm.

two wavelengths when the SOA is described through the UNM model or the 25-stage IG model. The input power was -25 dBm, a value for which the amplifier exhibits a nonlinear behavior, see Fig. 4.12. It can be seen that the two models are in good agreement, showing a mismatch within 0.5 dB. Therefore, in the remainder of this work, we will focus on a 25-stage IG model.

4.4 SOA-amplified optical line

In this section, we investigate the performance of an SOA-amplified optical line in terms of SNR. We considered the transmission of 3 equally-spaced WDM channels spanning the C band, centered at 1520 nm, 1540 nm, and 1560 nm respectively. The modulation format was single-polarization QPSK, and the symbol rate was 49 Gbaud. In order to obtain accurate results, we transmitted sequences of 65536 symbols, which were enough to capture the maximum walk-off among channels. At the receiver side, chromatic dispersion compensation, matched filtering, and average carrier phase recovery were performed. At first, we calculate the SNR at the output of a single SOA.

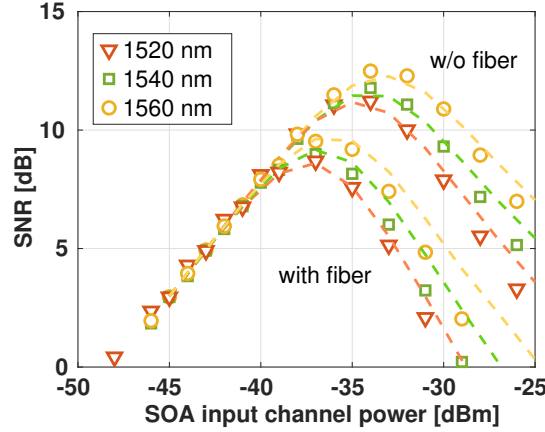


Figure 4.14: SNR vs. channel power at the SOA input, after 1 span. WDM signal with channels centered at 1520 nm, 1540 nm, and 1560 nm. Single-polarization QPSK modulation format. Markers: UNM SOA model. Dashed: 25-stage IG model.

Then, we concatenate SOAs both in the presence and absence of the optical fiber.

The field propagation inside the SOA was simulated by means of the UNM model, hence assuming zero reflectivity, or by the multi-stage IG model with lumped losses. For each input power value, both steady-state models were solved to identify the initial conditions for the dynamic evolution.

On the other hand, the propagation along the optical fiber was simulated through the SSFM with a separate-field point of view, and CLE criterion with maximum tolerable FWM phase shift in the first step [28] equal to 20 rad. In this work, we limited the analysis to SMFs whose parameters are the same as in Tab. 3.1.

In Fig. 4.15 we reported the SNR bell-curves as a function of the power at the SOA input, with and without the optical fiber before the amplifier, measured after the first SOA in the link. The power at the amplifier output is controlled such that the total average power is constant, and the power profile at the amplifier output is frequency-flat. The figure shows that the 25-stage IG model (dashed line) provides an excellent match up to the nonlinear threshold, i.e., up to the maximum SNR value. However, it can be seen that the SOA under test, having parameters as in Tab. A.1

	(a)	(b)	(c)
$B_{\text{rad}} [\text{m}^3\text{s}^{-1}]$	2.6×10^{-6}	2×10^{-6}	1.6×10^{-6}
$F [\text{dB}] @ 1520 \text{ nm}$	9.1	9.7	10.5
$F [\text{dB}] @ 1540 \text{ nm}$	9.2	9.9	10.9
$F [\text{dB}] @ 1560 \text{ nm}$	9.5	10.5	11.7

Table 4.1: Pseudo-linear SOAs characteristics.

of Appendix A, is not well-suited for long-haul transmissions. In fact, its nonlinear regime is triggered for very small input powers compared to typical powers of long-haul transmissions (about -5 to 0 dBm).

For this reason, we tested the performance of the so-called *pseudo-linear* SOAs, which differ from the SOA of Tab. A.1 in three parameters: the bias current $I = 60$ mA, the linewidth factor $\alpha_H = 0$, and the bimolecular radiative recombination B_{rad} . We considered three pseudo-linear SOAs having different values of B_{rad} : (a) 2.6×10^{-6} , (b) 2×10^{-6} and (c) $1.6 \times 10^{-6} \text{ m}^3\text{s}^{-1}$. Unfortunately, this linearization of the amplifier behavior implies an increase in the noise figure F compared to the SOA adopted in Fig. 4.14, which was 7 dB. The noise figure values are summarized in Tab. 4.1 for the three amplifiers under test.

The left-hand side of Fig. 4.15 shows the SNR after the concatenation of 10 amplifiers (without optical fibers) versus the transmitted channel power, using the UNM model. For the sake of comparison, we reported also the results obtained with EDFA-based amplification. The EDFA has the same noise figure as the pseudo-linear SOA. It can be seen that, in the absence of the optical fiber, the EDFA provides a linear SNR over the whole power range, while for the SOA the nonlinear regime is triggered at different powers according to the B_{rad} value. Since we are considering pseudo-linear SOAs with the aforementioned parameters, the nonlinear regime is characterized by

a mild SNR decrease typical of a bell-curve near the NLT.

The right-hand side of Fig. 4.15 shows also the SNR curves in the presence of an optical fiber in each span. In this case, we increased the transmitted power by 20 dB in order to have the same power at the amplifier input as in the case without optical fibers. It can be seen that the NLT in the bell-curves with EDFA-based amplification is around 2 dBm, hence -18 dBm at the amplifier input. We note from the left part of the figure that the nonlinear regime is triggered approximately at -22 , -20 and -18 dBm of input power, respectively, for the cases (a), (b), and (c). As a consequence, the SNR values of the SOA-based amplification are smaller than those estimated for the EDFA-based link around the NLT, for the cases (a) and (b).

The gap between the SOA and EDFA results can be reduced considering even more linear SOA devices whose nonlinear regime is triggered at higher input powers. For the most linear SOA under test, the SOA nonlinearities are triggered at input power levels comparable to the NLT of the EDFA-based SNR curve. Hence, for scenario (c) in Fig. 4.15, the bell-curves are almost overlapped. In the limit case in which the SOA operates in a linear regime across the whole range of power values under test, the SNRs curves coincide. However, the linear SOA obtained with this procedure has a noise figure that is too high for transmissions purposes, and it serves only as a sanity check.

Solving the algorithm in Sec. 4.1.1 for each amplifier within the optical line provides an accurate estimation of the system performance, yet time-consuming. We are thus motivated to investigate the accuracy of the IG model for a multi-span link. Although the initial fitting stage requires solving the reference UNM model for the extraction of the model parameters, this step is done only once and the computational effort required to simulate the propagation along the multi-span link is then reduced by a factor of 3 compared to the UNM model.

Figure 4.16 shows a comparison between the SNR obtained with the UNM model and the multi-stage IG model, for the pseudo-linear SOA (c), hence the most linear SOA under test. We reported with dashed lines the results obtained with the 10 and 25-stage IG model with lumped losses, wherein the static and dynamic parameters are extracted in advance as discussed in Sec. 4.3.1. The top image of Fig. 4.17 shows

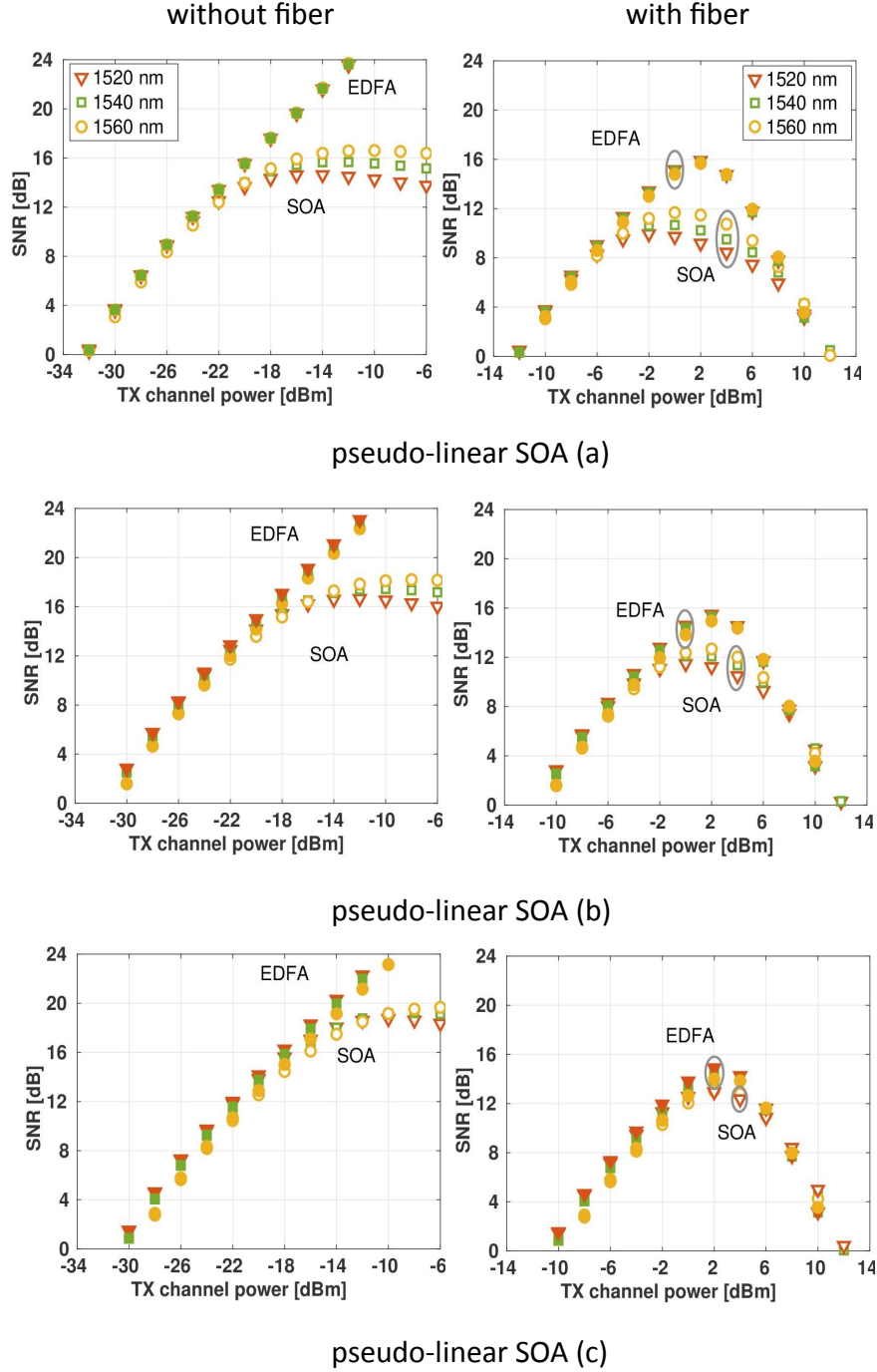


Figure 4.15: SNR vs. transmitted channel power. 10×100 km SMFs link, amplified with pseudo-linear SOAs (empty markers) or EDFA (full markers). (a), (b) and (c): different kind of pseudo-linear SOAs, see Tab. 4.1.

a very good agreement between the UNM model and the 25-stage IG model for all the wavelengths. On the other hand, we note that reducing the number of stages to 10 provides worse matching, as already observed in Fig. 4.12, although it allows slightly faster simulations.

Finally, for the sake of completeness, in Fig. 4.17 we report the results obtained with the UNM model and the 25-stage IG, with and without the optical fiber, for

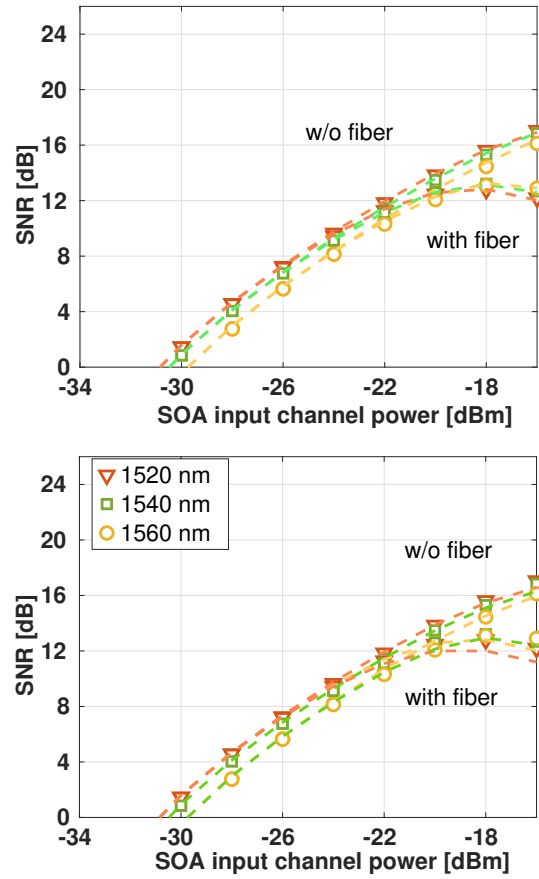


Figure 4.16: SNR vs. amplifier input channel power. 10×100 km SMFs link with pseudo-linear SOA (c) as per Tab. 4.1. Markers: UNM model. Dashed: Multi-stage IG model with 25 (top) and 10 stages (bottom).

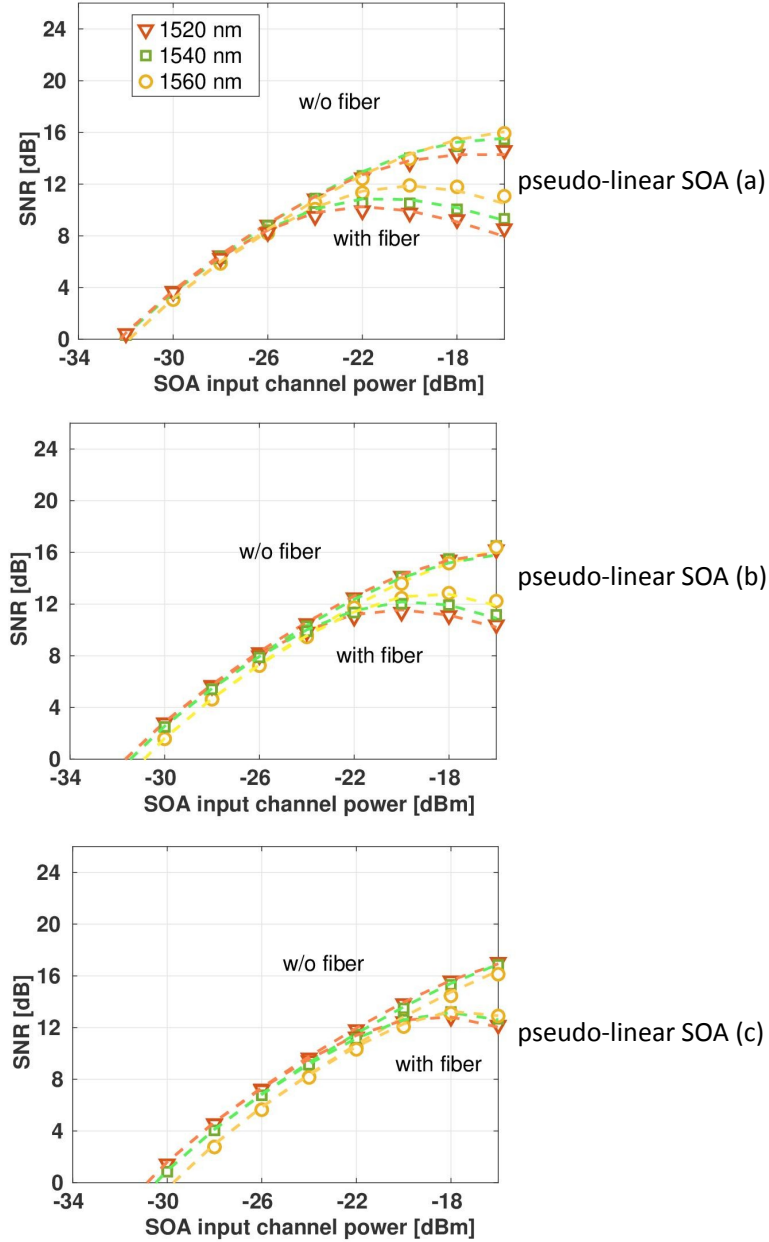


Figure 4.17: SNR vs. transmitted channel power. 10×100 km SMFs link with pseudo-linear SOA amplification, see Tab. 4.1. Markers: UNM model. Dashed: 25-stage IG model with a pre-run for fitting.

the three pseudo-linear SOAs in Tab. 4.1. We reported the SNR as a function of the channel power at the amplifier input in a range between -34 dBm and -16 dBm, which corresponds to channel powers at the amplifier output between -14 dBm and 4 dBm, respectively. The figure shows that, for realistic power values, the proposed multi-stage IG model with lumped losses can be safely used for the SNR estimation of an optical link based on pseudo-linear SOA with zero reflectivity.

Part II

Spatial multiplexing

Chapter 5

EGN model with polarization dependent loss

In spatially multiplexed links the polarizations of the propagating field generally experience different effects. The problem is present even in dual-polarization transmissions. A particular polarization-related problem is due to the losses introduced by an optical device, which can depend on the state of polarization of the input field. Such a phenomenon is known as polarization-dependent loss (PDL) [8], and manifests itself as an unequal loss of energy. In particular, contrary to the optical fiber, the PDL introduced by lumped devices such as EDFAs and WSSs is typically non-negligible [95]. As amplifiers and switches densely populate the optical networks, it follows that the accumulation of PDL along the distance can be particularly detrimental in dual-polarization transmissions.

In optical communication systems, the PDL experienced by the signal can be equalized at the receiver through digital-signal processing [96,97]. However, the PDL accumulated along the link in the ASE and the NLI partially remains. The distributed generation of the two effects along the distance results in different PDL impairments. As a result, the estimation of the system performance in terms of SNR must account for PDL even in the presence of equalization. Since PDL randomly fluctuates over time, it might not be sufficient to assess the average system performance, since for

some PDL realizations the SNR can fall below a given threshold causing the system outage. The performance analysis should thus focus on such an outage probability. Because of the random fluctuations, the problem of estimating such a probability is particularly challenging, especially in numerical simulations where low outage values call for many time-consuming simulations, but also in experiments where collecting many observations may require a huge amount of resources to save and post-process the results.

Such difficulties in the numerical complexity stimulated the development of theoretical models for quick estimation of the PDL effects. Most of the literature focused on the interplay between PDL and ASE noise in the linear regime [17, 98–101]. On the other hand, the interplay between PDL and the fiber nonlinear Kerr effect has been overlooked and received much less attention, and most of the literature focused on numerical/experimental investigations [102–109], with contrasting results.

To analytically handle the impact of PDL on NLI in a simple way, the first choice is to include PDL in the perturbative models introduced in Chapter 2. The scalar theory of the GN model was extended to include polarization effects in [110], and first included PDL in the GN model framework in [111] for a quick estimation of the PDF of the NLI.

In this chapter, besides providing a novel mathematical formalism to cope with PDL in the GN model, we derive a PDL-aware EGN model, hence including PDL even in the FON and HON contributions to the NLI. The general theory will be double-checked against SSFM simulations. The work presented in this chapter was partially published in [43].

5.1 Polarization-dependent loss

The PDL is defined as the ratio of the maximum and minimum transmission intensities [8, 17]:

$$\zeta = \frac{\mathcal{R}_{\max}}{\mathcal{R}_{\min}} \quad (5.1)$$

where the transmission coefficient \mathcal{R} of an optical device is given by the product

$\mathcal{R} \triangleq \langle s^\dagger | \mathbf{T}^\dagger \mathbf{T} | s \rangle$ with \mathbf{T} the PDL matrix of the device, and $|s\rangle$ the SOP of the field at the device input. The PDL matrix \mathbf{T} can be described by a matrix exponential, whose expression is given in [8]:

$$\mathbf{T} = e^{-\alpha_p/2} \exp\left(\frac{\hat{\alpha}_p \cdot \vec{\sigma}}{2}\right) \quad (5.2)$$

where $\vec{\alpha}_p = \hat{\alpha}_p \alpha_p$ is the 3D PDL vector in the Stokes space [8]. α_p represents the PDL loss common to the two polarization while $\hat{\alpha}_p$ is a unit Stokes vector pointing in the direction for maximum transmission. The vector $\vec{\sigma}$ in Eq. (5.2) is the well-known spin vector of Pauli matrices [8]. Such a PDL matrix is typically expanded in the following form:

$$\mathbf{T} = e^{-\alpha_p/2} \left(\cosh\left(\frac{\alpha_p}{2}\right) \mathbf{I} + \sinh\left(\frac{\alpha_p}{2}\right) (\hat{\alpha}_p \cdot \vec{\sigma}) \right). \quad (5.3)$$

Applying the definition of transmission coefficient, it follows that:

$$\mathcal{R} = \frac{1}{1+\Gamma} (1 + \Gamma (\hat{\alpha}_p \cdot \hat{s})) \quad (5.4)$$

where we defined $\Gamma \triangleq \tanh \alpha_p$ as in [8]. Such a parameter is related to the extrema of transmission through $(\mathcal{R}_{\max} - \mathcal{R}_{\min})/(\mathcal{R}_{\max} + \mathcal{R}_{\min})$. From Eq. (5.4) it can be seen that the maximum transmission occurs when the input SOP is aligned with the PDL maximum axes, hence $\hat{\alpha}_p \cdot \hat{s} = 1 \rightarrow \mathcal{R}_{\max} = 1$. On the other hand, the minimum is such that: $\hat{\alpha}_p \cdot \hat{s} = -1 \rightarrow \mathcal{R}_{\min} = e^{-2\alpha_p}$. It is common in the literature to express the PDL as a function of the parameter Γ in a dB scale [107, 108], namely:

$$\zeta_{\text{dB}} = 10 \log_{10} \left(\frac{1+\Gamma}{1-\Gamma} \right). \quad (5.5)$$

5.2 The PDL-EGN model

In the presence of PDL, the linear operator in Eq. (1.52) must be modified as:

$$\mathcal{F}\left\{e^{\mathcal{L}z}\right\} = e^{-j\left(\frac{\beta_2}{2}\omega^2 + \frac{\beta_3}{6}\omega^3\right)z}\mathbf{U}(z) \quad (5.6)$$

$$\mathcal{F}\left\{e^{-\mathcal{L}z}\right\} = e^{j\left(\frac{\beta_2}{2}\omega^2 + \frac{\beta_3}{6}\omega^3\right)z}\mathbf{U}^{-1}(z) \quad (5.7)$$

where, in this chapter, the matrix \mathbf{U} accounts only for a frequency-independent PDL accumulated up to coordinate z . We assume that PDL is introduced only by lumped devices, such as EDFA and WSS, placed at coordinates $z_p : p = 0, \dots, N-1$, with $z_0 = 0$ and N the total number of PDL blocks. The average power loss induced by PDL is recovered after each block by an equalizer having gain G_E , such that the matrix \mathbf{U} depends on the k th device at coordinate $z_k \leq z$ with PDL matrix \mathbf{M}_k by [8]:

$$\mathbf{U}(z) = \mathbf{M}_p \mathbf{M}_{p-1} \cdots \mathbf{M}_0, \quad z_p < z < z_{p+1} \quad (5.8)$$

$$\mathbf{M}_k \triangleq G_{E,k} \mathbf{T}_k = \mathbf{W}_k^\dagger \begin{bmatrix} \sqrt{1+\Gamma_k} & 0 \\ 0 & \sqrt{1-\Gamma_k} \end{bmatrix} \mathbf{W}_k \quad (5.9)$$

where the \mathbf{W}_k are matrices uniformly distributed in the set of the 2×2 unitary random matrices (Haar matrices), and \mathbf{T} is the PDL matrix before applying the equalizer gain, as in Eq. (5.3). With many identically distributed PDL elements, the resulting PDL of the link, expressed in dB, follows a Maxwellian distribution [99] with an average value scaling with \sqrt{N} .

In the absence of noise, the power imbalance between polarizations due to PDL can be removed at the receiver side. For instance, a zero-forcing equalizer can be applied. Unfortunately, PDL remains in both the ASE and the NLI after the equalization. Focusing on the ASE noise for the sake of simplicity, the motivation can be grasped from Fig. 5.1. The figure sketches the PDL effects on the signal and ASE noise along a 3 span link with 3 lumped PDL elements. In particular, in the figure the zero-forcing equalization is represented as a round-trip propagation of the signal, thus returning equal to itself at reception. However, it can be seen that ASE follows an

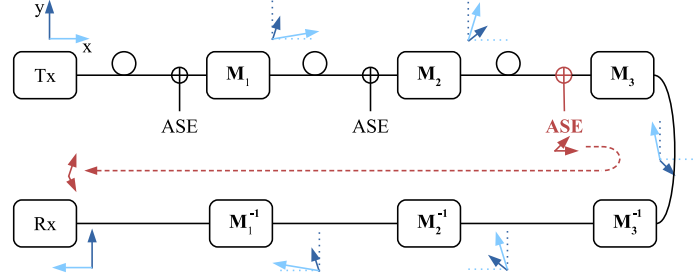


Figure 5.1: Sketch of the effects experienced by the signal $|A\rangle$ and the ASE $|w\rangle$ along propagation, including zero-forcing equalization by the cascade of \mathbf{M}_k^{-1} matrices, for a link having 3 lumped PDL elements.

incomplete round-trip since it is introduced in the link in a distributed way, resulting in residual PDL.

After linear equalization, the received ASE $|\tilde{w}(\omega)\rangle$ is related to the ASE $|\tilde{w}_m(\omega)\rangle$ emitted by in-line amplifiers at coordinate z_m by:

$$|\tilde{w}(\omega)\rangle = \sum_{m=1}^M \mathbf{M}_0^{-1} \mathbf{M}_1^{-1} \cdots \mathbf{M}_{m-1}^{-1} |\tilde{w}_m(\omega)\rangle \quad (5.10)$$

with M the number of amplifiers. In the example depicted in Fig. 5.1 where the transmitted has no PDL, i.e., $\mathbf{M}_0 = \mathbf{I}$, we have:

$$\begin{aligned} |\tilde{w}(\omega)\rangle &= |\tilde{w}_1(\omega)\rangle + \mathbf{M}_1^{-1} |\tilde{w}_2(\omega)\rangle + \mathbf{M}_1^{-1} \mathbf{M}_2^{-1} |\tilde{w}_3(\omega)\rangle \\ &= |\tilde{w}_1(\omega)\rangle + \mathbf{U}(z_1)^{-1} |\tilde{w}_2(\omega)\rangle + \mathbf{U}(z_2)^{-1} |\tilde{w}_3(\omega)\rangle . \end{aligned} \quad (5.11)$$

Regarding the NLI, Eq. (2.2) must be generalized to include PDL. The derivation follows the same idea of Chapter 2, with the proper linear operator defined in Eq. (5.6) hence obtaining:

$$\begin{aligned} |\tilde{n}(\omega)\rangle &= -j \frac{8}{9} \gamma \iint_{-\infty}^{\infty} \sum_{p=0}^{N-1} \eta^{(p)}(\omega, \omega_1, \omega_2) \\ &\quad \times \langle \tilde{A}(\omega + \omega_1 + \omega_2) | \mathbf{P}(z_p) | \tilde{A}(\omega + \omega_2) \rangle | \tilde{A}(\omega + \omega_1) \rangle \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} \end{aligned} \quad (5.12)$$

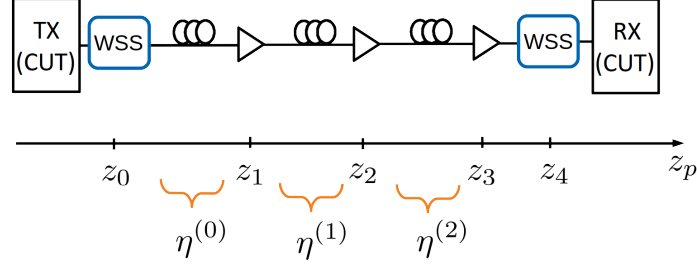


Figure 5.2: Sketch of a 3 spans link having 5 PDL elements (3 amplifiers, 2 WSS). Bottom: PDL elements coordinates identifying link segments. Each segment k containing an optical fiber is associated with a fiber kernel $\eta^{(k)}$. Segments containing only PDL blocks have zero kernel.

where $\eta^{(p)}$ is the kernel of the optical fiber in the link segment $[z_p, z_{p+1}]$. To compact the notation, we introduced the following definition:

$$\mathbf{P}(z_p) \triangleq \mathbf{U}^\dagger(z_p) \mathbf{U}(z_p). \quad (5.13)$$

Note that the in Eq. (5.12) we exploited the relationship $\mathbf{I} = \mathbf{U}^{-1}(z_p) \mathbf{U}(z_p)$ which holds only for frequency-independent matrices \mathbf{U} .

As an example to clarify the notation, Fig. 5.2 sketches a 3 span link, which can be seen as a segment of a wider network. Such a link is composed of 5 PDL elements: 3 inline amplifiers and 2 WSS for add/drop operation. Hence, in this example, $p = 0, \dots, 4$ with PDL element coordinates from z_0 to z_4 , highlighted in the bottom part of the figure. It is worth noting that not every segment $[z_p, z_{p+1}]$ contains an optical fiber, due to the presence of cascades of PDL elements, as the segment $[z_3, z_4]$ in Fig. 5.2. Since the segments contain only lumped elements, they do not contribute to the kernel, hence they have $\eta^{(p)} = 0$ in Eq. (5.12). Figure 5.2 highlights the three segments having an optical fiber and hence non-zero $\eta^{(p)}$. In the example under test, the overall link kernel is $\sum_p \eta^{(p)} = \eta^{(0)} + \eta^{(1)} + \eta^{(2)}$, where each $\eta^{(k)} = \eta_1 e^{j\Delta\beta kL}$ for segments including an optical fiber, with η_1 the single-span kernel in Eq. (2.16).

Both the NLI and the ASE accumulate linearly with the propagation distance, both being additive under the model assumptions. Moreover, by comparing Eq. (5.10)

and Eq. (5.12), we observe that after zero-forcing equalization they both depend on the PDL accumulated before their generation. However, such a dependence follows completely different relationships. In particular, the quadratic dependence on the entries of matrix $\mathbf{U}(z_p)$ in the NLI is expected to induce larger random fluctuations of the SNR compared to the linear dependence in the ASE case.

By following the same steps as in Chapter 2, we get the following discrete-time channel model relating the transmitted atom a_i to the received one

$$u_i = a_i + w_i + n_i. \quad (5.14)$$

The sampled ASE is defined as

$$w_i = \int_{-\infty}^{\infty} \langle \tilde{G}_i(\omega) | \tilde{w}(\omega) \rangle \frac{d\omega}{2\pi} \quad (5.15)$$

while the sampled NLI n_i is given in Eq. (2.8), with the only difference that the function $\mathcal{X}_{\text{hmni}}$ weighting the mixing at the symbol level defined in Eq. (2.9) must be generalized to include PDL, consistently with Eq. (5.12). The novel expression of this function is given by:

$$\begin{aligned} \mathcal{X}_{\text{hmni}} &= -j \frac{8}{9} \left(\frac{1}{\sqrt{2}} \right)^3 \gamma \sum_{p=0}^{N-1} \iiint_{-\infty}^{\infty} \eta^{(p)}(\omega, \omega_1, \omega_2) \\ &\quad \times \langle \tilde{G}_h(\omega + \omega_1 + \omega_2) | \mathbf{P}(z_p) | \tilde{G}_m(\omega + \omega_2) \rangle \\ &\quad \times \langle \tilde{G}_i(\omega) | \tilde{G}_n(\omega + \omega_1) \rangle \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} \frac{d\omega}{2\pi} \\ &= \left(\frac{1}{2\sqrt{2}} \right) \sum_{p=0}^{N-1} P_{h_3 m_3}(z_p) \delta_{i_3 n_3} \mathcal{S}_{\text{hmni}}(z_p). \end{aligned} \quad (5.16)$$

where the δ indicates Kronecker's delta. In the final identity we expanded the tensor in terms of the function $\mathcal{S}_{\text{hmni}}$ weighting the FWM interaction at the scalar level in the segment $[z_p, z_{p+1}]$:

$$\begin{aligned} \mathcal{S}_{\mathbf{h}\mathbf{m}\mathbf{n}\mathbf{i}}(z_p) \triangleq & -j\frac{8}{9} \iiint_{-\infty}^{\infty} \eta^{(p)}(\omega, \omega_1, \omega_2) \tilde{G}_{h_1 h_2}^*(\omega + \omega_1 + \omega_2) \\ & \times \tilde{G}_{m_1 m_2}(\omega + \omega_2) \tilde{G}_{i_1 i_2}^*(\omega) \tilde{G}_{n_1 n_2}(\omega + \omega_1) \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} \frac{d\omega}{2\pi} \end{aligned} \quad (5.17)$$

where $\mathcal{S}_{\mathbf{h}\mathbf{m}\mathbf{n}\mathbf{i}}(z_p)$ is a segment of the scalar function $\mathcal{S}_{\mathbf{h}\mathbf{m}\mathbf{n}\mathbf{i}}$ in Eq. (2.9), hence $\sum_p \mathcal{S}_{\mathbf{h}\mathbf{m}\mathbf{n}\mathbf{i}}(z_p) = \mathcal{S}_{\mathbf{h}\mathbf{m}\mathbf{n}\mathbf{i}}$. It is worth noting that the presence of the term $P_{h_3 m_3}(z_p)$ does not allow to express the link kernel in a closed-form as in Eq. (2.15). Nevertheless, the term $P_{h_3 m_3}(z_p)$ is frequency independent and can be factored out of the frequency integrals.

5.2.1 ASE variance

Several efforts have been made in the literature to address the PDL impact on ASE noise, for instance [8, 17, 101, 102]. We now review the key results regarding the ASE noise variance, consistently with the adopted notation.

By definition, the variance of the ASE atom \mathbf{i} is $\sigma_{\text{ASE}}^2 = \mathbb{E}[w_{\mathbf{i}} w_{\mathbf{i}}^*]$. From Eq. (5.15), we start from the following object:

$$\begin{aligned} w_{\mathbf{i}} w_{\mathbf{i}}^* &= \iint_{-\infty}^{\infty} \langle \tilde{G}_{\mathbf{i}}(\mu) | \tilde{w}(\mu) \rangle \langle \tilde{w}(\omega) | \tilde{G}_{\mathbf{i}}(\omega) \rangle \frac{d\omega}{2\pi} \frac{d\mu}{2\pi} \\ &= \sum_{m,p} \iint_{-\infty}^{\infty} \langle \tilde{G}_{\mathbf{i}}(\mu) | \mathbf{U}^{-1}(z_{m-1}) | \tilde{w}_m(\mu) \rangle \\ &\quad \times \langle \tilde{w}_p(\omega) | (\mathbf{U}^{-1}(z_{p-1}))^\dagger | \tilde{G}_{\mathbf{i}}(\omega) \rangle \frac{d\omega}{2\pi} \frac{d\mu}{2\pi}. \end{aligned} \quad (5.18)$$

The ASE sources along the link are assumed independent and identically distributed, with [112]:

$$\mathbb{E} [|\tilde{w}_m(\omega)\rangle \langle \tilde{w}_p(\mu)|] = \frac{N_0}{2} \delta(\omega - \mu) \delta_{mp} \mathbf{I} \quad (5.19)$$

where the two δ indicate Dirac/Kronecker's delta, respectively, while $N_0 = h\nu FG$ is the one-sided, dual-polarization, PSD of ASE per amplifier. Therefore, taking the expectation \mathbb{E} of Eq. (5.18) yields the following expression of the ASE variance for the generic $i_3 \in (x, y)$:

$$\sigma_{\text{ASE}i_3}^2 = \frac{N_0 B}{2} \sum_{p=1}^M P_{i_3 i_3}^{-1}(z_p). \quad (5.20)$$

Please note that the matrices $\mathbf{P}(z_p)$ and hence their elements $P_{i_3 j_3}(z_p)$ are not independent but related by the concatenation rule of PDL [17].

5.2.2 NLI variance

The general expression of the NLI variance derived in Eq. (2.28) holds even in the presence of PDL. In this section, we will derive expressions for the GN, F4, Q4, and Q6 variances by applying the novel PDL-aware tensor to the theoretical framework of Chapter 2.

GN term

In order to extend the GN term variance expression derived in Eq. (2.30) to the presence of PDL, we start from the general expression in terms of the $\mathcal{X}_{\mathbf{h}\mathbf{m}\mathbf{n}\mathbf{i}}$. We then substitute to this tensor its expression in Eq. (5.16), namely:

$$\begin{aligned} \sigma_{\text{GN}i_3}^2 &= \kappa_{1;1}^3 \sum_{\mathbf{h}, \mathbf{m}, \mathbf{n}} \mathcal{X}_{\mathbf{h}\mathbf{m}\mathbf{n}\mathbf{i}} (\mathcal{X}_{\mathbf{h}\mathbf{m}\mathbf{n}\mathbf{i}}^* + \mathcal{X}_{\mathbf{h}\mathbf{n}\mathbf{m}\mathbf{i}}^*) \\ &= \frac{1}{8} \kappa_{1;1}^3 \sum_{h_1, m_1, n_1} \sum_{h_2, m_2, n_2} \sum_{p, \ell=0}^{N-1} \mathcal{S}_{\mathbf{h}\mathbf{m}\mathbf{n}\mathbf{i}}(z_p) \mathcal{S}_{\mathbf{h}\mathbf{m}\mathbf{n}\mathbf{i}}^*(z_\ell) \\ &\quad \times \left(\sum_{h_3, m_3} P_{h_3 m_3}(z_p) P_{h_3 m_3}^*(z_\ell) + \sum_{h_3} P_{h_3 i_3}(z_p) P_{h_3 i_3}^*(z_\ell) \right) \\ &= \sum_{p, \ell=0}^{N-1} \mathcal{C}_{\text{GN}}(p, \ell) \left(\sum_{h_3, m_3} P_{h_3 m_3}(z_p) P_{h_3 m_3}^*(z_\ell) + \sum_{h_3} P_{h_3 i_3}(z_p) P_{h_3 i_3}^*(z_\ell) \right) \end{aligned} \quad (5.21)$$

where the scalar cross-correlation $\mathcal{C}_{\text{GN}}(p, \ell)$ is the same as in Chapter 2 in the absence of PDL, here reported for the reader convenience:

$$\begin{aligned}
\mathcal{C}_{\text{GN}}(p, \ell) &\triangleq \left(\frac{8}{81} \right) \frac{1}{T^3} \gamma^2 \kappa_{1;1}^3 \sum_{h,m,n} \iiint_{-\infty}^{\infty} \eta^{(p)}(\omega, \omega_1, \omega_2) \left(\eta^{(\ell)}(\omega, \omega_1, \omega_2) \right)^* \\
&\quad \times |\tilde{p}_h(\omega + \omega_1 + \omega_2 - \omega_h)|^2 |\tilde{p}_m(\omega + \omega_2 - \omega_m)|^2 \\
&\quad \times |\tilde{p}_i(\omega - \omega_i)|^2 |\tilde{p}_n(\omega + \omega_1 - \omega_n)|^2 \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} \frac{d\omega}{2\pi}.
\end{aligned} \tag{5.22}$$

It is worth noting that, in the absence of PDL, the GN term variance in Eq. (2.30) involves products between identity matrices, resulting in a weighting factor 3. In the presence of PDL, this weighting factor generalizes to the expression in Eq. (5.21).

We note that the absence of the modulus square of the kernel $|\eta|^2$ is induced by the presence of PDL. Namely, each cross-correlation $\mathcal{C}_{\text{GN}}(p, \ell)$ must be evaluated separately and thus a spatial disaggregation cannot be used like in the scalar case. Fortunately, the PDL-dependent part can be factored out of the frequency integrals. From the numeric point of view, the evaluation of each PDL-independent $\mathcal{C}_{\text{GN}}(p, \ell)$ is straightforward by using the methods described in Chapter 2. These terms can be pre-computed since they do not depend on the PDL realization. Once the $\mathcal{C}_{\text{GN}}(p, \ell)$ are available, the impact of PDL reduces to the computation of the 2×2 matrices \mathbf{P} . Such a weight must be evaluated for several PDL realizations. Thanks to its simple expression, it can be done very quickly.

The GN-term variance result can be easily generalized to the spatial-covariance matrix. We introduce the 2×2 GN covariance matrix between the polarizations, \mathbf{K}_{GN} : $K_{i_3 j_3}^{(\text{GN})} = \mathbb{E}[n_i' n_j'^*]$, $i_{1,2} = j_{1,2}$, $(i_3, j_3) \in (x, y)$, which takes the elegant form¹ [111]:

$$\mathbf{K}_{\text{GN}} \triangleq \begin{bmatrix} \text{var}(\text{NLI}_x^{(\text{GN})}) & \text{cov}(\text{NLI}_x^{(\text{GN})}, \text{NLI}_y^{(\text{GN})}) \\ \text{cov}(\text{NLI}_x^{(\text{GN})}, \text{NLI}_y^{(\text{GN})}) & \text{var}(\text{NLI}_y^{(\text{GN})}) \end{bmatrix} \tag{5.23}$$

$$= \sum_{p, \ell=0}^{N-1} \mathcal{C}_{\text{GN}}(p, \ell) \left(\text{Tr} [\mathbf{P}(z_p) \mathbf{P}^\dagger(z_\ell)] \mathbf{I} + \mathbf{P}(z_p) \mathbf{P}^\dagger(z_\ell) \right) \tag{5.24}$$

¹ n_i' is the NLI after recovering the average carrier phase rotation [111].

with the matrix trace defined as $\sum_{i,j} A_{ij} B_{ij}^* = \text{Tr}[\mathbf{A}\mathbf{B}^\dagger]$. Note that, in the absence of PDL, the last line in the above equation reduces to $3\mathbf{I}$ since $\mathbf{P}(z_p) = \mathbf{I}$ for each p , in agreement with [41].

FON contributions

We now focus on the F4 contribution, which represents the most significant EGN model correction as discussed in Chapters 2 and 3. We generalize the variance expression in Eq. (2.43), which was derived in the absence of polarization-dependent effects, by using the proper PDL-dependent weights. In particular, the product of identity matrices leading to a multiplicative factor 5 in Eq. (2.43) must be modified by substituting the proper entries of the matrix \mathbf{P} . Such an extended-expression can be written as:

$$\begin{aligned}
 \sigma_{\text{F4}i_3}^2 &= \sum_{\mathbf{h}, \mathbf{n}} \kappa_{2;2}^{(\mathbf{h})} \kappa_{1;1}^{(\mathbf{n})} |\mathcal{X}_{\mathbf{h}\mathbf{h}\mathbf{n}\mathbf{i}} + \mathcal{X}_{\mathbf{h}\mathbf{n}\mathbf{h}\mathbf{i}}|^2 \\
 &= \frac{1}{8} \sum_{h_1, n_1} \sum_{h_2, n_2} \kappa_{2;2}^{(\mathbf{h})} \kappa_{1;1}^{(\mathbf{n})} \sum_{p, \ell=0}^{N-1} \mathcal{S}_{\mathbf{h}\mathbf{h}\mathbf{n}\mathbf{i}}(z_p) \mathcal{S}_{\mathbf{h}\mathbf{h}\mathbf{n}\mathbf{i}}^*(z_\ell) \\
 &\quad \times \left(\sum_{h_3} P_{h_3 h_3}(z_p) P_{h_3 h_3}^*(z_s) + 2P_{i_3 i_3}(z_p) P_{i_3 i_3}^*(z_s) + \sum_{n_3} P_{i_3 n_3}(z_p) P_{i_3 n_3}^*(z_s) \right) \quad (5.25) \\
 &= \sum_{p, \ell=0}^{N-1} \mathcal{C}_{\text{F4}}(p, \ell) \left(\sum_{h_3} P_{h_3 h_3}(z_p) P_{h_3 h_3}^*(z_s) \right. \\
 &\quad \left. + 2P_{i_3 i_3}(z_p) P_{i_3 i_3}^*(z_s) + \sum_{n_3} P_{i_3 n_3}(z_p) P_{i_3 n_3}^*(z_s) \right)
 \end{aligned}$$

where $\mathcal{C}_{\text{F4}}(p, \ell)$ is the same as in Eq. (2.45) while its weight reduces to 5 in the case of no PDL, i.e., with $\mathbf{P} = \mathbf{I}$.

Similarly to the GN term, the result can be extended to the covariance matrix, which can be written as:

$$\mathbf{K}_{\text{F4}} = \mathbb{E}[n_i' n_j'^*] \Big|_{\text{only F4}} = \sum_{p, \ell=0}^{N-1} \mathcal{C}_{\text{F4}}(p, \ell) \mathbf{F}(z_p, z_\ell) \quad (5.26)$$

where the matrix $\mathbf{F}(z_p, z_\ell)$ has the following entries:

$$\begin{aligned} F_{11} &= 4P_{11}(z_p)P_{11}^*(z_\ell) + P_{22}(z_p)P_{22}^*(z_\ell) + P_{12}(z_p)P_{12}^*(z_\ell) \\ F_{12} &= P_{22}(z_p)P_{21}^*(z_\ell) + P_{12}(z_p)P_{11}^*(z_\ell) \\ F_{21} &= P_{11}(z_p)P_{12}^*(z_\ell) + P_{21}(z_p)P_{22}^*(z_\ell) \\ F_{22} &= 4P_{22}(z_p)P_{22}^*(z_\ell) + P_{11}(z_p)P_{11}^*(z_\ell) + P_{21}(z_p)P_{21}^*(z_\ell). \end{aligned} \quad (5.27)$$

Regarding the other FON term, i.e., the Q4 contribution, the extended variance becomes:

$$\begin{aligned} \sigma_{Q4i_3}^2 &= \sum_{\mathbf{h}, \mathbf{n}} \kappa_{2;2}^{(\mathbf{h})} \kappa_{1;1}^{(\mathbf{n})} |\mathcal{X}_{\mathbf{hnni}}|^2 \\ &= \frac{1}{8} \sum_{h_1, n_1} \sum_{h_2, n_2} \kappa_{2;2}^{(\mathbf{h})} \kappa_{1;1}^{(\mathbf{n})} \sum_{p, \ell=0}^{N-1} \mathcal{S}_{\mathbf{hnni}}(z_p) \mathcal{S}_{\mathbf{hnni}}^*(z_\ell) \sum_{h_3} P_{h_3 i_3}(z_p) P_{h_3 i_3}^*(z_s) \\ &= \sum_{p, \ell=0}^{N-1} \mathcal{C}_{Q4}(p, \ell) \sum_{h_3} P_{h_3 i_3}(z_p) P_{h_3 i_3}^*(z_s) \end{aligned} \quad (5.28)$$

where it is easy to see that, in the absence of PDL, the cross-correlation in Eq. (5.28) is weighted by a factor 1, hence obtaining the same Q4 variance derived in Chapter 2. The covariance matrix is then given by the following equation:

$$\mathbf{K}_{Q4} = \mathbb{E}[n_i' n_j'^*] \Big|_{\text{only Q4}} = \sum_{p, \ell=0}^{N-1} \mathcal{C}_{Q4}(p, \ell) \mathbf{Q}(z_p, z_\ell) \quad (5.29)$$

where the entries of the matrix $\mathbf{Q}(z_p, z_\ell)$ can be expressed as:

$$\begin{aligned} Q_{11} &= P_{11}(z_p)P_{11}^*(z_\ell) + P_{21}(z_p)P_{21}^*(z_\ell) \\ Q_{12} &= Q_{21} = 0 \\ Q_{22} &= P_{22}(z_p)P_{22}^*(z_\ell) + P_{12}(z_p)P_{12}^*(z_\ell) . \end{aligned} \quad (5.30)$$

Q6 term

The last contribution to the overall NLI variance is the HON term, called Q6 in Chapter 2. The Q6 variance generalized to include PDL writes as:

$$\begin{aligned}
 \sigma_{Q6i_3}^2 &= \sum_{\mathbf{h}} \kappa_{3;3}^{(\mathbf{h})} |\mathcal{X}_{\mathbf{h}\mathbf{h}\mathbf{h}\mathbf{i}}|^2 \\
 &= \frac{1}{8} \sum_{h_1, n_1} \sum_{h_2} \kappa_{3;3}^{(\mathbf{h})} \sum_{p, \ell=0}^{N-1} \mathcal{S}_{\mathbf{h}\mathbf{h}\mathbf{h}\mathbf{i}}(z_p) \mathcal{S}_{\mathbf{h}\mathbf{h}\mathbf{h}\mathbf{i}}^*(z_\ell) P_{i_3 i_3}(z_p) P_{i_3 i_3}^*(z_s) \\
 &= \sum_{p, \ell=0}^{N-1} \mathcal{C}_{Q6}(p, \ell) P_{i_3 i_3}(z_p) P_{i_3 i_3}^*(z_s).
 \end{aligned} \tag{5.31}$$

Extending the results from the variance of one polarization to the covariance matrix, we obtain:

$$\mathbf{K}_{Q6} = \mathbb{E}[n'_i n'_j] \Big|_{\text{only Q6}} = \sum_{p, \ell=0}^{N-1} \mathcal{C}_{Q6}(p, \ell) \mathbf{H}(z_p, z_\ell) \tag{5.32}$$

with $\mathcal{C}_{Q6}(p, \ell)$ the scalar Q6 cross-correlation between trunk p and trunk ℓ , while matrix \mathbf{H} has entries:

$$\begin{aligned}
 H_{ii} &= P_{ii}(z_p) P_{ii}^*(z_\ell), \quad i = 1, 2 \\
 H_{ij} &= 0, \quad i \neq j.
 \end{aligned} \tag{5.33}$$

Please note that in the absence of PDL we have $\mathbf{H} = \mathbf{I}$.

5.2.3 Per-polarization PDL-aware SNR

In the presence of PDL, and in the absence of joint channel equalization and FEC decoding, it is more meaningful to deal with the per-polarization SNR, because of the asymmetrical behavior of noise power. Neglecting ASE-NLI interaction, the overall SNR of the generic polarization $i_3 \in (x, y)$, follows the usual concatenation rule:

$$\frac{1}{\text{SNR}^{i_3}} = \frac{1}{\text{SNR}_{\text{ASE}}^{i_3}} + \frac{1}{\text{SNR}_{\text{NLI}}^{i_3}} \tag{5.34}$$

where $\text{SNR}_{\text{ASE}}^{i_3}$ is the ASE-only SNR given by:

$$\text{SNR}_{\text{ASE}}^{i_3} = \frac{P/2}{\sigma_{\text{ASE}i_3}^2} = \frac{\text{SNR}_{\text{ASE}}^{i_3}(\text{PDL}=0)}{\frac{1}{M} \sum_{p=1}^M P_{i_3 i_3}^{-1}(z_p)} \quad (5.35)$$

with P the $(x+y)$ signal power and $\text{SNR}_{\text{ASE}}^{i_3}(\text{PDL}=0) = P/(N_0 MB)$.

The NLI-only SNR in Eq. (5.34) can be computed as:

$$\text{SNR}_{\text{NLI}}^{i_3} = \frac{1}{(K_{i_3 i_3}^{(\text{GN})} + K_{i_3 i_3}^{(\text{F4})} + K_{i_3 i_3}^{(\text{Q4})} + K_{i_3 i_3}^{(\text{Q6})}) (P_{i_3}/2)^2}, \quad i_3 \in (x, y). \quad (5.36)$$

The denominator of Eq. (5.35), equal to the span-average of $P_{i_3 i_3}^{-1}$, is the random PDL loss/gain per polarization.

5.3 Numerical results

We validated the proposed model against SSFM based simulations. As a first test, we neglected ASE noise and focused only on the PDL-impaired nonlinear interference, which represents the main novelty. The WDM signal was composed of 11×50 GHz channels with RRC pulses with roll-off 0.01 having a symbol rate of 49 Gbaud and channel spacing of 50 GHz. The lasers' SOP was randomly chosen on the Poincaré sphere. To test the validity of the PDL-GN model we first considered PDM signals with Gaussian distributed symbols.

The links under test were either 10×100 km or 20×100 km links of SMFs, with fiber parameter as in Tab. 3.1. Two dispersion-management schemes were investigated: a DM scheme with 30 ps/nm of residual dispersion after each span, or DU. The residual dispersion at the link-end was then compensated at the receiver input. In any case, each amplifier in the link introduced 0.5 dB of PDL.

The propagation along such links was simulated by the SSFM and compared with the prediction of the PDL-GN model. In the PDL-GN matched filtering, zero-forcing PDL equalization and CPE are implicit in the model. On the other hand, in the SSFM case we implemented matched filtering followed by a 1-tap zero-forcing equalizer, and by a CPE recovering the average phase rotation induced by the fibers.

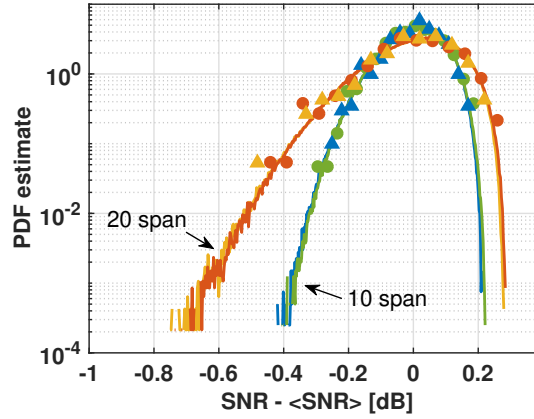


Figure 5.3: Estimate of the PDF of the SNR per polarization by SSFM simulations (triangles for DM30, circles for DU) and the corresponding (almost overlapped) PDL-GN PDFs (solid lines).

We estimated the PDF of the received SNR by Monte Carlo simulations by simulating 1000 different random PDL realizations, both with SSFM runs and with the PDL-GN model. For each realization, the SOP of the channel lasers was varied as well. For the SSFM simulations, we used a symmetric step with a first step accumulating an FWM phase of 20 rad [28] and CLE updating rule. The number of symbols was set to 65536, in order to capture the largest walk-off among channels and to have a negligible error from the Monte Carlo estimation [107]. In the PDL-GN, the frequency integrals were solved by means of Monte Carlo integration.

Figure 5.3 depicts the estimated PDF of the SNR per polarization at the central WDM channel by SSFM (symbols) and the PDL-GN model (lines). In order to compare the different links, we plotted the PDFs versus the SNR offset from its mean. The model is in excellent agreement with SSFM results for both numbers of spans, thus confirming its validity. Although it is not shown in the figure, the mean SNR of SSFM simulations was within 0.1 dB of that from the PDL-GN model, for both links.

In particular, we note that dispersion management does not affect the PDF shape, while it strongly impacts the average value (here not shown). In fact, the DM30 sce-

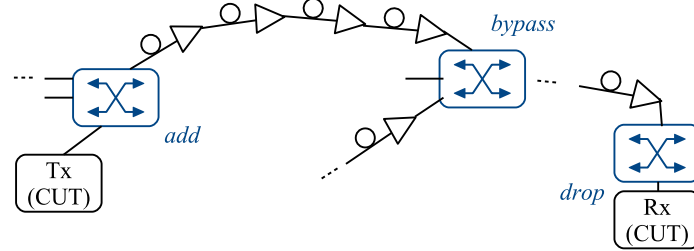


Figure 5.4: Network under investigation. 3200 km with initial add block, final drop block and a ROADM for optical bypass every 400 km.

nario is affected by a higher average NLI compared to the DU case, due to the higher correlation of the NLI among different spans.

In a second test, we investigated a different and more realistic distribution of PDL along the optical link. We thus focused on a 3200 km network segment. Differently from the previous case, we included the presence of ROADMs for the add and drop of the CUT at the beginning and end of the segment, as well as ROADMs placed every 4 spans for cross-connections. We assumed that each inline ROADM performed optical bypass, hence two WSSs are crossed. The PDL was 0.1 dB within EDFAs and 0.4 dB within each WSS, respectively [108]. Figure 5.4 sketches the link under test, as a segment of a wider network.

We investigated the transmission of 21 channels carrying PDM-star 8QAM modulation format. In this setup, we included ASE with a frequency-flat gain recovering fiber losses and a noise figure of 5 dB per amplifier. It is worth noting that, with ideal equalization, the last PDL element (hence the drop block) impacts equally ASE, NLI, and signal, hence with no implications on the statistics.

The PDF of the per-polarization SNR offset, estimated by the PDL-EGN and by SSFM at power 1 dBm per channel maximizing the SNR is reported in Fig. 5.5. Even in this scenario, we observe a good match between the models and SSFM simulations. Although it is not shown in the figure, we observed that the PDL-EGN model exhibits a bias of 0.1 dB compared to the SSFM results.

It is useful to compare the computational times of the PDL-EGN model and the

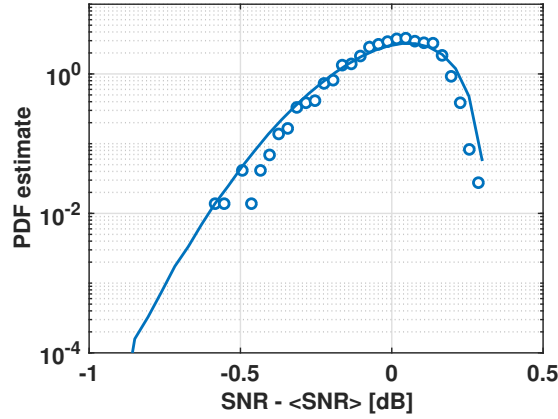
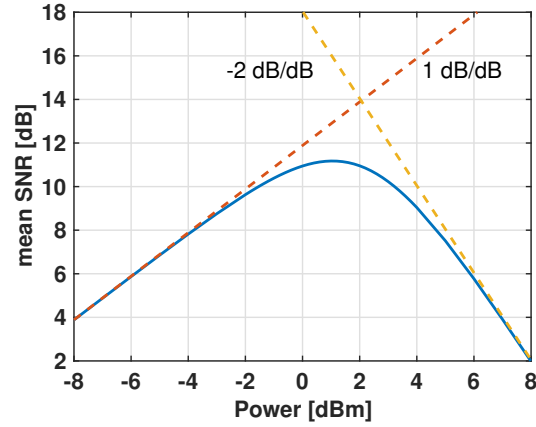


Figure 5.5: PDF estimate of the SNR offset from its mean, per polarization, by the PDL-EGN (solid line) and SSFM (symbols). 32×100 km SMF link with star-8QAM-modulation format, with different PDL between EDFA and ROADMs.

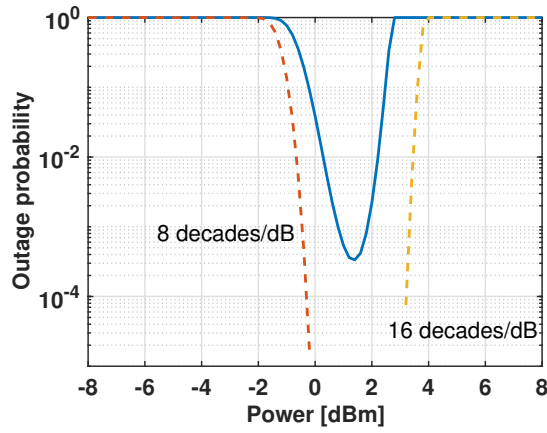
SSFM. As a reference, an SSFM simulation, with step setup as in [28], took 1 day to run 125 PDL seeds on a cluster using INTEL XEON E5- 2683v4 2.1GHz 32 cores central processing units (CPU) with 128 GB of RAM and NVIDIA Tesla P100 GPU. The same seeds have been simulated in a fraction of a second with the PDL-EGN, plus the overhead for the computation of the span cross-correlations of the order of seconds for the PDL-GN and a few minutes for the PDL-EGN. Not surprisingly, with the EGN we were able to simulate 10^6 PDL seeds, while with SSFM only 10^3 .

5.3.1 Outage probability

We now leverage the small computation time of the proposed model to estimate the probability that the system is in outage. The outage probability is here defined as the probability that the SNR falls below the threshold of 10.56 dB, corresponding to a Q-factor of 6.5 dB for star 8QAM formats. Figure 5.6(a) shows the mean per-polarization SNR versus launched channel power. We observe that the optimal channel power is 1 dBm. Moreover, the curve exhibits the typical slope of +1 dB/dB in the linear regime and -2 dB/dB in the nonlinear regime, as discussed in Fig. 1.11.



(a)



(b)

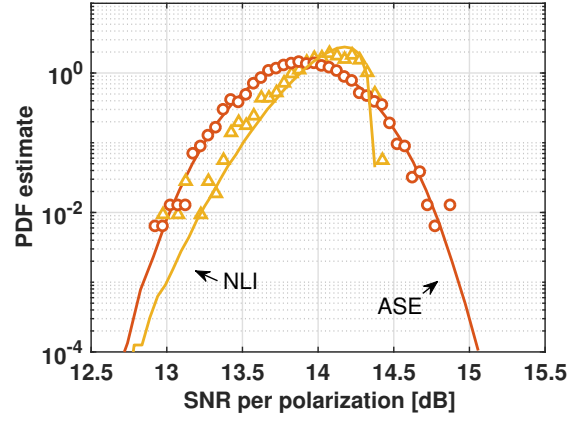
Figure 5.6: (a): mean SNR per polarization vs. power. (b): outage probability @ Q-factor = 6.5 dB. All curves with the proposed model with only ASE or NLI (dashed lines) or with both ASE and NLI (solid lines). Star-8QAM 32×100 km SMF link with different PDL between EDFA and ROADMs, see text.

On the other hand, Fig. 5.6(b) reports the outage probability versus the same power, showing several interesting features. First, we observe that the minimum outage probability is $3 \cdot 10^{-4}$, a non-negligible value indicating the importance of including PDL in the link design. Second, the best power for the mean SNR does not coincide with the best power for the outage probability, with a gap of 0.4 dB. This is strictly related to the nonlinear relation between the outage probability and the SNR. Third, we still observe a factor 2 between the slopes of the asymptotes on the outage probability graphs. Such an observation can be very useful for quickly scaling the outage probability with power.

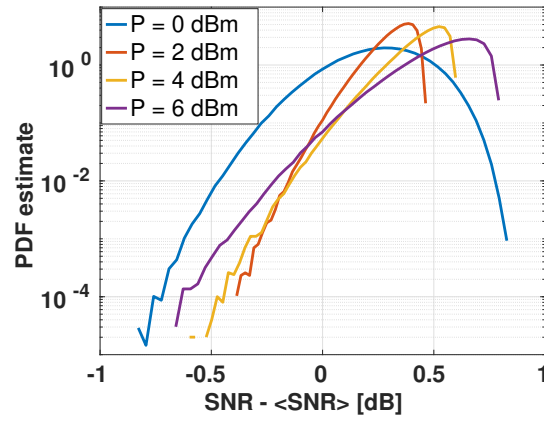
5.3.2 NLI vs ASE

In the theoretical section of this chapter, we pointed out that the NLI and the ASE noise are impacted by PDL in different ways. We now compare the statistics of the linear SNR in Eq. (5.35) and the nonlinear SNR in Eq. (5.36) for the same link considered in Fig. 5.6. We perform the comparison at a channel power yielding the same average noise variances, which was 2 dBm for this setup. The PDF of the ASE-only and NLI-only SNR are depicted in Fig. 5.7(a). We used symbols to represent SSFM results and lines for the PDL-EGN model. We observe that, in this setup, the PDF of the ASE-only case is slightly larger than that of the NLI-only case, although the average values are similar. Most important, the ASE-only and the NLI-only PDFs have different shapes, hence the NLI cannot be treated as an equivalent extra-ASE distributed along the link. Such observation also suggests that the shape of the PDF of the SNR depends on the transmitted power which determines whether the NLI or the ASE noise is dominant. As an example, Fig. 5.7(b) shows the PDF estimate of the SNR (including NLI and ASE noises) at different channel powers.

It is worth noting that if instead of plotting the per-polarization SNR PDF, we plot the PDF of the overall PDM SNR, as usually done in the literature, we obtain a sharply different behavior, as depicted in Fig. 5.8 with the PDL-EGN model. To estimate the PDM SNR, or simply SNR, it is sufficient to substitute the per-polarization power with the $(x + y)$ power, both for the signal and the noise. This difference is due to the fact that the PDFs of the PDL-impacted linear SNR on the best and worst-



(a)



(b)

Figure 5.7: PDF estimate of (a) ASE-only and NLI-only per-polarization SNR using the proposed PDL-EGN model (solid line) and SSFM simulations (circles: ASE, triangles: NLI). Power 2 dBm. (b) per-polarization SNR offset at different channel powers, including both ASE and NLI.

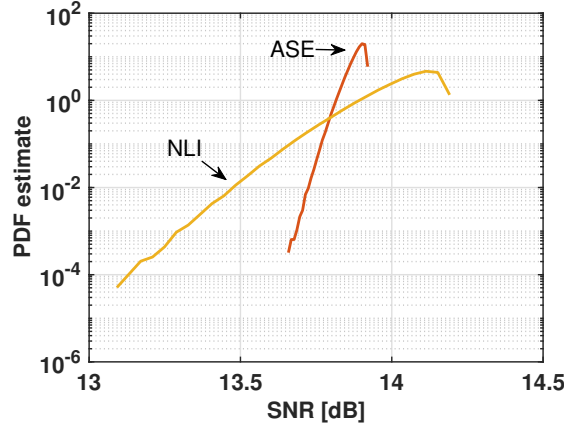


Figure 5.8: Estimate of the PDF of the linear and nonlinear PDM SNR usually adopted in the literature by the proposed PDL-EGN model. Same setup of Fig. 5.7(a)

performing polarization are antithetic, as was shown in [107]. On the other hand, SPM- and XPM-like contributions operate through a common scalar nonlinear phase on both polarizations, thus yielding similar PDFs for the two polarizations having tails in the same direction. As a result, the $(x+y)$ PDF in the NLI-only case is broader than the ASE-only case.

To further stress the importance of properly modeling the PDL impact on NLI, we now explore two alternative approximated methods for the estimation of the SNR's PDF. First, we investigate the possibility of modeling the overall noise as an ASE-like noise having an equivalent noise figure F_{eq} such that, on average with respect to the PDL realizations, the variance of the equivalent noise $\sigma_{\text{eq}i3}^2$ matches the variance of the actual noise, i.e., ASE+NLI, in the absence of PDL. Namely, the equivalent noise variance is calculated as in Eq. (5.20):

$$\sigma_{\text{eq}i3}^2 \triangleq \frac{h\nu F_{\text{eq}} GB}{2} \sum_{p=1}^M P_{i3i3}^{-1}(z_p) \quad (5.37)$$

with the following equivalent noise figure:

$$F_{\text{eq}} = F + \frac{2\sigma_{\text{NLI}i_3}^2 (\text{NOPDL})}{Mh\nu GB}. \quad (5.38)$$

In this strategy, the NLI is treated as an extra ASE noise and hence undergoes PDL in the same manner. To further simplify the modeling, one could account for PDL only in the ASE noise and treat the NLI as if it would be unaffected by PDL, hence estimated via the classic GN/EGN model in Chapter 2.

Figure 5.9 shows the PDF estimate of the $(x+y)$ SNR for the same setup of Fig. 5.8. In solid line we reported the PDL-EGN estimate, in dashed line the equivalent noise approach, and in dotted line the EGN model estimate with PDL included only in the ASE noise. Although the average SNR values coincide in the three approaches, the PDF shape is clearly different. In particular, neglecting any impact of PDL on the Kerr effect is particularly detrimental yielding a much narrower PDF.

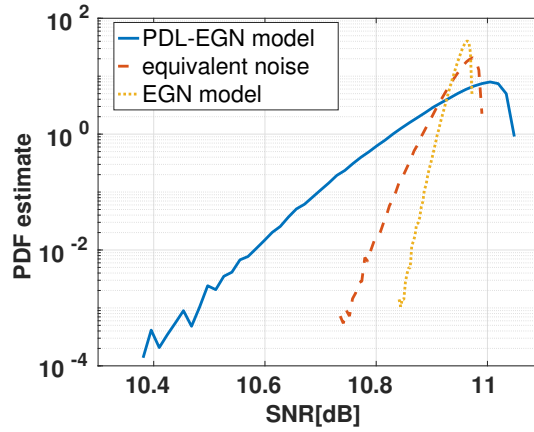


Figure 5.9: Estimate of the PDF of PDM SNR. Solid: PDL-EGN model estimate. Dashed: linear SNR with inflated average ASE variance. Dotted: EGN model estimate with PDL included only in the ASE noise.

5.3.3 Modulation format

We next investigated the impact of the modulation format. The link is the same as Fig. 5.7 with both ASE and NLI, and channel power 2 dBm. The PDF was estimated with the proposed PDL-EGN model for different modulation formats and shown in Fig. 5.10. Consistently with the literature [39–41] and the other chapters of this thesis, we observe that the GN model underestimates the average value of the other modulation formats.

The maximum mismatch is associated with the QPSK case, due to the significant FON contribution, with an error up to 0.77 dB. However, we note that the shapes of the PDFs are less impacted, with a slightly wider PDF for the QPSK modulation format. As a comparison metric, we can consider the difference between the average SNR and the SNR at a fixed PDF estimate equal to 10^{-4} , here called ΔSNR . For instance, this ΔSNR in the GN case is 0.58 dB, while for QPSK, by using the same random seeds, we estimated $\Delta\text{SNR} = 0.76$ dB.

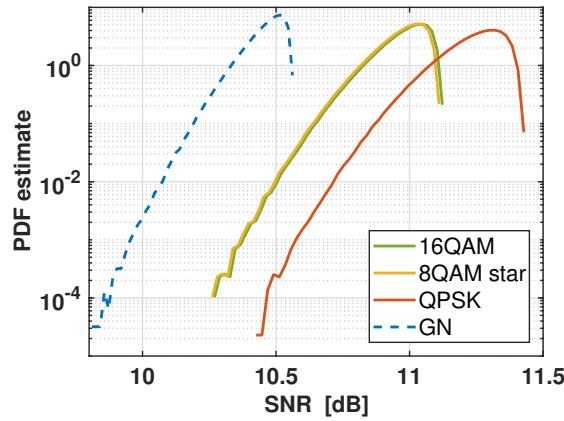


Figure 5.10: Same setup of Fig. 5.7 with ASE and NLI, for different modulation formats.

5.3.4 Number of channels

We tested the impact of the number of channels on the SNR PDF. To this aim, instead of performing time-consuming SSFM simulations, we exploited the proposed PDL-GN model which was validated in the previous sections.

The link under test was the same as Fig. 5.4 with the setup used for Fig. 5.5, except for the number of channels and the channel power. The WDM comb was composed of 3, 21, or 101 channels. For the link under test, the optimal power maximizing the SNR for such a WDM transmission (i.e., the NLT) was $P = 1.4$, 0.7 , and 0.4 dBm respectively. In order to compare the impact of PDL on the PDFs, Fig. 5.11 shows the PDF of the per-polarization SNR offset with respect to the mean value, instead of the PDF of the SNR. It can be seen that the PDFs are almost identical in the three cases under test. A similar result is obtained when increasing the power by 3 dB from the optimal one.

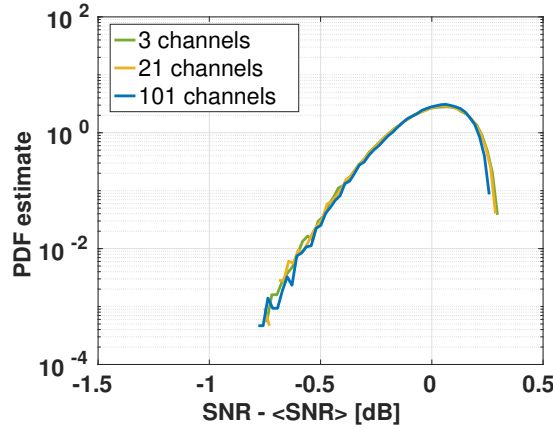


Figure 5.11: PDF estimate of the SNR offset from its mean, for a WDM comb of 3, 21 or 101 channels. PDL-GN model. 32×100 km SMF link with different PDL between EDFA and ROADMs. Each SNR curve evaluated at the relative optimal power (NLT).

5.3.5 PDL profile impact

As a last investigation, we compare the outage probability of links having a different PDL profile. For the sake of simplicity, we considered the presence of a PDL element at each span-end. The total amount of PDL was fixed to 30 dB, and we considered three different distributions of such total PDL:

1. The first PDL element introduces 2 dB of PDL and the remaining PDL is equally partitioned among the other PDL elements (0.47 dB each);
2. All the devices introduce the same amount of PDL equal to 0.5 dB;
3. The last PDL block has 2 dB of PDL. Similarly to the first case, the remainder is split among the other devices.

All the links under test were 60×100 km of SMF, having parameters listed in Tab. 3.1. The transmitted signal was a WDM comb of 51 PDM-QPSK modulated channels. We estimated the outage probability by means of the proposed PDL-EGN model, with 10^6 PDL realizations. The results are reported in Fig. 5.12.

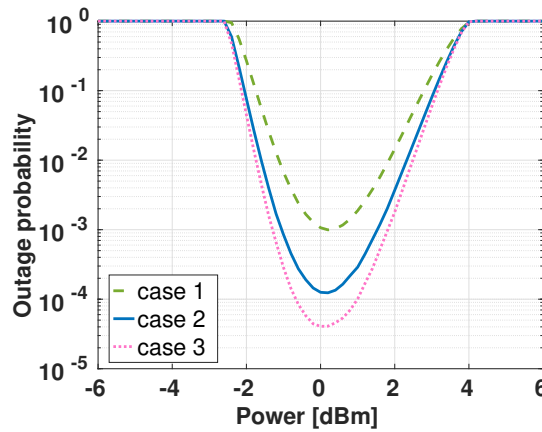


Figure 5.12: Outage probability estimated with the PDL-EGN model for a 60×100 km link with PDL distributed in three different ways (see text).

The figure shows that the PDL profile has a non-negligible impact on the system performance. First of all, comparing the curve relative to case 1 and 3 we note that switching the first and the last PDL elements do not yield the same outage probability. In particular, it is more detrimental to have a higher amount of PDL in the first element, with a non-negligible outage probability of 10^{-3} at the best power, instead of the last one. The reason is that the PDL introduced by the last block can be equalized as well on the NLI and the ASE noise, while the PDL introduced in the first block remains on the noises even after equalization. For the same reason, the link with higher PDL concentrated in the last block performs better than the homogeneous link, in which a slightly higher PDL is introduced by the blocks that cannot be recovered.

Chapter 6

The ergodic GN model with modal dispersion

Perturbations of the symmetrical structure of a single-mode fiber yield birefringence, polarization mixing, and PMD. As a result of PMD, a random DGD between the two polarizations arises [8]. Such a random DGD accumulates with the square root of length [17], hence becoming potentially detrimental for long-haul transmissions.

Although the effects of PMD on the useful signal can be compensated with a multiple-input multiple-output (MIMO) based DSP, its impact remains in the nonlinear interference. Unfortunately, the numerical estimation of the interaction between PMD and the fiber Kerr effect has great complexity, due to the random nature of PMD which calls for the evaluation of many random seeds. Despite their complexity, numerical and experimental studies were performed in the literature [113, 114], highlighting a beneficial effect of PMD in mitigating the NLI. Nevertheless, for practical values of PMD in modern optical fibers, the impact on the system performance for typical DU links is almost negligible. For this reason, the inclusion of PMD in the GN model was overlooked in the single-mode transmission literature.

On the other hand, in a multi-mode transmission, the modal dispersion (MD) can introduce a random mixing and delay among the polarizations of all the guided modes [20]. In such SDM fibers, the values of modal dispersion are orders of magni-

tude larger than in an SMF [115, 116]. As a consequence, neglecting the interaction between MD and the Kerr effect might not be justified in this context. The presence of MD was neglected in the GN model extension to SDM in [117, 118], while it was included in [119] for limiting scenarios of large MD values by neglecting the impact of MD within a channel bandwidth. Arbitrary values of MD were taken into account in [120] only with respect to self-channel nonlinear effects.

In this chapter, we first propose an extension of the GN model to fully account for the impact of arbitrary values of MD on the NLI. The model is validated against SSFM simulations and exploited to investigate the impact of MD for different system setups. Then, we present a simplified expression for the XPM variance, at arbitrary MD values, derived under the assumption of [119].

The work reported in this chapter was partially published in [121].

6.1 Modal dispersion

The phenomenon of modal dispersion manifests as a distortion of the pulse transmitted at the fiber input, due to the induced mode-dependent delay, and as a frequency-dependent change of the SOP at the fiber output.

6.1.1 The PMD vector

In an SMF, the SOPs that, to first order, do not change in frequency after crossing the fiber are called principal states of polarization (PSP). Light launched along the two PSP sees a different fiber refractive index, hence it experiences a different delay. The DGD measures this differential delay between the fast and slow PSP. In the literature it is customary to represent the effects of PMD through the PMD vector [8, 122], which can be expressed by the following Stokes vector [8]:

$$\vec{\tau} = \tau \hat{p} \quad (6.1)$$

having the same orientation \hat{p} of the slow PSP and length equal to the DGD τ . The PMD vector evolution along the distance is governed by the following equation:

$$\frac{\partial \vec{\tau}}{\partial z} = \frac{\partial \vec{\beta}}{\partial \omega} + \vec{\beta} \times \vec{\tau} \quad (6.2)$$

where $\vec{\beta}$ is the vector of local birefringence [8]. The random nature of the birefringence vector yields a random PMD vector, usually modeled with independent and identically distributed (i.i.d) Gaussian entries. As a consequence, its length, i.e. the DGD, becomes a random variable. Due to the Gaussian distribution of the PMD vector entries, the DGD has a Maxwellian distribution [17] whose mean value scales with the square-root of distance:

$$\mathbb{E}[\tau(z)] = \kappa_{\text{PMD}} \sqrt{z}. \quad (6.3)$$

Random perturbations also cause the PMD vector to depend on frequency. Its frequency autocorrelation function was derived in [123].

6.1.2 The MD vector

Such a formalism was extended to SDM transmissions in [124], where the authors introduced the MD vector. The MD vector is a generalization of the PMD vector in Eq. (6.1) in a $4N_m^2 - 1$ dimensional space, with N_m the number of spatial modes. Its length thus generalizes to

$$\tau \triangleq \sqrt{2N_m \sum_{n=1}^{2N_m} t_n^2} \quad (6.4)$$

where t_n are the individual delays of the polarizations $n = 1, \dots, 2N_m$. Similarly to the single-mode case, the MD vector is a random vector which evolves along distance according to Eq. (6.2), generalized to the higher dimensional Stokes space. Modeling its entries as i.i.d Gaussian random variables, the MD length in Eq. (6.4) is now a chi-distributed random variable [124, 125].

To quantify MD it is customary in the literature to focus on the mode-averaged intensity impulse response (IIR) $I(t)$, i.e., the output power when a single-mode carrying white noise is excited on input [125]. The variance T_I^2 of such a Gaussian-shaped

IIR is related to $\mathbb{E}[\tau^2]$ in Eq. (6.3) and the spatial mode dispersion (SMD) parameter in [126], by:

$$T_I^2 = \frac{\mathbb{E}[\tau^2]}{4N_m^2} = \frac{1}{2N_m} \sum_{n=1}^{2N_m} \mathbb{E}[t_n^2] \triangleq \frac{\text{SMD}^2 \cdot z}{4}. \quad (6.5)$$

6.1.3 The linear fiber response with MD

Due to the random nature of MD, the linear operator in Eq. (1.52) can be generalized to the following expression:

$$\mathcal{F}\{e^{\mathcal{L}z}\} = e^{-j\left(\frac{\beta_2}{2}\omega^2 + \frac{\beta_3}{6}\omega^3\right)z} \mathbf{U}(z, \omega) \quad (6.6)$$

where $\mathbf{U}(z, \omega)$ is a $2N_m \times 2N_m$ unitary matrix that accounts for random coupling among modes and modal dispersion. Such a transfer matrix is fast-varying along the propagation length, due to the fast variation of birefringence. A possible strategy to describe its evolution along distance is represented by the so-called *waveplate model* [8]. The idea behind this approach is to model the optical fiber as a concatenation of sections, or waveplates, over which the axes of birefringence are constant. Along this correlation length L_{corr} , the propagation can be modeled as ideal, i.e., unperturbed, in the proper reference system. Then, the random rotation of the axes of birefringence makes the next waveplates uncorrelated to the previous one. The overall fiber transfer matrix can thus be seen as the product of N_w independent random matrices, one for each section of the waveplate model:

$$\mathbf{U}(z, \omega) = \prod_{k=1}^{N_w} \mathbf{M}_k(\omega). \quad (6.7)$$

The matrix characterizing each waveplate $k = 1, \dots, N_w$ can be expressed, in frequency, as a matrix exponential [20] of the matrix $\mathbb{B}^{(1)}$ in Eq. (1.47), and can be diagonalized as follow:

$$\begin{aligned}
\mathbf{M}_k(\omega) &= e^{-j\mathbb{B}_k^{(1)}L_{\text{corr}}} \\
&= \mathbf{V}_k \begin{bmatrix} e^{-j\omega t_1} & & 0 \\ & \ddots & \\ 0 & & e^{-j\omega t_{2N_m}} \end{bmatrix} \mathbf{V}_k^\dagger
\end{aligned} \tag{6.8}$$

where \mathbf{V}_k is a frequency-independent unitary matrix accounting for the change of basis into the waveplate reference system, while the diagonal matrix describes the uncoupled propagation along the waveplate length.

Alternatively to the discretized approach adopted by the waveplate model, the evolution of the transfer matrix can be modeled by means of the following differential equation [125, 127]:

$$\frac{d\mathbf{U}(z, \omega)}{dz} = -\frac{j\omega\mu}{2N_m}\mathbb{B}^{(1)}\mathbf{U}(z, \omega) \tag{6.9}$$

where the generic elements of the matrix $d\mathbf{W} \equiv d\mathbf{W}(z) = \mathbb{B}^{(1)}dz$ describes a Brownian motion, having properties:

$$\begin{aligned}
dW_{kn} &= dW_{nk}^* \\
\mathbb{E}[dW_{kn}] &= 0 \\
\mathbb{E}[dW_{ij}dW_{kn}^*] &= 2\delta_{ik}\delta_{jn} \cdot dz.
\end{aligned} \tag{6.10}$$

The parameter μ in Eq. (6.9) identifies the strength of MD, and is related to the SMD coefficient in Eq. (6.5) through the following equation:

$$\mu = \sqrt{\frac{N_m^3}{4N_m^2 - 1}} \text{SMD}. \tag{6.11}$$

In particular, for an SMF the PMD parameter in Eq. (6.3) is related to the parameter μ through [125, 127]:

$$\mu = \sqrt{\frac{\pi}{8}} \kappa_{\text{PMD}}. \quad (6.12)$$

6.2 The ergodic MD-GN model

In order to include MD in the GN model, we proceed by substituting the extended linear operator of Eq. (6.6) in the NLI general expression in Eq. (2.1). As a result, the tensor of Eq. (2.9) generalizes to the following:

$$\begin{aligned} \mathcal{X}_{\text{hmmi}} = & -j \left(\frac{1}{\sqrt{2}} \right)^3 \kappa \gamma \int_0^L e^{-\alpha z} \iiint_{-\infty}^{\infty} e^{j\Delta\beta z} \langle \tilde{G}_{\mathbf{h}}(0, \omega + \omega_1 + \omega_2) | \mathbf{P}(z, \omega + \omega_2, \omega_1) \\ & | \tilde{G}_{\mathbf{m}}(0, \omega + \omega_2) \rangle \langle \tilde{G}_{\mathbf{i}}(0, \omega) | \mathbf{P}^\dagger(z, \omega, \omega_1) | \tilde{G}_{\mathbf{n}}(0, \omega + \omega_1) \rangle \frac{d\omega}{2\pi} \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} dz \end{aligned} \quad (6.13)$$

where the Manakov coefficient κ is redefined as in Eq. (1.48), and the matrix \mathbf{P} is defined in terms of the transfer matrix \mathbf{U} similarly to Eq. (5.13):

$$\mathbf{P}(z, \omega, \omega_1) \triangleq \mathbf{U}^\dagger(z, \omega + \omega_1) \mathbf{U}(z, \omega). \quad (6.14)$$

Note that, contrary to the case with only PDL in Eq. (5.12), the matrix \mathbf{P} is frequency-dependent and appears two times in Eq. (6.13).

The novel tensor is then substituted in the general expression of the variance of the SON/GN term, as per Eq. (2.28). After applying the Poisson formula in Eq. (2.31), as in Chapter 2, we can remove the summation over time indexes and drop three frequency integrals. The variance of the GN term on the generic polarization i_3 can then be written as:

$$\begin{aligned}
\sigma_{\text{GN}i_3}^2 &= \kappa_{1;1}^3 \sum_{\mathbf{h}, \mathbf{m}, \mathbf{n}} \mathcal{X}_{\mathbf{h}\mathbf{m}\mathbf{n}i} (\mathcal{X}_{\mathbf{h}\mathbf{m}\mathbf{n}i}^* + \mathcal{X}_{\mathbf{h}\mathbf{m}\mathbf{n}i}^*) \\
&= \frac{1}{8T^3} \kappa^2 \gamma^2 \kappa_{1;1}^3 \sum_{h_2, m_2, n_2} \iiint_{-\infty}^{\infty} \left(|\eta_{i_3 i_3}^{(1)}(\omega, \omega_1, \omega_2)|^2 + |\eta_{i_3 i_3}^{(2)}(\omega, \omega_1, \omega_2)|^2 \right) \\
&\quad \times |\tilde{p}_{h_2}(\omega + \omega_1 + \omega_2 - \omega_{h_2})|^2 |\tilde{p}_{m_2}(\omega + \omega_2 - \omega_{m_2})|^2 \\
&\quad \times |\tilde{p}_{n_2}(\omega - \omega_{n_2})|^2 |\tilde{p}_{n_2}(\omega + \omega_1 - \omega_{n_2})|^2 \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} \frac{d\omega}{2\pi}.
\end{aligned} \tag{6.15}$$

Contrary to the general theory presented in Chapter 2, the two terms $\mathcal{X}_{\mathbf{h}\mathbf{m}\mathbf{n}i} \mathcal{X}_{\mathbf{h}\mathbf{m}\mathbf{n}i}^*$ and $\mathcal{X}_{\mathbf{h}\mathbf{m}\mathbf{n}i} \mathcal{X}_{\mathbf{h}\mathbf{m}\mathbf{n}i}^*$ are not degenerate due to the presence of MD. In particular, they are associated with a fiber kernel $\eta_{i_3 i_3}^{(1)}$ and $\eta_{i_3 i_3}^{(2)}$, respectively. The modulus square of such fiber kernels, also known as the FWM efficiency, appearing in Eq. (6.15) is obtained through a double integration along distance as follow:

$$|\eta_{i_3 i_3}^{(1)}(\omega, \omega_1, \omega_2)|^2 = \int_0^L \int_0^L e^{-(\alpha+j\Delta\beta)z} e^{-(\alpha-j\Delta\beta)s} \Lambda_{i_3 i_3}(z, s, \omega, \omega_1, \omega_2) dz ds \tag{6.16a}$$

$$|\eta_{i_3 i_3}^{(2)}(\omega, \omega_1, \omega_2)|^2 = \int_0^L \int_0^L e^{-(\alpha+j\Delta\beta)z} e^{-(\alpha-j\Delta\beta)s} \Xi_{i_3 i_3}(z, s, \omega, \omega_1, \omega_2) dz ds \tag{6.16b}$$

for a single-span with fiber length L , where $\Lambda_{i_3 i_3}$ and $\Xi_{i_3 i_3}$ indicate the (i_3, i_3) entry of the following matrices:

$$\begin{aligned}
\mathbf{\Lambda} &\triangleq \text{Tr} [\mathbf{P}(z, \omega + \omega_2, \omega_1) \mathbf{P}^\dagger(s, \omega + \omega_2, \omega_1)] \mathbf{P}^\dagger(z, \omega, \omega_1) \mathbf{P}(s, \omega, \omega_1) \\
\mathbf{\Xi} &\triangleq \mathbf{P}^\dagger(z, \omega, \omega_2) \mathbf{P}^\dagger(s, \omega + \omega_2, \omega_1) \mathbf{P}(z, \omega + \omega_1, \omega_2) \mathbf{P}(s, \omega, \omega_1).
\end{aligned} \tag{6.17}$$

Please note that the matrix \mathbf{P} reduces to a $2N_m \times 2N_m$ identity matrix in the absence of MD. In such a case, the above matrices take the simple expressions $\mathbf{\Lambda} = 2N_m \mathbf{I}$ and $\mathbf{\Xi} = \mathbf{I}$, hence yielding an overall $2N_m + 1$ weight that can be factored out of the integrals. Such a weight coincides with the factor 3 reported in Chapter 2 for a single-mode fiber and is consistent with [41]. We recall that in the absence polarization-dependent effects it is

$$\frac{|\eta_{i_3 i_3}^{(1)}|^2}{2N_m} = |\eta_{i_3 i_3}^{(2)}|^2 = |\eta_1|^2 = \frac{1 - e^{-\alpha L}(e^{-j\Delta\beta L} + e^{j\Delta\beta L}) + e^{-2\alpha L}}{\alpha^2 + \Delta\beta^2} \approx \frac{1}{\alpha^2 + \Delta\beta^2} \quad (6.18)$$

where η_1 is the single-span fiber kernel as defined in Eq. (2.16), and the last approximation holds for any sufficiently long optical fiber having $L \gg 1/\alpha$.

We note that the only novelty is introduced by matrices $\mathbf{\Lambda}$ and $\mathbf{\Xi}$. However, the random nature of these matrices makes it unfeasible to approach Eq. (6.15) numerically. Intending to reduce the problem complexity we thus propose to average the NLI variance with respect to the MD realizations. In practice, we propose to replace $\mathbf{\Lambda}$ and $\mathbf{\Xi}$ in the FWM efficiency terms with $\mathbb{E}[\mathbf{\Lambda}]$ and $\mathbb{E}[\mathbf{\Xi}]$, respectively. Such a calculation will be the topic of the next section.

6.2.1 Ergodic NLI variance

The stochastic differential equation (SDE) in Eq. (6.9) can be expressed in integral form as:

$$\mathbf{U}(z) = \mathbf{U}(0) - \frac{j\omega\mu}{2N_m} \int_0^z \mathbf{U}(\zeta) \circ d\mathbf{W}(\zeta) \quad (6.19)$$

where the symbol \circ indicates the Stratonovich integral [31] defined as the limit of the partial sums

$$\sum_{\ell} \frac{\mathbf{U}(\zeta_{\ell}) + \mathbf{U}(\zeta_{\ell-1})}{2} (\mathbf{W}(\zeta_{\ell}) - \mathbf{W}(\zeta_{\ell-1})) \quad (6.20)$$

with ℓ the index of the integration interval. For our purposes, it is more convenient to interpret stochastic integrals in the Ito sense [18, 31, 127]. Such a stochastic integral, expressed as $\int_0^z \mathbf{U}(\zeta) d\mathbf{W}(\zeta)$, is defined as the limit of

$$\sum_{\ell} \mathbf{U}(\zeta_{\ell-1}) (\mathbf{W}(\zeta_{\ell}) - \mathbf{W}(\zeta_{\ell-1})) \quad (6.21)$$

where the increment points "towards the future" [31, p. 92].

As a first step, we thus convert Eq. (6.9) governing the evolution of the transfer matrix in Ito's form by applying the conversion [31] to each element of the matrix:

$$d\mathbf{U} = -\frac{\omega^2 \mu^2}{2N_m} \mathbf{U} dz - \frac{j\omega\mu}{2N_m} (d\mathbf{W}) \mathbf{U} \quad (6.22)$$

where the first term on the right-hand side is the Ito's correction term. It is convenient to get rid of such a term by moving in the following reference system:

$$\mathbf{U}(z, \omega) \triangleq e^{-\frac{\omega^2 \mu^2 z}{2N_m}} \mathbf{R}(z, \omega) \quad (6.23)$$

with the matrix \mathbf{R} evolving along distance according to the following SDE:

$$d\mathbf{R} = -\frac{j\omega\mu}{2N_m} (d\mathbf{W}) \mathbf{R}. \quad (6.24)$$

We thus define the matrix \mathbf{Q} which plays the role of \mathbf{P} in the novel reference system:

$$\mathbf{Q}_{ij}(z) \triangleq \mathbf{R}^\dagger(z, v_i) \mathbf{R}(z, v_j) \quad (6.25)$$

where we exploited the following compact notation for the involved frequencies:

$$\begin{aligned} v_1 &\triangleq \omega + \omega_1 + \omega_2, & v_2 &\triangleq \omega + \omega_2 \\ v_3 &\triangleq \omega + \omega_1, & v_4 &\triangleq \omega. \end{aligned} \quad (6.26)$$

The SDE for the matrix $\mathbf{Q}_{ij}(z)$ in Eq. (6.25) can be computed as:

$$\begin{aligned} d\mathbf{Q}_{ij}(z) &= (d\mathbf{R}^\dagger(z, v_i)) \mathbf{R}(z, v_j) + \mathbf{R}^\dagger(z, v_i) (d\mathbf{R}(z, v_j)) \\ &\quad + (d\mathbf{R}^\dagger(z, v_i)) (d\mathbf{R}(z, v_j)). \end{aligned} \quad (6.27)$$

Expressing now the matrices of interest in the new notation yields:

$$\begin{aligned} \mathbf{\Lambda} &\triangleq e^{-\frac{\mu^2(z+s)}{2N_m} (v_1^2 + v_2^2 + v_3^2 + v_4^2)} \mathbf{\Lambda}' \\ \mathbf{\Xi} &\triangleq e^{-\frac{\mu^2(z+s)}{2N_m} (v_1^2 + v_2^2 + v_3^2 + v_4^2)} \mathbf{\Xi}' \end{aligned} \quad (6.28)$$

where we defined:

$$\begin{aligned}\Lambda' &= \text{Tr} \left[\mathbf{Q}_{12}(z) \mathbf{Q}_{12}^\dagger(s) \right] \mathbf{Q}_{34}^\dagger(z) \mathbf{Q}_{34}(s) \\ \Xi' &= \mathbf{Q}_{42}(z) \mathbf{Q}_{12}^\dagger(s) \mathbf{Q}_{31}^\dagger(z) \mathbf{Q}_{34}(s)\end{aligned}\quad (6.29)$$

having the same structure of Eq. (6.17). Treating s as a fixed spatial coordinate, the SDE of matrices Λ' and Ξ' with respect to coordinate z can be written as:

$$\begin{aligned}d\Lambda' &= \text{Tr} \left[(d\mathbf{Q}_{12}(z)) \mathbf{Q}_{12}^\dagger(s) \right] \mathbf{Q}_{34}^\dagger(z) \mathbf{Q}_{34}(s) + \text{Tr} \left[\mathbf{Q}_{12}(z) \mathbf{Q}_{12}^\dagger(s) \right] \\ &\quad \times (d\mathbf{Q}_{34}^\dagger(z)) \mathbf{Q}_{34}(s) + \text{Tr} \left[(d\mathbf{Q}_{12}(z)) \mathbf{Q}_{12}^\dagger(s) \right] (d\mathbf{Q}_{34}^\dagger(z)) \mathbf{Q}_{34}(s)\end{aligned}\quad (6.30)$$

$$\begin{aligned}d\Xi' &= (d\mathbf{Q}_{42}(z)) \mathbf{Q}_{12}^\dagger(s) \mathbf{Q}_{31}^\dagger(z) \mathbf{Q}_{34}(s) + \mathbf{Q}_{42}(z) \mathbf{Q}_{12}^\dagger(s) (d\mathbf{Q}_{31}^\dagger(z)) \mathbf{Q}_{34}(s) \\ &\quad + (d\mathbf{Q}_{42}(z)) \mathbf{Q}_{12}^\dagger(s) (d\mathbf{Q}_{31}^\dagger(z)) \mathbf{Q}_{34}(s).\end{aligned}\quad (6.31)$$

We now take the average of the above equations. Such an averaging process involves the evaluation of $\mathbb{E}[d\mathbf{Q}_{ij}]$, that is the expectation of Eq. (6.27). It is worth noting that in such computation, only the average of the second order product survives $(d\mathbf{R}^\dagger(z, v_i))(d\mathbf{R}(z, v_j))$. Hence, the expectation reduces to:

$$\mathbb{E}[d\mathbf{Q}_{ij}] = \mathbb{E}[(d\mathbf{R}^\dagger(z, v_i))(d\mathbf{R}(z, v_j))] = 2N_m \frac{v_i v_j \mu^2}{2N_m^2} \mathbb{E}[\mathbf{Q}_{ij}] \quad (6.32)$$

where we substituted $d\mathbf{R}^\dagger(z, v_i)$ with the right-hand side of Eq. (6.24), and exploited the following property $\mathbb{E}[\mathbf{A}(d\mathbf{W})^\dagger(d\mathbf{W})\mathbf{B}] = 2N_m \mathbf{A}\mathbf{B}dz$, along with the properties of the matrix $d\mathbf{W}$ in Eq. (6.10). Similar reasoning can be applied to compute the expectation of the last addend in Eq. (6.30) and Eq. (6.31) as:

$$\mathbb{E} \left[\text{Tr} \left[(d\mathbf{Q}_{12}(z)) \mathbf{Q}_{12}^\dagger(s) \right] (d\mathbf{Q}_{34}^\dagger(z)) \mathbf{Q}_{34}(s) \right] = \frac{(v_1 - v_2)(v_3 - v_4)\mu^2}{2N_m^2} \mathbb{E}[\Xi'] \quad (6.33)$$

$$\mathbb{E} \left[(d\mathbf{Q}_{42}(z)) \mathbf{Q}_{12}^\dagger(s) (d\mathbf{Q}_{31}^\dagger(z)) \mathbf{Q}_{34}(s) \right] = \frac{(v_4 - v_2)(v_3 - v_1)\mu^2}{2N_m^2} \mathbb{E}[\Lambda'] \quad (6.34)$$

where we exploited the following properties:

$$\begin{aligned}\mathbb{E} [\mathbf{A}(\mathrm{d}\mathbf{W})^\dagger \mathbf{B} \mathrm{d}\mathbf{W}] &= 2\mathrm{Tr}[\mathbf{B}]\mathbf{A} \mathrm{d}z \\ \mathbb{E} [\mathrm{Tr}[(\mathrm{d}\mathbf{W})\mathbf{A}] \mathbf{B} \mathrm{d}\mathbf{W}^\dagger] &= 2\mathbf{B} \mathrm{d}z.\end{aligned}\tag{6.35}$$

Finally, after returning back to the original frequency notation through Eq. (6.26) and hence to the original matrices \mathbf{A} and \mathbf{E} , the evolution with respect to $\xi = |z - s|$ is given by the following system of ordinary differential equations:

$$\begin{aligned}\frac{\mathrm{d}\mathbb{E}[\mathbf{A}]}{\mathrm{d}\xi} &= \frac{\mu^2}{N_m} \left(-\omega_1^2 \mathbb{E}[\mathbf{A}] + \frac{\omega_1^2}{2N_m} \mathbb{E}[\mathbf{E}] \right) \\ \frac{\mathrm{d}\mathbb{E}[\mathbf{E}]}{\mathrm{d}\xi} &= \frac{\mu^2}{N_m} \left(\frac{\omega_2^2}{2N_m} \mathbb{E}[\mathbf{A}] - \omega_2^2 \mathbb{E}[\mathbf{E}] \right).\end{aligned}\tag{6.36}$$

Such a linear system can be written in a vectorial form in terms of its 2×2 matrix:

$$\mathbf{A} = \begin{bmatrix} -\omega_1^2 & \frac{\omega_1^2}{2N_m} \\ \frac{\omega_2^2}{2N_m} & -\omega_2^2 \end{bmatrix}\tag{6.37}$$

and a solution can be found by applying the matrix exponential $e^{\frac{\mu^2}{N_m} \mathbf{A} \xi}$ to the following initial values:

$$\mathbb{E}[\mathbf{A}(\xi = 0)] = 2N_m \mathbf{I}\tag{6.38}$$

$$\mathbb{E}[\mathbf{E}(\xi = 0)] = \mathbf{I}.\tag{6.39}$$

The problem allows a closed-form solution provided that the matrix exponential is computed analytically. To this aim, we rewrite such a matrix exponential applying the Cayley-Hamilton theorem, thus obtaining the simplified expression:

$$e^{\frac{\mu^2}{N_m} \mathbf{A} \xi} = e^{-p \frac{\mu^2}{N_m} \xi} \left[\left(\cosh \left(q \frac{\mu^2}{N_m} \xi \right) + p \frac{\sinh \left(q \frac{\mu^2}{N_m} \xi \right)}{q} \right) \mathbf{I} + \frac{\sinh \left(q \frac{\mu^2}{N_m} \xi \right)}{q} \mathbf{A} \right]\tag{6.40}$$

where the parameters p and q depend on the frequency involved in the FWM process and are defined as:

$$p \triangleq -\frac{\text{Tr}[\mathbf{A}]}{2} = \frac{\omega_1^2 + \omega_2^2}{2} \quad (6.41)$$

$$q \triangleq \sqrt{p^2 - \omega_1^2 \omega_2^2 \left(1 - \frac{1}{4N_m^2}\right)}. \quad (6.42)$$

We thus find that the system of ODEs in Eq. (6.36) allows the following compact solution:

$$\begin{aligned} \mathbb{E}[\mathbf{\Lambda}] &= m_1 \left(e^{(q-p)\frac{\mu^2}{N_m}|z-s|} (1+c_1) + e^{-(q+p)\frac{\mu^2}{N_m}|z-s|} (1-c_1) \right) \\ \mathbb{E}[\mathbf{\Xi}] &= m_2 \left(e^{(q-p)\frac{\mu^2}{N_m}|z-s|} (1+c_2) + e^{-(q+p)\frac{\mu^2}{N_m}|z-s|} (1-c_2) \right) \end{aligned} \quad (6.43)$$

where we defined:

$$c_\ell \triangleq \begin{cases} \frac{p}{q} - \frac{\omega_1^2}{q} \left(1 - \frac{1}{4N_m^2}\right) & \ell = 1 \\ \frac{p}{q} & \ell = 2 \end{cases} \quad (6.44)$$

and

$$m_\ell \triangleq \begin{cases} N_m & \ell = 1 \\ \frac{1}{2} & \ell = 2. \end{cases} \quad (6.45)$$

6.2.2 FWM efficiency

The next step towards evaluating the NLI variance is to substitute $\Lambda_{i_3 i_3}(z, s, \omega, \omega_1, \omega_2)$ and $\Xi_{i_3 i_3}(z, s, \omega, \omega_1, \omega_2)$ in Eqs. (6.16a)-(6.16b) with their average value at each coordinate, as computed in the previous section and reported in Eq. (6.43). Thanks to

the structure of the system solution in Eq. (6.43), the average value of the FWM efficiency of generic type ℓ in Eqs. (6.16a)-(6.16b) can be written as:

$$\begin{aligned} \mathbb{E} \left[|\eta^{(\ell)}(\omega, \omega_1, \omega_2)|^2 \right] &= m_\ell \int_0^L \int_0^L e^{-\alpha(z+s)} e^{j\Delta\beta(z-s)} \\ &\quad \times \left(e^{(q-p)\frac{\mu^2}{N_m}|z-s|} (1+c_\ell) + e^{-(q+p)\frac{\mu^2}{N_m}|z-s|} (1-c_\ell) \right) dz ds. \end{aligned} \quad (6.46)$$

We note that the novel FWM efficiency is a linear combination of exponential functions, and therefore can be evaluated in a closed-form similarly to the case without MD. Solving the double integral along distance yields the following expression:

$$\begin{aligned} \mathbb{E} \left[|\eta^{(\ell)}(\omega, \omega_1, \omega_2)|^2 \right] &= m_\ell \left\{ (1+c_\ell) \left[\frac{\eta_1(\alpha_1)}{\alpha_3 + j\Delta\beta} + \left(\frac{\alpha - \alpha_1}{(\alpha_1 + j\Delta\beta)(\alpha_3 + j\Delta\beta)} \right) \right. \right. \\ &\quad \times \left. \frac{1 - e^{-2\alpha L}}{2\alpha} - \frac{e^{-(\alpha_1 + j\Delta\beta)L} - e^{-2\alpha L}}{(\alpha_3 - j\Delta\beta)(\alpha_1 + j\Delta\beta)} \right] \\ &\quad + (1-c_\ell) \left[\frac{\eta_1(\alpha_2)}{\alpha_4 + j\Delta\beta} + \left(\frac{\alpha - \alpha_2}{(\alpha_2 + j\Delta\beta)(\alpha_4 + j\Delta\beta)} \right) \right. \\ &\quad \times \left. \left. \frac{1 - e^{-2\alpha L}}{2\alpha} - \frac{e^{-(\alpha_2 + j\Delta\beta)L} - e^{-2\alpha L}}{(\alpha_4 - j\Delta\beta)(\alpha_2 + j\Delta\beta)} \right) \right] \right\} \end{aligned} \quad (6.47)$$

where the coefficients α_n with $n = 1, \dots, 4$ are MD-dependent equivalent attenuation coefficients defined as:

$$\begin{aligned} \alpha_1 &\triangleq \alpha - (q-p)\frac{\mu^2}{N_m}, & \alpha_3 &\triangleq \alpha + (q-p)\frac{\mu^2}{N_m} \\ \alpha_2 &\triangleq \alpha + (q+p)\frac{\mu^2}{N_m}, & \alpha_4 &\triangleq \alpha - (q+p)\frac{\mu^2}{N_m}. \end{aligned} \quad (6.48)$$

$\eta_1(\alpha_n)$ in Eq. (6.47) must be interpreted as the fiber kernel without MD, as in Eq. (2.16), evaluated at the equivalent attenuation in Eq. (6.48). The novel FWM efficiency expression in Eq. (6.47) is less compact than the well-known formula without

MD. Nevertheless, it involves the same basic operations and can thus be computed quickly. For any sufficiently long optical fiber, such averaged FWM efficiency can be further simplified in the elegant expression:

$$\mathbb{E} \left[|\eta^{(\ell)}(\omega, \omega_1, \omega_2)|^2 \right] \approx m_\ell \left\{ (1 + c_\ell) \frac{\alpha_1}{\alpha} |\eta_1(\alpha_1)|^2 + (1 - c_\ell) \frac{\alpha_2}{\alpha} |\eta_1(\alpha_2)|^2 \right\} \quad (6.49)$$

with $|\eta_1(\alpha_n)|^2$ the FWM efficiency without MD, as approximated in Eq. (6.18) for a long fiber, evaluated at the equivalent attenuation α_n . Once the FWM efficiency is known, the variance can be efficiently calculated through the Monte Carlo method discussed in Chapter 2.

6.3 Inter-channel MD simplification

A possible simplification to the model proposed in Sec. 6.2 is represented by the *inter-channel MD* assumption. Such a simplification consists in approximating the MD-induced phase shift of each frequency component within a WDM channel with the one experienced by the channel central frequency, hence neglecting the MD effects within a channel. The idea is sketched in Fig. 6.1, which shows the propagation constant as a step-wise function of frequency, with steps having the width of the channel bandwidth. In other words, the transfer matrix $\mathbf{U}(z, \omega)$ is sampled only at the channel center frequency, and the same value is applied to all the other channel frequencies.

Under the inter-channel MD simplification, the variance of SPM is not affected by MD. From a mathematical point of view, such simplification yields $\mathbf{P} = \mathbf{I}$ when evaluating SPM, because all the unitary transfer matrices \mathbf{U} are sampled at the same frequency, i.e., the central frequency of the channel under test. We thus focus on applying the inter-channel MD approximation to the derivation of the XPM variance.

The tensor indexes associated with a valid XPM process are \mathbf{hhii}_2 and \mathbf{hihi}_2 , where we used a unique subscript 2 as a short-hand notation to indicate channel indexes. The two sets of indexes are associated with different integration regions in

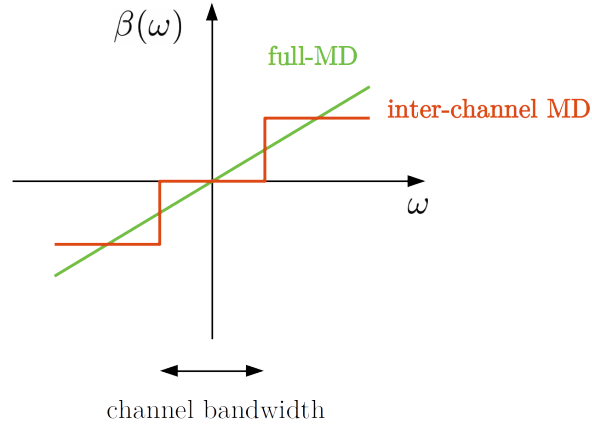


Figure 6.1: Sketch of the frequency-dependent propagation constant induced by MD on a generic mode of interest, and its inter-channel approximation, flat in each channel bandwidth.

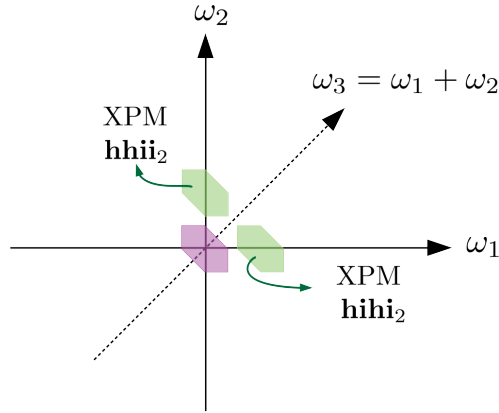


Figure 6.2: Frequency integration domains relative to XPM and SPM (see Fig. 2.1) for a 2 channel WDM signal, with emphasis on the different combinations of channel indexes for XPM.

the (ω_1, ω_2) plane of Fig. 2.1. In particular, we recall that the XPM integration regions lie along the axis. According to our notation, the combination **hhi₂** is related to the integration regions along the ω_2 axis, as sketched in Fig. 6.2 for a 2 channel signal.

frequency combination	short-hand notation	\mathbf{hhii}_2 sampling	\mathbf{hihi}_2 sampling
$\omega + \omega_1 + \omega_2$	v_1	$\Delta\omega$	$\Delta\omega$
$\omega + \omega_2$	v_2	$\Delta\omega$	0
$\omega + \omega_1$	v_3	0	$\Delta\omega$
ω	v_4	0	0

Table 6.1: Frequency sampling induced by the inter-channel MD approximation.

Since the XPM variance is additive in the number of channels, we can focus on a 2 channel setup with generic channel spacing $\Delta\omega$. Therefore, the computation of the XPM variance involves the calculation of the following terms:

$$(\mathcal{X}_{\mathbf{hhii}_2} + \mathcal{X}_{\mathbf{hihi}_2})(\mathcal{X}_{\mathbf{hhii}_2}^* + \mathcal{X}_{\mathbf{hihi}_2}^*) \quad (6.50)$$

which, in turn, requires evaluating four FWM efficiency terms: two of the type in Eq. (6.16a), and two of the type in Eq. (6.16b).

Applying the inter-channel MD idea implies performing different samplings for the two integration regions depicted in Fig. 6.2, hence for the tensors in Eq. (6.50), as summarized in Tab. 6.1. Such a sampling induces the following simplification in the matrices \mathbf{A} and \mathbf{E} , defined in Eq. (6.17), involved in the FWM efficiency:

$$\begin{aligned}
\mathcal{X}_{\mathbf{hhii}_2} \mathcal{X}_{\mathbf{hhii}_2}^* &\rightarrow \mathbf{A} = \text{Tr}[\mathbf{I}]\mathbf{I} = 2N_m \mathbf{I} \\
\mathcal{X}_{\mathbf{hihi}_2} \mathcal{X}_{\mathbf{hihi}_2}^* &\rightarrow \mathbf{A} = \text{Tr} \left[\mathbf{U}^\dagger(z, \Delta\omega) \mathbf{U}(s, \Delta\omega) \right] \mathbf{U}(z, \Delta\omega) \mathbf{U}^\dagger(s, \Delta\omega) \\
\mathcal{X}_{\mathbf{hihi}_2} \mathcal{X}_{\mathbf{hhii}_2}^* &\rightarrow \mathbf{E} = \mathbf{I} \\
\mathcal{X}_{\mathbf{hhii}_2} \mathcal{X}_{\mathbf{hihi}_2}^* &\rightarrow \mathbf{E} = \mathbf{I}.
\end{aligned} \quad (6.51)$$

We thus note that out of the four terms in Eq. (6.51) which contribute in Eq.

(6.50), three are unaffected by inter-channel MD. In particular, only in the computation of $\mathcal{X}_{\mathbf{h}\mathbf{h}\mathbf{i}\mathbf{i}_2}\mathcal{X}_{\mathbf{h}\mathbf{h}\mathbf{i}\mathbf{i}_2}^*$ the product of matrices \mathbf{P} accounting for MD does not reduce to the identity matrix. This observation allows us to significantly reduce the complexity of the calculations carried out in Sec. 6.2.1. Following the same steps, we can compute the average value of the only MD-affected term in Eq. (6.51), as:

$$\mathbb{E}[\mathbf{A}] = \frac{1}{2N_m} \left(1 + (4N_m^2 - 1)e^{-\frac{\Delta\omega^2\mu^2}{N_m}|z-s|} \right) \quad (6.52)$$

and use it to compute the average FWM efficiency $\mathbb{E}|\eta^{(1)}(\omega, \Delta\omega, 0)|^2$, see Eq. (6.16a). In the other three contributions in Eq. (6.51), which are not affected by MD, the FWM process is weighted by the FWM efficiency $|\eta_1|^2$ in Eq. (6.18).

In the variance computation, all four contributions must be taken into account. In particular, the summation of the four FWM efficiency terms requires the evaluation of the following expression:

$$\begin{aligned} & \frac{1}{2N_m} \int_0^L \int_0^L e^{-\alpha(z+s)} e^{j\Delta\beta(z-s)} \left(1 + (4N_m^2 - 1)e^{-\frac{\Delta\omega^2\mu^2}{N_m}|z-s|} \right) dz ds + (2N_m + 2)|\eta_1|^2 \\ &= \frac{(2N_m + 1)^2}{2N_m} |\eta_1|^2 - \frac{(4N_m^2 - 1)\Delta\omega^2\mu^2}{N_m^2(\alpha + j\Delta\beta)^2 - \Delta\omega^4\mu^4} \left(\frac{1 - e^{-2\alpha L}}{2\alpha} \right) \\ &+ \frac{4N_m^2 - 1}{2N_m} \left(\frac{1 - e^{-\alpha L + j\Delta\beta L - \frac{\Delta\omega^2\mu^2}{N_m}L}}{\alpha^2 - (\frac{\Delta\omega^2\mu^2}{N_m} - j\Delta\beta)^2} - \frac{e^{-\alpha L - j\Delta\beta L - \frac{\Delta\omega^2\mu^2}{N_m}L} - e^{-2\alpha L}}{\alpha^2 - (\frac{\Delta\omega^2\mu^2}{N_m} + j\Delta\beta)^2} \right) \\ &\approx \frac{2N_m + 1}{2N_m} \left((2N_m + 1)|\eta_1(\alpha)|^2 + \frac{(2N_m - 1)\alpha_m}{\alpha} |\eta_1(\alpha_m)|^2 \right) \end{aligned} \quad (6.53)$$

where the last approximation holds for a span length $L \gg 1/\alpha$. We introduced the following equivalent attenuation in Eq. (6.53):

$$\alpha_m \triangleq \alpha + \frac{\Delta\omega^2\mu^2}{N_m} \quad (6.54)$$

to compact the notation. We can now express the XPM ergodic variance, as a result of the frequency integration, in the following elegant expression:

$$\sigma_{\text{GNi}_3, \text{XPM}}^2 \approx \frac{2N_m + 1}{2N_m} \left((2N_m + 1) \sigma_{\text{GNi}_3, \text{XPM}, 1}^2(\alpha) + \frac{(2N_m - 1)\alpha_m}{\alpha} \sigma_{\text{GNi}_3, \text{XPM}, 1}^2(\alpha_m) \right) \quad (6.55)$$

where $\sigma_{\text{GNi}_3, \text{XPM}, 1}^2$ is the single-span XPM variance without any weighting factor. For the reader convenience, we report here the expression of such term:

$$\begin{aligned} \sigma_{\text{GNi}_3, \text{XPM}, 1}^2 &\triangleq \frac{1}{T^3} \kappa^2 \gamma^2 \kappa_{1;1}^3 \sum_{h,m,n} \iiint_{-\infty}^{\infty} |\eta(\omega, \omega_1, \omega_2)|^2 |\tilde{p}_h(\omega + \omega_1 + \omega_2 - \omega_h)|^2 \\ &\quad \times |\tilde{p}_m(\omega + \omega_2 - \omega_m)|^2 |\tilde{p}_i(\omega - \omega_i)|^2 |\tilde{p}_n(\omega + \omega_1 - \omega_n)|^2 \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} \frac{d\omega}{2\pi} \\ &= \frac{\kappa^2}{8/81} \mathcal{C}_{\text{GN}}(1, 1) \end{aligned} \quad (6.56)$$

with $\mathcal{C}_{\text{GN}}(1, 1)$ defined in Eq. (2.37).

The proposed simplified formula for the XPM variance with inter-channel MD sampling can be seen as an extension of the work in [119] for an arbitrary MD. If we let the parameter μ tend to 0 or to infinity, we obtain the following expressions:

$$\sigma_{\text{GNi}_3, \text{XPM}}^2 = \begin{cases} 2(2N_m + 1) \sigma_{\text{GNi}_3, \text{XPM}, 1}^2(\alpha) & \mu \rightarrow 0 \\ \frac{(2N_m + 1)^2}{2N_m} \sigma_{\text{GNi}_3, \text{XPM}, 1}^2(\alpha) & \mu \rightarrow \infty \end{cases} \quad (6.57)$$

thus matching the findings of [119].

The ultimate simplification enabled by the inter-channel MD model is the possibility to evaluate the XPM variance in a closed-form expression relying on formulas from the literature for $\sigma_{\text{GNi}_3, \text{XPM}, 1}^2$ computed at equivalent attenuation values.

6.4 Numerical results

In this section, we first test the validity of the GN model based on the inter-channel MD assumption by comparison with SSFM simulations performed in the same framework. Then, we perform full-MD simulations to investigate the goodness of the inter-channel MD assumption and to check the full-MD GN model proposed in Sec. 6.2.1.

The GN model was implemented exploiting Monte Carlo integration as described in Chapter 2 with the novel FWM efficiency, while SSFM simulations were based on a waveplate model of the optical fiber and were repeated for various realizations of the random waveplates. In this framework, we set up the SSFM such that the fiber discretization in terms of the step length and the number of waveplates adjusts to the desired value of modal dispersion. To this aim, we chose the first SSFM step h_1 such that the worst-case walk-off was equal to $1/10$ of the symbol time, and the length of each waveplate was $2h_1$. Namely, we computed the first step value by solving the equation:

$$|\beta_2|2\pi B_{\text{WDM}}h_1 + \text{SMD}\sqrt{\frac{h_1}{2}} = \frac{1}{10R} \quad (6.58)$$

and the SSFM step was then updated through the CLE criterion.

For all the simulated scenarios, the transmitted channels carried 131072 Gaussian distributed symbols, hence complying with the GN model assumptions.

6.4.1 The inter-channel MD GN model

As a first test, we validate the model for a single-mode transmission with arbitrary PMD. As discussed in the previous section, SPM is unaffected by the inter-channel PMD. Therefore, we focus on investigating the validity of the XPM variance expression in Eq. (6.55). To this aim, we limited the simulated scenarios to a 2 channels transmission. The XPM variance was isolated in SSFM simulations by a heterogeneous power allocation. Namely, the transmitted power on the CUT was $P_{\text{CUT}} = -30$ dBm while the interfering channel had power $P_{\text{INT}} = 0$ dBm, such that the SPM contribution to the NLI variance on the CUT can be safely neglected. All the XPM

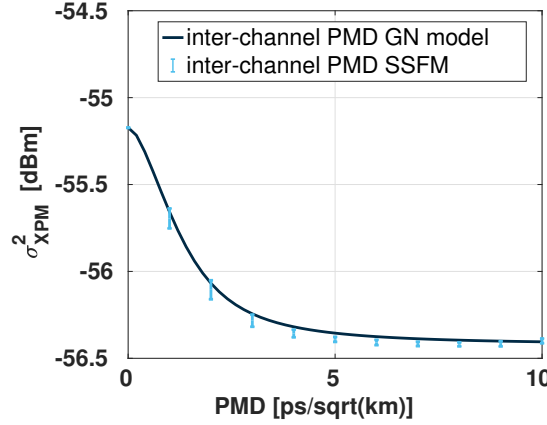


Figure 6.3: XPM variance per-polarization vs. PMD parameter for a 2 channel transmission with $\Delta f = 50$ GHz over 1 span of SMF. Extended GN model and max-min bars of SSFM for 1000 seeds, with only inter-channel PMD.

results reported in this section are then scaled to match the XPM variance of homogeneous power transmission with channel power $P = P_{\text{INT}}$, hence they are normalized to $P_{\text{CUT}} P_{\text{INT}}^2$ and then multiplied by P^3 (in linear scale).

Figure 6.3 shows the XPM variance estimation for a 50 GHz spacing transmission at 49 Gbaud, over a single-span of SMF with variable inter-channel PMD. Solid lines represent the inter-channel MD GN model results, while SSFM results are represented through bars spanning between the maximum and minimum values observed in the 1000 random realizations. It can be seen that XPM is mitigated by the presence of PMD, which contributes to channel decorrelation. In the limit of high PMD values, the XPM variance is reduced by 1.25 dB, consistently with Eq. (6.57). Apart from the excellent agreement in the whole range of values, the figure shows a small impact of the waveplates' randomness, with a maximum excursion of ≈ 0.1 dB between the maximum and minimum value, hence justifying the approach of averaging the FWM efficiency.

We then varied the main system parameters involved in the novel XPM variance expression, such as channel spacing and chromatic dispersion. To better describe

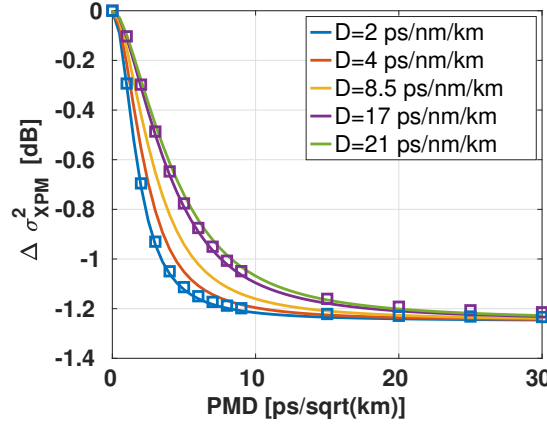


Figure 6.4: Normalized XPM variance vs. PMD parameter for a 2 channel transmission with $\Delta f = 50$ GHz over 1 span of SMF with variable dispersion. Extended ergodic GN model (lines) and SSFM (markers), with only inter-channel PMD.

the impact of PMD on XPM, we represent the results in terms of XPM variance normalized to its value without PMD, namely:

$$\Delta\sigma_{\text{XPM}}^2 \triangleq \frac{\sigma_{\text{GN}i_3, \text{XPM}}^2}{\sigma_{\text{GN}i_3, \text{XPM, no PMD}}^2}. \quad (6.59)$$

Figure 6.4 shows the normalized XPM variance $\Delta\sigma_{\text{XPM}}^2$ for the same setup of Fig. 6.3 with a variable dispersion. The GN model prediction suggests that the beneficial effect of PMD on the XPM variance is less evident at high dispersion values, where GVD masks the decorrelation yielded by PMD. As a consequence, the normalized variance curves in Fig. 6.4 require a higher value of PMD to saturate at their asymptotic value of -1.25 dB as the dispersion increases. To confirm such a behavior, we performed SSFM simulations for 100 random seeds and reported the average results with markers for the cases $D = 2$ ps/(nm·km) and $D = 17$ ps/(nm·km).

In the same spirit, we fixed the dispersion to $D = 17$ ps/(nm·km) and varied the channel spacing Δf from 50 GHz to 1 THz. The results are collected in Fig. 6.5 in terms of normalized variance as per Eq. (6.59). The figure shows that PMD is more

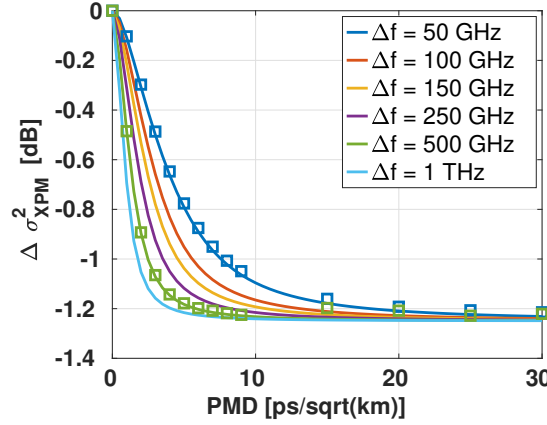


Figure 6.5: Normalized XPM variance vs. PMD parameter for a 2 channel transmission with variable channel spacing over 1 span of SMF with $D = 17$ ps/(nm·km). Extended ergodic GN model (lines) and SSFM (markers), with only inter-channel PMD.

effective in mitigating XPM for far apart channels since the phase-shift induced by modal dispersion increases with the frequency separation. In fact, it can be seen that $\Delta\sigma_{\text{XPM}}^2$ converges more rapidly to its asymptotic value for increasing values of channel spacing. The results were double-checked against SSFM simulation (markers) with 100 random seeds for $\Delta f = 50$ GHz and 500 GHz. On average, the matching with the ergodic GN model is good at both spacing values.

Finally, we validated the GN model in the presence of inter-channel MD for an SDM transmission having a variable number of strongly-coupled spatial modes. Figure 6.6 shows the XPM variance for a 2 channel transmission having $\Delta f = 100$ GHz. Solid lines refer to the GN model estimation while markers to SSFM results averaged over 10 random seeds. First of all, we note an excellent agreement, thus confirming the model validity for SDM systems with variable MD values, provided that the intra-channel MD effects are negligible. Secondly, it can be seen that all the curves eventually saturate to an asymptotic value, as for the dual-polarization case. The saturation value depends now on the number of spatial modes, namely $(2N_m + 1)/4N_m$, which

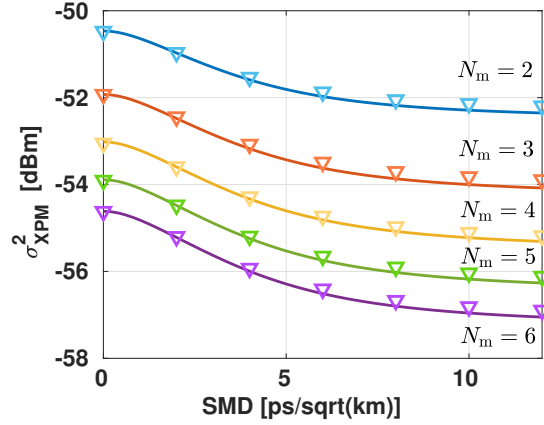


Figure 6.6: XPM variance per-polarization vs. SMD parameter for a 2 channel transmission with $\Delta f = 100$ GHz over 1 span at a variable number of modes N_m . Extended ergodic GN model (lines) and SSFM (markers), with only inter-channel MD.

coincides with -1.25 dB for $N_m = 1$.

Finally, Fig. 6.6 shows that the XPM variance is mitigated by an increasing number of spatial modes, consistently with the results reported in [10, 119] for strongly coupled modes. In particular, Eq. (6.57) suggests the following scaling of XPM with the number of modes:

$$\frac{\sigma_{\text{GN}i_3, \text{XPM}}^2}{\sigma_{\text{GN}i_3, \text{XPM}}^2(N_m = 1)} = \begin{cases} \frac{3}{2N_m + 1} & \mu \rightarrow 0 \\ \frac{1}{N_m} & \mu \rightarrow \infty \end{cases} \quad (6.60)$$

which is confirmed by the results in Fig. 6.6.

6.4.2 Full MD

We now remove the inter-channel MD simplification and devote this section to numerical results obtained with the ergodic MD-GN model presented in Sec. 6.2. To validate such a model, we performed realistic full-MD SSFM simulations as a benchmark, contrary to the previous section.

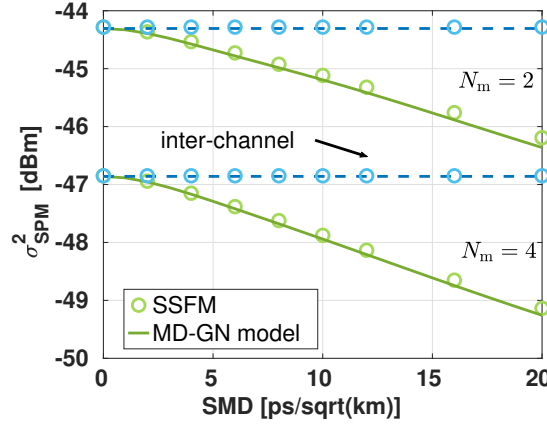


Figure 6.7: SPM variance per-polarization vs. SMD parameter over 1 span of MMF with $N_m = 2$ and $N_m = 4$ modes. Ergodic MD-GN (solid: full-MD, dashed: inter-channel MD) model and SSFM (markers).

As a first test, we simulated the transmission of a 2 channel signal with spacing 100 GHz and symbol rate 49 Gbaud, over 100 km of an MMF supporting 2 or 4 spatial modes. We estimated both the SPM and the XPM variance at variable SMD parameter values. Figure 6.7 shows excellent agreement between the SPM variance estimated by the ergodic MD-GN model (solid lines) and the SSFM simulations (markers). For the sake of completeness, we also reported the results obtained with the inter-channel MD simplification. As previously discussed, in this framework SPM is not impacted by the MD, yielding a flat line in Fig. 6.7. We note that such an approximation introduce less than 0.5 dB of error for SMD values below $\approx 5 \text{ ps}/\sqrt{\text{km}}$, which is a practical value for several deployed fibers [115, 128]. However, for increasing values of the SMD parameter, SPM vanishes when the full-MD is taken into account. Such behavior was already observed in [120], where the authors proposed a heuristic scaling of σ_{SPM}^2 with SMD to account for the intra-channel MD effects.

The estimated XPM variance for the same setup is reported in Fig. 6.8. From the figure, we note the different behavior of the XPM curves when the full-MD is taken into account. In particular, the curves do not exhibit saturation to the asymptotic

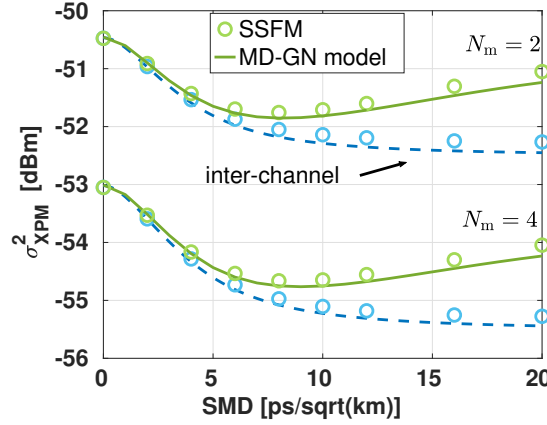


Figure 6.8: XPM variance per-polarization vs. SMD parameter for a 2 channel transmission with $\Delta f = 100$ GHz over 1 span of MMF with $N_m = 2$ and $N_m = 4$ modes. Ergodic MD-GN (solid: full-MD, dashed: inter-channel MD) model and SSFM (markers).

value predicted by inter-channel MD assumption in Eq. (6.57) in the considered range of SDM values. The ergodic MD-GN model curves well match the inter-channel MD results up to $\approx 5 \text{ ps}/\sqrt{\text{km}}$, as observed for SPM in Fig. 6.7, and they reach a plateau around $\approx 8 \text{ ps}/\sqrt{\text{km}}$. For increasing SMD values, the XPM variance is then inflated by the presence of a non-negligible intra-channel MD, thus diverging from the inter-channel MD results. Such peculiar behavior, confirmed by full-MD SSFM simulations, can be attributed to an MD-reduced walk-off.

For the same setup of Fig. 6.8 we extended the SMD values range and we ran SSFM simulation for 1000 random seeds, hence waveplates realizations. Figure 6.9 shows the histogram of the SSFM results (blurred region) compared with the model prediction. By extending the SMD axis, we observe a peak value around an SMD of $\approx 35 \text{ ps}/\sqrt{\text{km}}$, followed by a monotonic decrease for very large values of the SMD, due to the beneficial effect of MD in mitigating the accumulation of the Kerr nonlinearity.

We then tested the accuracy of the ergodic MD-GN model at variable channel spacing and fixed SMD parameter, for a transmission of $N_m = 2$ spatial modes. The

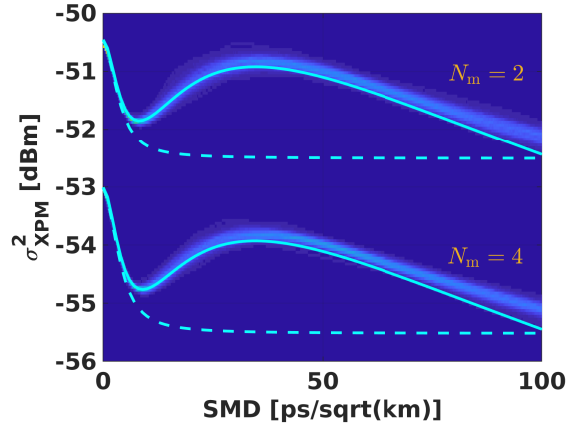
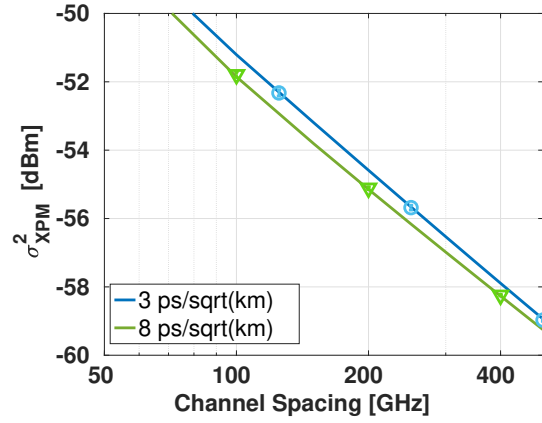


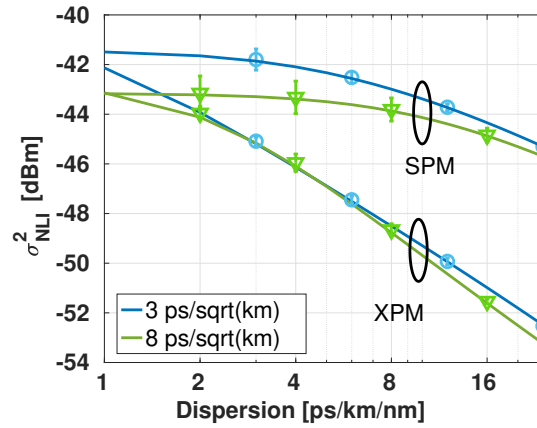
Figure 6.9: XPM variance per-polarization vs. SMD parameter for a 2 channel transmission with $\Delta f = 100$ GHz over 1 span of MMF with $N_m = 2$ and $N_m = 4$ modes. Ergodic MD-GN model (solid line), inter-channel MD GN model (dashed), and 2D histogram of the SSFM results.

SSFM and ergodic MD-GN model results at SMD 3 and 8 ps/ $\sqrt{\text{km}}$ are reported in Fig. 6.10(a). We observe an excellent agreement in the considered range of Δf up to 500 GHz. The figure also shows the max-min bars for SSFM results on top of the markers, which indicate the average values over 1500 random seeds. Such bars are very narrow, thus suggesting that the average approach used in the calculation of the fiber kernel is sufficient to capture the MD effect on XPM at realistic values of SMD. The randomness becomes more evident at high SMD (roughly > 20 ps/ $\sqrt{\text{km}}$), as visible in the histogram reported in Fig. 6.9. Similarly, Fig. 6.10(b) shows the results for a fixed channel spacing $\Delta f = 100$ GHz and variable dispersion values. Here we also plotted the SPM contribution. As expected, it can be seen that dispersion is more effective in mitigating XPM compared to SPM, thanks to the walk-off. However, we note that the slope of the variance curves versus dispersion depends on the value of the SMD parameter, being steeper at 8 ps/ $\sqrt{\text{km}}$.

Having checked the model against SSFM simulations, we now exploit it to investigate the impact of full-MD on the XPM variance by varying the system parameters.



(a)



(b)

Figure 6.10: (a) XPM variance per-polarization vs. channel spacing or (b) XPM and SPM variance vs. dispersion. 2 channel transmission over 1 span of MMF with $N_m = 2$. SMD parameter equal to 3 or 8 ps/ $\sqrt{\text{km}}$. Ergodic MD-GN model (solid) and SSFM (markers with max-min bars for 1500 seeds).

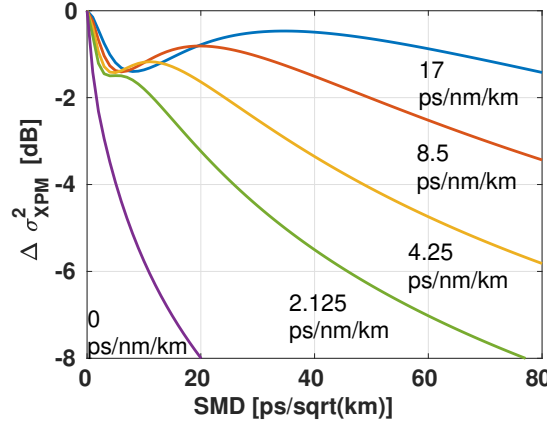


Figure 6.11: Normalized XPM variance vs. SMD parameter for a 2 channel transmission with $\Delta f = 100$ GHz over 1 span of MMF with $N_m = 2$. Results obtained with the ergodic MD-GN model.

Figure 6.11 shows $\Delta\sigma_{\text{XPM}}^2$ at different dispersion values as a function of the SMD parameter. The channel spacing between the two transmitted channels was fixed to $\Delta f = 100$ GHz, and the number of spatial modes was $N_m = 2$. The figure shows that, except for the case without dispersion, all the curves exhibit the same qualitative behavior of those reported in Fig. 6.9 for $D = 17$ ps/(nm·km). In particular, the curves exhibit a peak associated with an SMD value depending on the dispersion coefficient. For a decreasing dispersion, such inflation becomes less evident and starts to manifest at smaller SMD. In the limit case of zero-dispersion, where the decorrelation yielded by MD is not masked by the walk-off, $\Delta\sigma_{\text{XPM}}^2$ rapidly vanishes.

On the other hand, Fig. 6.12 collects the results obtained with the model for a variable channel spacing and a fixed dispersion value $D = 17$ ps/(nm·km), versus the SMD parameter. First, we note that all the curves exhibit a local minimum around the same SMD parameter ≈ 8 ps/ $\sqrt{\text{km}}$ regardless of the channel spacing. Second, it can be seen that, at high channel spacing, $\Delta\sigma_{\text{XPM}}^2$ reaches a plateau around ≈ 5 ps/ $\sqrt{\text{km}}$. Such a plateau coincides with the asymptotic value predicted by the inter-channel MD simplification, meaning that at high channel spacing the inter-channel MD effect

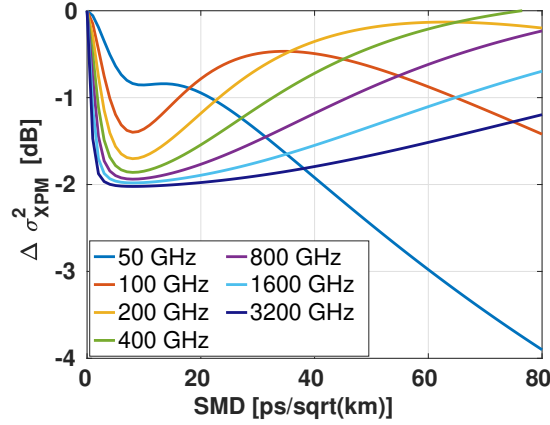


Figure 6.12: Normalized XPM variance vs. SMD parameter for a 2 channel transmission over 1 span of MMF with $D = 17$ ps/(nm·km) and $N_m = 2$. Results obtained with the ergodic MD-GN model.

dominates over the intra-channel one, for such SMD parameter values. Eventually, the XPM inflation is kicked in by a very large SMD parameter.

For the sake of completeness, we then tested the ergodic MD-GN model by varying the number of WDM channels. All the other parameters are the same as Fig. 6.9. We considered a fixed SMD value of 3 or 8 ps/√km, and varied the number of channels from 2 to 21. The results obtained with the ergodic MD-GN model (solid lines) and the SSFM (markers), in a logarithmic scale, are reported in Fig. 6.13, along with the prediction of the GN model without MD (dashed lines). The figure shows both the XPM variance and the SPM+XPM variance. The figure confirms the excellent agreement between the proposed model and the SSFM simulations even at a higher channel number count. Moreover, we note that the presence of SMD affects the slope of the variance curves. In particular, it can be seen that in the presence of SMD the variance increases less rapidly with the number of channels, thanks to the beneficial effects of MD in mitigating the XPM variance.

As the last investigation, we tested the ergodic MD-GN model for a transmission over a 5 span link. To this aim, we performed SSFM simulations for 1000 random

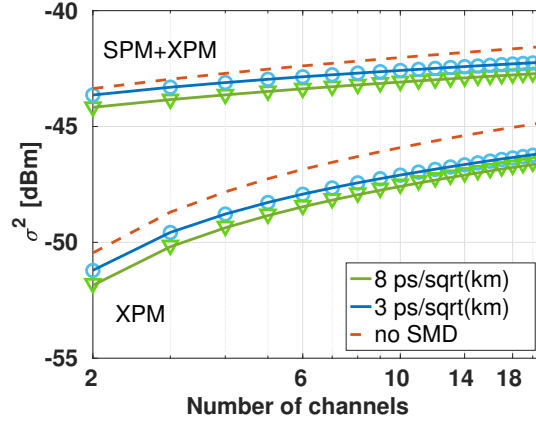


Figure 6.13: NLI variance per-polarization vs. channel spacing for a variable number of WDM channels spaced 100 GHz, over 1 span of MMF with $N_m = 2$. SMD parameter equal to 3 or 8 ps/ $\sqrt{\text{km}}$. Ergodic MD-GN model (solid), standard GN model without SMD (dashed) and SSFM (markers).

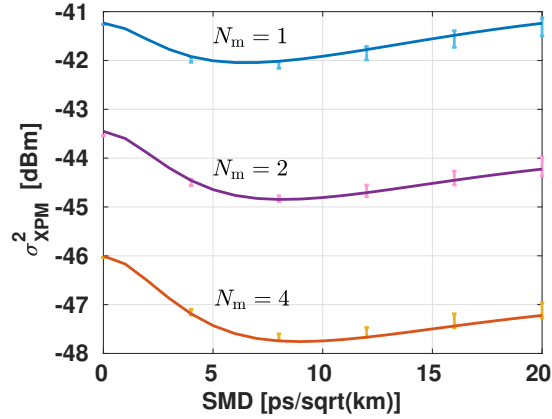


Figure 6.14: XPM variance per-polarization vs. SMD parameter for a 2 channel transmission over 5 spans, with $N_m = 1, 2, 4$ spatial modes. Ergodic MD-GN model (solid) and max-min bars of SSFM for 1000 seeds.

seeds. All the transmission parameters are the same as Fig. 6.9, with a variable number of spatial modes $N_m = 1, 2, 4$. Figure 6.14 reports the SSFM results by means of max-min bars and the ergodic MD-GN model prediction in solid lines. For the sake of simplicity, the multi-span ergodic MD-GN model is implemented relying on the phased-array term in the absence of MD, as per Eq. (2.17), while the single-span kernel is given in Eq. (6.47). The excellent agreement shown in Fig. 6.14 suggests that neglecting the MD accumulated along the spans is a good approximation. Moreover, we note that the qualitative behavior of the XPM variance curves observed for $N = 1$ is maintained even for longer links.

Conclusions

In this work, we addressed physical layer modeling problems in the context of ultra-wideband transmissions and space-division multiplexing, for high capacity optical communication systems. The main goal of the thesis was to propose models for the system performance estimation, with a particular focus on the extension of analytical models for the estimation of the NLI variance. The first part of the thesis was devoted to ultra-wideband transmissions, while the second part tackled the topic of space-division multiplexing.

In Chapter 2 we presented in detail the GN and EGN model, thus laying the theoretical foundation for the investigation of UWB and SDM extensions.

In Chapter 3 we addressed the problem of including the SRS in the EGN model. At first, we proceeded by assuming the presence of a DGE at each span-end, as typically done in the literature. Contrary to other works, we fully accounted for the interaction between the Kerr effect and the SRS in a modulation-format-aware manner. The proposed model was validated against 10 THz SSFM simulations, showing a very good agreement. We highlighted the importance of taking into account the modulation format for an accurate estimation of the frequency-tilted SNR.

In the same chapter, we extended the proposed model to practical scenarios having arbitrary positioning of the DGEs along the link for SRS gain equalization. In addition, we derived a simplified version of the model which allowed us to significantly reduce the computational time. In the framework of the GN-model, we extended closed-form expressions of the NLI variance available in the literature, to the case of an arbitrary DGE positioning. We validated the proposed model against 10

THz SSFM simulations considering a variable number of spans between the equalizers and different modulation formats, observing an excellent agreement. Most importantly, we showed that the DGE placement along the link has a serious impact on the NLI variance and on the overall SNR, and cannot be neglected.

The proposed model not only allowed us to gain a deep insight into the interaction between SRS and the Kerr effect for different transmission scenarios but also opened the path to fast power allocation optimization. In particular, we investigated the optimization of a signal power pre-emphasis in the form of an opposite SRS gain, leveraging the simplicity of the closed-form expressions extended to include such a pre-emphasis, without the need of advanced numerical optimization algorithms as usually done in the literature.

In Chapter 4 we moved the focus from the optical fiber to the optical amplifier. In particular, we investigated models for the signal evolution along the length of a semiconductor optical amplifier. The work was carried out during a six-month internship at Nokia Bell Labs (France) and was motivated by the interest that such amplifiers are gaining in the UWB context. The goal of this work was to analyze SOA models from the literature and to find a trade-off between simplicity and accuracy.

At first, we proceeded by analyzing the SOA's main characteristics and their impact on the amplifier modeling. Then, we proposed a parametric model able to capture the main SOA dynamics. In particular, we found that the space-resolution, the wavelength dependency of the gain, and the inclusion of material absorption play an important role in the SOA dynamics. Therefore, the proposed model includes space-resolution via stage concatenation, material absorption through lumped losses, and wavelength-dependency of all the parameters. After a fitting stage, the proposed model showed good accuracy in the estimation of the SNR for an SOA-amplified link with reduced computational times with respect to our benchmark.

In the remainder of the thesis, we focused on analytical models of the NLI variance in the presence of mode-dependent effects, starting with addressing the impact of PDL in a dual-polarization scenario. As a major result, in Chapter 5 we extended the EGN model to account for PDL. We presented a detailed derivation of the extended formulas and we validated the model against SSFM simulations for many

random realizations. The proposed model showed excellent accuracy and allowed to compute the system performance statistics, such as the outage probability and the PDF of the SNR, in a short time.

We showed that the impact of PDL on the NLI is different from its impact on the ASE noise. To confirm the importance of properly addressing PDL in the GN/EGN model, we showed that treating the PDL impact on the NLI in an approximated way, as typically done in the literature, yields an inaccurate estimation of the system outage probability, thus affecting the allocation of the system margins.

In Chapter 6 we extended the GN model to include modal dispersion. This work was motivated by the fact that in SDM transmissions the values of MD are typically much higher than the PMD values for single-mode transmissions. In the literature, the MD was included in the GN model only for very high values, far from practical scenarios. In this chapter, we presented a detailed derivation of the GN model extension which holds for arbitrary MD values in a context of strongly coupled modes, along with a simplified version that neglects the MD impact within the channel bandwidth, consistently with the approximations done in the literature. In both cases, we evaluated the average of the NLI variance with respect to random realizations. We referred to the novel GN model as the ergodic GN model, following a similar definition generally used for the ergodic capacity of wireless networks.

We carried out numerical validations to test the accuracy of the proposed model and to verify that the averaging process was justified. We found an excellent agreement between the SSFM and GN model average values, and we reported small fluctuations in the SSFM results with respect to the random seeds, thus confirming that the evaluation of the average NLI variance was enough to capture the impact of MD.

We then exploited the proposed model to investigate the impact of MD on the NLI for several transmission scenarios. We showed that MD has a beneficial effect in mitigating both SPM and XPM. We thus confirmed the importance of fully accounting for MD in the modeling of the NLI. Finally, we extended closed-form formulas of the GN model available in the literature for SMF to the SDM case with small, yet practical, values of modal dispersion. The novel formulas can be used to design and analyze the upcoming SDM networks based on strongly coupled modes.

Appendix A

Connelly model: extra equations and parameters

The SOA material gain can be written as $g = g' + g''$, with g' a gain coefficient and g'' an absorption coefficient, respectively:

$$g'_k(n) \triangleq \frac{\lambda_k^2}{4\sqrt{2}\pi^{3/2}n_1\tau(n)} \left(\frac{2m_cm_{hh}}{\hbar(m_c + m_{hh})} \right)^{3/2} \sqrt{\frac{\omega_k}{2\pi} - \frac{H_g(n)}{2\pi\hbar}} f_{c,k}(n)(1 - f_{v,k}(n)) \quad (\text{A.1})$$

$$g''_k(n) \triangleq \frac{\lambda_k^2}{4\sqrt{2}\pi^{3/2}n_1\tau(n)} \left(\frac{2m_cm_{hh}}{\hbar(m_c + m_{hh})} \right)^{3/2} \sqrt{\frac{\omega_k}{2\pi} - \frac{H_g(n)}{2\pi\hbar}} f_{v,k}(n)(1 - f_{c,k}(n)) \quad (\text{A.2})$$

where $H_g(n)$ is the band-gap energy, with coefficients defined in Tab A.1:

$$H_g(n) = q(1.35 - 0.775r_{As} + 0.149r_{As}^2) - qK_g n^{1/3}. \quad (\text{A.3})$$

$f_{c,k}, f_{v,k}$ in Eq. (A.1)-(A.2) are the Fermi-Dirac distributions in the CB and VB, respectively:

$$f_{c,k}(n) = \left\{ \exp\left(\frac{H_a(n) - H_{fc}(n)}{\kappa T}\right) + 1 \right\}^{-1} \quad (\text{A.4})$$

$$f_{v,k}(n) = \left\{ \exp\left(\frac{H_b(n) - H_{fv}(n)}{\kappa T}\right) + 1 \right\}^{-1} \quad (\text{A.5})$$

depending on the following functions of the carrier density:

$$H_a(n) \triangleq (\hbar\omega_k - H_g(n)) \frac{m_{hh}}{(m_e + m_{hh})} \quad (\text{A.6})$$

$$H_b(n) \triangleq -(\hbar\omega_k - H_g(n)) \frac{m_e}{(m_e + m_{hh})} \quad (\text{A.7})$$

$$H_{fc}(n) = \frac{n}{n_c} \kappa T \left(64 + 0.05524 \frac{n}{n_v} \left(64 + \sqrt{\frac{n}{n_c}} \right) \right)^{-1/4} + \ln \left(\frac{n}{n_c} \right) \quad (\text{A.8})$$

$$H_{fv}(n) = \frac{n}{n_v} \kappa T \left(64 + 0.05524 \frac{n}{n_v} \left(64 + \sqrt{\frac{n}{n_v}} \right) \right)^{-1/4} - \ln \left(\frac{n}{n_v} \right) \quad (\text{A.9})$$

with κ the Boltzmann constant, and:

$$n_v \triangleq 2 \left(\frac{m_{dh} \kappa T}{2\pi \hbar^2} \right)^{3/2}, \quad n_c \triangleq 2 \left(\frac{m_e \kappa T}{2\pi \hbar^2} \right)^{3/2}, \quad m_{dh} \triangleq \left(m_{hh}^{3/2} + m_{\ell h}^{3/2} \right)^{2/3}. \quad (\text{A.10})$$

Finally, the recombination carrier lifetime τ in Eq. (A.1)-(A.2) can be approximated as [89]:

$$\tau(n) = \frac{1}{A_{\text{rad}} + B_{\text{rad}} n}. \quad (\text{A.11})$$

All the SOA coefficients are defined in Tab. A.1 along with their values considered in this work.

Table A.1: Definitions of the symbols adopted in the SOA model and SOA parameters value used in this work.

Symbol	Description	Value in this work
T_e	Temperature [K]	300
L_a	Length [m]	700×10^{-6}
A_a	Area [m^2]	0.16×10^{-12}
V_a	Volume [m^3]	112×10^{-18}
Γ	Confinement factor	0.45
α_H	Linewidth enhancement factor	6
K_0	Carrier independent absorption [m^{-1}]	6200
K_1	Carrier dependent absorption [m^2]	7500
I	Bias current [A]	130×10^{-3}
r_1	Input facet reflectivity	5×10^{-5}
r_2	Output facet reflectivity	5×10^{-5}
α_c	Coupling losses	0.5
K_g	Band-gape shrinkage coefficient [eV]	0.9×10^{-10}
n_{eq0}	Equivalent refractive index at $n = 0$	3.22
dn_{eq}/dn	Differential of Equivalent refractive index	-1.34×10^{-26}
n_1	Active region refractive index	3.22

A_{rad}	Linear radiative recombination [s^{-1}]	1×10^7
A_{nrad}	Linear non-radiative recombination [s^{-1}]	3.5×10^8
B_{rad}	Bimolecular radiative recombination [m^3s^{-1}]	5.6×10^{-16}
B_{nrad}	Bimolecular non-radiative recombination [m^3s^{-1}]	0
C_{aug}	Auger recombination [m^6s^{-1}]	3×10^{-41}
D_{leak}	Leakage recombination [$\text{m}^{13.5}\text{s}^{-1}$]	0
m_e	Effective mass of electron in the CB [kg]	4.10×10^{-32}
m_{hh}	Effective mass of heavy hole in the VB [kg]	4.19×10^{-31}
$m_{\ell h}$	Effective mass of light hole in the VB [kg]	5.06×10^{-32}
r_{As}	Fraction of Arsenide in the active region	0.892

Bibliography

- [1] P. J. Winzer, D. T. Neilson, and A. R. Chraplyvy, “Fiber-optic transmission and networking: the previous 20 and the next 20 years,” *Opt. Express*, vol. 26, no. 18, pp. 24 190–24 239, Sep. 2018.
- [2] P. J. Winzer and D. T. Neilson, “From Scaling Disparities to Integrated Parallelism: A Decathlon for a Decade,” *J. Lightw. Technol.*, vol. 35, no. 5, pp. 1099–1115, Mar. 2017.
- [3] 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.
- [4] A. Carena, V. Curri, G. Bosco, P. Poggiolini, and F. Forghieri, “Modeling of the Impact of Nonlinear Propagation Effects in Uncompensated Optical Coherent Transmission Links,” *J. Lightw. Technol.*, vol. 30, no. 10, pp. 1524–1539, May 2012.
- [5] D. Semrau, R. Killey, and P. Bayvel, “A Closed-Form Approximation of the Gaussian Noise Model in the Presence of Inter-Channel Stimulated Raman Scattering,” *J. Lightw. Technol.*, vol. 37, no. 9, pp. 1924–1936, May 2019.
- [6] J. Proakis, *Digital Communications*, 5th ed. McGraw Hill, 2007.
- [7] G. Agrawal, *Nonlinear Fiber Optics*, 3rd ed. Academic Press, 2001.

- [8] J. N. Damask, *Polarization Optics in Telecommunications*. New York, NY, USA: Springer, 2005.
- [9] A. Bononi, R. Dar, M. Secondini, P. Serena, and P. Poggiolini, "Fiber Nonlinearity and Optical System Performance," in *Handbook of Optical Networks*. Springer International Publishing, 2020.
- [10] C. Antonelli, M. Shtaif, and A. Mecozzi, "Modeling of nonlinear propagation in Space-Division Multiplexed fiber-optic transmission," *J. Lightw. Technol.*, vol. 34, no. 1, pp. 36–54, Jan. 2016.
- [11] E. F. Schubert, *Light-Emitting Diodes*, 2nd ed. Cambridge: Cambridge University Press, 2006.
- [12] G. P. Agrawal, *Fiber-Optics Communication Systems*, 3rd ed. New York, USA: John Wiley and Sons, Inc., 2001.
- [13] Q. Lin and G. P. Agrawal, "Vector theory of stimulated Raman scattering and its application to fiber-based Raman amplifiers," *J. Opt. Soc. Am. B*, vol. 20, no. 8, pp. 1937–1952, Aug. 2003.
- [14] C. Headley III and G. P. Agrawal, "Unified description of ultrafast stimulated Raman scattering in optical fibers," *J. Opt. Soc. Am. B*, vol. 13, no. 10, pp. 2170–2177, Oct. 1996.
- [15] R. H. Stolen and E. P. Ippen, "Raman gain in glass optical waveguides," *Appl. Phys. Lett.*, vol. 22, no. 6, pp. 276–278, Mar. 1973.
- [16] D. Semrau, E. Sillekens, R. Killey, and P. Bayvel, "Modelling the Delayed Nonlinear Fiber Response in Coherent Optical Communications," *J. Lightw. Technol.*, vol. 39, no. 7, pp. 68–76, Apr. 2021.
- [17] M. Karlsson and M. Petersson, "Quaternion Approach to PMD and PDL Phenomena in Optical Fiber Systems," *J. Lightw. Technol.*, vol. 22, no. 4, pp. 1137–1146, Apr. 2004.

- [18] P. K. A. Wai and C. R. Menyuk, "Polarization mode dispersion, decorrelation, and diffusion in optical fibers with randomly varying birefringence," *J. Lightw. Technol.*, vol. 14, no. 2, pp. 148–157, Feb. 1996.
- [19] R. Dar and P. J. Winzer, "Nonlinear Interference Mitigation: Methods and Potential Gain," *J. Lightw. Technol.*, vol. 35, no. 4, pp. 903–930, Feb. 2017.
- [20] K.-P. Ho and J. M. Kahn, "Mode coupling and its impact on spatially multiplexed systems," in *Optical Fiber Telecommunications VI B*. Amsterdam, The Netherlands: Elsevier, 2013.
- [21] A. Mecozzi and R.-J. Essiambre, "Nonlinear Shannon limit in pseudolinear coherent systems," *J. Lightw. Technol.*, vol. 30, no. 12, pp. 2011–2024, Jun. 2012.
- [22] A. Vannucci, P. Serena, and A. Bononi, "The RP Method: A New Tool for the Iterative Solution of the Nonlinear Schrödinger Equation," *J. Lightw. Technol.*, vol. 20, no. 7, pp. 1102–1112, Jul. 2002.
- [23] F. D. Tappert, "Numerical solutions of the Korteweg de Vries equation and its generalizations by the split-step Fourier method," *Lect. Appl. Math. Am. Math. Soc.*, vol. 15, pp. 215–216, 1974.
- [24] G. Bosco, A. Carena, V. Curri, R. Gaudino, P. Poggiolini, and S. Benedetto, "Suppression of spurious tones induced by the split-step method in fiber systems simulation," *IEEE Photon. Technol. Lett.*, vol. 12, no. 5, pp. 489–491, May 2000.
- [25] J. Shao, X. Liang, and S. Kumar, "Comparison of Split-Step Fourier schemes for simulating fiber optic communication systems," *IEEE Photon. J.*, vol. 6, no. 4, pp. 1–15, Aug. 2014.
- [26] 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. 302–315, Jan. 2008.

- [27] O. V. Sinkin, R. Holzlöhner, J. Zweck, and C. R. Menyuk, “Optimization of the Split-step Fourier Method in Modeling Optical Fiber Communications Systems,” *J. Lightw. Technol.*, vol. 21, no. 1, pp. 1–27, Jan. 2003.
- [28] S. Musetti, P. Serena, and A. Bononi, “On the Accuracy of Split-Step Fourier Simulations for Wideband Nonlinear Optical Communications,” *J. Lightw. Technol.*, vol. 38, no. 23, pp. 5669–5677, Dec. 2018.
- [29] Q. Zhang and M. I. Hayee, “An SSF scheme to achieve comparable global simulation accuracy in WDM systems,” *IEEE Photon. Technol. Lett.*, vol. 17, no. 9, pp. 1869–1871, Sep. 2005.
- [30] P. Serena, C. Lasagni, S. Musetti, and A. Bononi, “On Numerical Simulations of Ultra-Wideband Long-Haul Optical Communication Systems,” *J. Lightw. Technol.*, vol. 38, no. 5, pp. 1019–1031, Mar. 2020.
- [31] C. Gardiner, *Handbook of Stochastic methods*, 2nd ed. New York: Springer Verlag, 1985.
- [32] R. G. Gallager, *Information Theory and Reliable Communication*. Hoboken: Wiley, 1968.
- [33] A. Alvarado, T. Fehenberger, B. Chen, and F. M. J. Willems, “Achievable Information Rates for Fiber Optics: Applications and Computations,” *J. Lightw. Technol.*, vol. 36, no. 2, pp. 424–439, Jan. 2018.
- [34] G. Bosco, A. Carena, R. Cigliutti, V. Curri, P. Poggiolini, and F. Forghieri, “Performance prediction for WDM PM-QPSK transmission over uncompensated links,” in *Proc. Optical Fiber Communications (OFC)*, Los Angeles, USA, Mar. 2011, paper OTh07.
- [35] E. Grellier and A. Bononi, “Quality parameter for coherent transmissions with Gaussian-distributed nonlinear noise,” *Opt. Express*, vol. 19, no. 13, pp. 12 781–12 788, Jun. 2012.

- [36] A. Bononi, N. Rossi, and P. Serena, "On the nonlinear threshold versus distance in long-haul highly-dispersive coherent systems," *Opt. Express*, vol. 20, no. 26, pp. B204–B216, Dec. 2012.
- [37] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, 1948.
- [38] T. Fehenberger, A. Alvarado, P. Bayvel, and N. Hanik, "On achievable rates for long-haul fiber-optic communications," *Opt. Express*, vol. 23, no. 7, pp. 9183–9191, Apr. 2015.
- [39] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "The GN-model of fiber nonlinear propagation and its applications," *J. Lightw. Technol.*, vol. 32, no. 4, pp. 694–721, Feb. 2012.
- [40] A. Carena, G. Bosco, V. Curri, Y. Jiang, P. Poggiolini, and F. Forghieri, "EGN model of non-linear fiber propagation," *Opt. Express*, vol. 22, no. 13, pp. 16 335–16 362, Jun. 2014.
- [41] P. Serena and A. Bononi, "A Time-Domain Extended Gaussian Noise Model," *J. Lightw. Technol.*, vol. 33, no. 7, pp. 1459–1472, Apr. 2015.
- [42] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Properties of nonlinear noise in long, dispersion-uncompensated fiber links," *Opt. Express*, vol. 21, no. 22, pp. 25 685–25 699, Nov. 2013.
- [43] P. Serena, C. Lasagni, and A. Bononi, "The Enhanced Gaussian Noise Model Extended to Polarization-Dependent Loss," *J. Lightw. Technol.*, vol. 38, no. 20, pp. 5685–5694, Jun. 2020.
- [44] M. Nazarathy, J. Khurgin, R. Weidenfeld, Y. Meiman, P. Cho, R. Noe, I. Shpanzter, and V. Karagodsky, "Phased-array cancellation of nonlinear FWM in coherent OFDM dispersive multi-span links," *Opt. Express*, vol. 16, no. 20, pp. 15 778–15 810, Sep. 2008.

- [45] I. S. Reed, "On a Moment Theorem for Complex Gaussian Processes," *RE Trans. Inform. Theory*, vol. 8, no. 3, pp. 194–195, Apr. 1962.
- [46] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Accumulation of nonlinear interference noise in multi-span fiber-optic systems," *Opt. Express*, vol. 22, no. 12, pp. 14 199–14 211, Jun. 2014.
- [47] G. Bosco, R. Cigliutti, A. Nespola, A. Carena, V. Curri, F. Forghieri, Y. Yamamoto, T. Sasaki, Y. Jiang, and P. Poggiolini, "Experimental investigation of nonlinear interference accumulation in uncompensated links," *IEEE Photon. Technol. Lett.*, vol. 24, no. 14, pp. 1230–1232, Jul. 2012.
- [48] 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. Express*, vol. 20, no. 2, pp. 1022–1032, Jan. 2012.
- [49] P. Poggiolini, A. Nespola, Y. Jiang, G. Bosco, A. Carena, L. Bertignono, S. M. Bilal, S. Abrate, and F. Forghieri, "Analytical and Experimental Results on System Maximum Reach Increase Through Symbol Rate Optimization," *J. Lightw. Technol.*, vol. 34, no. 8, pp. 1872–1885, Apr. 2016.
- [50] P. Serena and A. Bononi, "On the accuracy of the Gaussian nonlinear model for dispersion-unmanaged coherent links," in *European Conference on Optical Communication (ECOC)*, London, U.K., Sep. 2013, paper Th1D.3.
- [51] R. Korn, E. Korn, and G. Kroisandt, *Monte Carlo Methods and Models in Finance and Insurance*. CRC Press, 2010.
- [52] K. S. Shanmugam and P. Balaban, "A modified Monte-Carlo simulation technique for the evaluation of error rate in digital communication systems," *IEEE Trans. Commun.*, vol. 28, no. 11, pp. 1916–1924, Nov. 1980.
- [53] J.-C. Chen, D. Lu, J. S. Sadowsky, and K. Yao, "On importance sampling in digital communications-Part I: Fundamentals," *IEEE J. Sel. Areas Commun.*, vol. 11, no. 3, pp. 289–299, Apr. 1993.

- [54] N. Rossi, P. Serena, and A. Bononi, "Stratified-Sampling Estimation of PDL-Induced Outage Probability in Nonlinear Coherent Systems," *J. Lightw. Technol.*, vol. 32, no. 24, pp. 4303–4309, Dec. 2014.
- [55] I. Roberts, J. M. Kahn, J. Harley, and D. W. Boertjes, "Channel Power Optimization of WDM Systems Following Gaussian Noise Nonlinearity Model in the presence of Stimulated Raman Scattering," *J. Lightw. Technol.*, vol. 35, no. 23, pp. 5237–5250, Dec. 2017.
- [56] M. Cantono, D. Pilori, A. Ferrari, C. Catanese, J. Thouras, J. Augé, and V. Curri, "On the Interplay of Nonlinear Interference Generation With Stimulated Raman Scattering for QoT Estimation," *J. Lightw. Technol.*, vol. 36, no. 15, pp. 3131–3141, Aug. 2018.
- [57] D. Semrau, R. Killey, and P. Bayvel, "The Gaussian Noise Model in the Presence of Inter-Channel Stimulated Raman Scattering," *J. Lightw. Technol.*, vol. 36, no. 14, pp. 3046–3055, Jul 2018.
- [58] P. Poggiolini, "A generalized GN-model closed-form formula," *arXiv:1810.06545v2*, Nov. 2018.
- [59] D. Semrau, E. Sillekens, R. Killey, and P. Bayvel, "Modeling and mitigation of fiber nonlinearity in wideband optical signal transmission," *J. Opt. Commun. Netw.*, vol. 12, no. 6, pp. 68–76, Jun. 2020.
- [60] M. Cantono, R. Schmogrow, M. Newland, V. Vusirikala, and T. Hofmeister, "Opportunities and Challenges of C+L Transmission Systems," *J. Lightw. Technol.*, vol. 38, no. 5, pp. 1050–1060, Mar. 2020.
- [61] D. Semrau, E. Sillekens, R. Killey, and P. Bayvel, "A Modulation Format Correction Formula for the Gaussian Noise Model in the Presence of Inter-Channel Stimulated Raman Scattering," *J. Lightw. Technol.*, vol. 37, no. 19, pp. 5122–5131, Oct. 2019.

- [62] C. Lasagni, P. Serena, and A. Bononi, "A Raman-aware enhanced GN-model to estimate the modulation format dependence of the SNR tilt in C+L band," in *European Conference on Optical Communication (ECOC)*, Dublin, Ireland, Sep. 2019, paper W1D.2.
- [63] C. Lasagni, P. Serena, and A. Bononi, "Impact of Sparse Gain Equalization in the Presence of Stimulated Raman Scattering," in *European Conference on Optical Communication (ECOC)*, Brussels, Belgium, Dec. 2020, paper We2F.4.
- [64] C. Lasagni, P. Serena, and A. Bononi, "Modeling Nonlinear Interference With Sparse Raman-tilt Equalization," *J. Lightw. Technol.*, vol. 29, no. 15, pp. 4980–4989, Aug. 2021.
- [65] C. Lasagni, P. Serena, A. Bononi, and J.-C. Antona, "Power Allocation Optimization in the Presence of Stimulated Raman Scattering," in *European Conference on Optical Communication (ECOC)*, Bordeaux, France, 2021, paper Tu1D.3.
- [66] D. N. Christodoulides and R. B. Jander, "Evolution of stimulated raman crosstalk in wavelength division multiplexed systems," *IEEE Photon. Technol. Lett.*, vol. 8, no. 12, pp. 1722–1724, Dec. 1996.
- [67] A. R. Chraplyvy, "Optical power limits in multi-channel wavelength-division-multiplexed systems due to stimulated Raman scattering," *Electron. Lett.*, vol. 20, no. 2, pp. 58–59, Jan. 1984.
- [68] E. Pincemin, E. Grot, L. Bathany, S. Gosselin, M. Joindot, S. Bordaïs, Y. Jaouen, and J.-M. Delavaux, "Raman gain efficiencies of modern terrestrial transmission fibers in S-, C- and L-band," in *Nonlinear Guided Waves Appl.*, Stresa, Italy, Sep. 2002, paper NLTuC2.
- [69] M. Zirngibl, "Analytical model of Raman gain effects in massive wavelength division multiplexed transmission systems," *Electron. Lett.*, vol. 34, no. 8, pp. 789–790, Apr. 1998.

- [70] D. Semrau, L. Galdino, E. Sillekens, D. Lavery, R. Killey, and P. Bayvel, "Modulation format dependent, closed-form formula for estimating nonlinear interference in S+C+L band systems," in *European Conference on Optical Communication (ECOC)*, Dublin, Ireland, Sep. 2019, paper W1D.1.
- [71] A. Quarteroni, R. Sacco, and F. Saleri, *Numerical Mathematics*, 2nd ed. New York, NY, USA: Springer-Verlag, 2007.
- [72] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover Publications Inc., 1974.
- [73] A. Ferrari, D. Pileri, E. Virgillito, and V. Curri, "Power control strategies in C+L optical line systems," in *Optical Fiber Communications Conference (OFC)*, San Diego, USA, Mar. 2019, paper W2A.48.
- [74] S. Bigo, Y. Frignac, G. Charlet, W. Idler, S. Borne, H. Gross, R. Dischler, W. Poehlmann, P. Tran, C. Simonneau, D. Bayart, G. Veith, A. Jourdan, and J.-P. Hamaide, "0.2 Tbit/s (256x42.7 Gbit/s PDM/WDM) transmission over 100 km TeraLight fiber with 1.28 bit/s/Hz spectral efficiency," in *Optical Fiber Communications Conference (OFC)*, Anaheim, USA, Mar. 2001, paper PD25.
- [75] J.-X. Cai, H. G. Batshon, M. V. Mzurczyk, C. R. Davidson, O. V. Sinkin, D. Wand, M. Paskov, W. W. Patterson, M. A. Bolshtyansky, and D. G. Foursa, "94.9 Tb/s Single Mode Capacity Demonstration over 1,900 km with C+L EDFAs and Coded Modulation," in *European Conference on Optical Communication (ECOC)*, Rome, Italy, Sep. 2018.
- [76] U. C. de Moura, F. Da Ros, A. M. R. Brusin, A. Carena, and D. Zibar, "Experimental demonstration of arbitrary Raman gain-profile designs using machine learning," in *Optical Fiber Communications Conference (OFC)*, San Diego, USA, Mar. 2020, paper W1K.3.
- [77] U. C. de Moura, M. A. Iqbal, M. Kamalian, L. Krzczanowicz, F. Da Ros, A. M. R. Brusin, A. Carena, W. Forysiak, S. Turitsyn, and D. Zibar,

- “Multi-Band Programmable Gain Raman Amplifier,” *J. Lightw. Technol.*, vol. 39, no. 2, pp. 429–438, Jan. 2021.
- [78] D. Zibar, A. Ferrari, V. Curri, and A. Carena, “Machine learning-based Raman amplifier design,” in *Optical Fiber Communications Conference (OFC)*, San Diego, USA, Mar. 2019, paper M1J.1.
- [79] X. Ye, A. Arnould, A. Ghazisaeidi, D. Le Gac, and J. Renaudier, “Experimental Prediction and Design of Ultra-Wideband Raman Amplifiers Using Neural Network,” in *Optical Fiber Communications Conference (OFC)*, San Diego, USA, Mar. 2020, paper W1K.3.
- [80] M. Ionescu, A. Ghazisaeidi, and J. Renaudier, “Machine Learning Assisted Hybrid EDFA-Raman Amplifier Design for C+L Bands,” in *European Conference on Optical Communication (ECOC)*, Brussels, Belgium, Dec. 2020, paper Tu1A.14.
- [81] M. Ionescu, D. Lavery, A. Edwards, E. Sillekens, D. Semrau, L. Galdino, R. Killey, W. Pelouch, S. Barnes, and P. Bayvel, “74.38 Tb/s Transmission Over 6300 km Single Mode Fibre Enabled by C+L Amplification and Geometrically Shaped PDM-64QAM,” *J. Lightw. Technol.*, vol. 38, no. 2, pp. 531–537, Jan. 2020.
- [82] M. Karlsson, “Transmission Systems with Low Noise Phase-Sensitive Parametric Amplifiers,” *J. Lightw. Technol.*, vol. 34, no. 5, pp. 1411–1423, Mar. 2016.
- [83] M. J. Connelly, *Semiconductor Optical Amplifiers*. Berlin, Germany: Springer, 2007.
- [84] D. F. Bendimerad and Y. Frignac, “Numerical investigation of SOA nonlinear impairments for coherent transmission systems based on SOA amplification,” *J. Lightw. Technol.*, vol. 35, no. 24, pp. 5286–5295, Dec. 2017.

- [85] J. Renaudier, A. Carbo Meseguer, A. Ghazisaeidi, P. Tran, R. Rios Muller, R. Brenot, A. Verdier, F. Blache, K. Mekhazni, B. Duval, H. Debregeas, M. Achouche, A. Boutin, F. Morin, L. Letteron, N. Fontaine, Y. Frignac, and G. Charlet, "First 100-nm continuous-band WDM transmission system with 115 Tb/s transport over 100 km using novel ultra-wideband semiconductor optical amplifiers," in *European Conference on Optical Communication (ECOC)*, Goteborg, Sweden, Sep. 2017, paper Th.PDP.A3.
- [86] A. Arnould, A. Ghazisaeidi, D. Le Gac, P. Brindel, M. Makhsian, K. Mekhazni, F. Blache, M. Achouche, and J. Renaudier, "Experimental Characterization of Nonlinear Distortions of Semiconductor Optical Amplifiers in the WDM Regime," *J. Lightw. Technol.*, vol. 38, no. 2, pp. 509–513, Jan. 2020.
- [87] J. Renaudier, A. Arnould, A. Ghazisaeidi, D. Le Gac, P. Brindel, E. Awwad, M. Makhsian, K. Mekhazni, F. Blache, A. Boutin, L. Letteron, Y. Frignac, N. Fontaine, D. Neilson, and M. Achouche, "Recent Advances in 100+nm Ultra-Wideband Fiber-Optic Transmission Systems Using Semiconductor Optical Amplifiers," *J. Lightw. Technol.*, vol. 38, no. 5, pp. 1071–1079, Mar. 2020.
- [88] A. Ghazisaeidi, "Theory of Coherent WDM Systems Using In-Line Semiconductor Optical Amplifiers," *J. Lightw. Technol.*, vol. 37, no. 17, pp. 4188–4200, Sep. 2019.
- [89] M. J. Connelly, "Wide-band Semiconductor Optical Amplifier Steady-State Numerical Model," *IEEE J. Quantum Electron.*, vol. 37, no. 3, pp. 439–447, Mar. 2001.
- [90] G. P. Agrawal and N. A. Olsson, "Self-phase modulation and spectral broadening of optical pulses in semiconductor laser amplifiers," *IEEE J. Quantum Electron.*, vol. 25, no. 11, pp. 2297–2306, Nov. 1989.
- [91] D. Cassioli, S. Scotti, and A. Mecozzi, "A time-domain computer simulator of the nonlinear response of semiconductor optical amplifiers," *IEEE J. Quantum Electron.*, vol. 36, no. 9, pp. 1072–1080, Sep. 2000.

- [92] F. Vacondio, A. Ghazisaeidi, A. Bononi, and L. Rusch, ““Low-complexity compensation of SOA nonlinearity for single-channel PSK and OOK,” *J. Lightw. Technol.*, vol. 28, no. 3, pp. 277–286, Feb. 2010.
- [93] C. Antonelli, A. Mecozzi, W. Li, and L. A. Coldren, “Efficient and accurate modeling of multiwavelength propagation in SOAs: A generalized coupled-mode approach,” *J. Lightw. Technol.*, vol. 34, no. 9, pp. 2188–2197, May 2016.
- [94] W. Mathlouthi, P. Lemieaux, M. Salsi, A. Vannucci, A. Bononi, and L. Rush, “Fast and Efficient Dynamic WDM Semiconductor Optical Amplifier Model,” *J. Lightw. Technol.*, vol. 24, no. 11, pp. 4353–4365, Nov. 2006.
- [95] L. E. Nelson, C. Antonelli, A. Mecozzi, M. Birk, P. Magil, A. Schex, and L. Rapp, “Statistics of polarization dependent loss in an installed long-haul WDM system,” *Opt. Express*, vol. 19, no. 7, pp. 6790–6796, Mar. 2011.
- [96] A. Dumenil, E. Awwad, and C. Masson, “Polarization Dependent Loss: Fundamental Limits and How to Approach Them,” in *OSA Advanced Photonics Congress*, New Orleans, USA, Jul. 2017, paper SpM4F.1.
- [97] A. Nafta, E. Meron, and M. Shtaif, “Capacity limitations in fiber-optic communication systems as a result of polarization-dependent loss,” *Opt. Lett.*, vol. 34, no. 23, pp. 3613–3615, Dec. 2009.
- [98] N. Gisin, “Statistics of polarization dependent losses,” *Opt. Comm.*, vol. 114, no. 5-6, pp. 399–405, Feb. 1995.
- [99] A. Mecozzi and M. Shtaif, “The statistics of polarization-dependent loss in optical communication systems,” *IEEE Photon. Technol. Lett.*, vol. 14, no. 3, pp. 313–315, Mar. 2002.
- [100] A. Mecozzi and M. Shtaif, “Signal to noise ratio degradation caused by polarization dependent loss and the effect of dynamic gain equalization,” *J. Lightw. Technol.*, vol. 22, no. 8, pp. 1856–1871, Aug. 2004.

- [101] M. Shtaif, "Performance degradation in coherent polarization multiplexed systems as a result of polarization dependent loss," *Opt. Express*, vol. 16, no. 18, pp. 873–879, Sep. 2008.
- [102] Z. Tao, L. Dou, T. Hoshida, and J. C. Rasmussen, "A fast method to simulate the PDL impact on dual-polarization coherent systems," *IEEE Photon. Technol. Lett.*, vol. 21, no. 24, pp. 1882–1884, Dec. 2009.
- [103] C. Xie, "Polarization-dependent loss induced penalties in PDM-QPSK coherent optical communication systems," in *Optical Fiber Communications Conference (OFC)*, San Diego, USA, Mar. 2010, paper OWE6.
- [104] O. Vassilieva, S. Oda, T. Hoshida, J. C. Rasmussen, and M. Sekiya, "Experimental investigation of the statistics of the interplay between nonlinear and PDL effects in coherent polarization multiplexed systems," in *Optical Fiber Communications Conference (OFC)*, Anaheim, USA, Mar. 2013, paper OM3B.6.
- [105] I. F. J. Ruiz, A. Ghazisaeidi, E. Awwad, P. Tran, and G. Charlet, "Polarization effects in nonlinearity compensated links," in *European Conference on Optical Communication (ECOC)*, Dusseldorf, Germany, Sep. 2016, paper We1.3.3.
- [106] H.-M. Chin, D. Charlton, A. Borowiec, M. Reimer, C. Laperle, M. O'Sullivan, and S. J. Savory, "Probabilistic Design of Optical Transmission Systems," *J. Lightw. Technol.*, vol. 35, no. 4, pp. 931–940, Dec. 2017.
- [107] N. Rossi, S. Musetti, P. Ramantanis, and S. Almonacil, "The Impact of Kerr Nonlinearity on the SNR Variability Induced by Polarization Dependent Loss," *J. Lightw. Technol.*, vol. 37, no. 19, pp. 5048–5055, Oct. 2017.
- [108] A. I. A. El-Rahman and J. C. Cartledge, "Implications of Distributed Link Polarization-Dependent Loss on Bitwise Achievable Information Rates for Probabilistically Shaped and Uniform DP 64-QAM," *J. Lightw. Technol.*, vol. 37, no. 4, pp. 1187–1194, Feb. 2019.

- [109] J. C. Cartledge and A. I. A. El-Rahman, "Implications of Polarization Dependent Loss on the Performance of Probabilistically Shaped 64-QAM," in *European Conference on Optical Communication (ECOC)*, Dublin, Ireland, Sep. 2019, paper Th2B.5.
- [110] 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.
- [111] P. Serena, S. Musetti, S. Almonacil, S. Bigo, A. Bononi, P. Jenneve, N. Rossi, and P. Ramantanis, "The Gaussian noise model extended to polarization dependent loss and its application to outage probability estimation," in *European Conference on Optical Communication (ECOC)*, Rome, Italy, Sep. 2018, paper Tu4G.4.
- [112] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. New York, NY, USA: McGraw-Hill, 1991.
- [113] O. Betran-Pardo, J. Renaudier, G. Charlet, P. Tran, H. Mardoyan, M. Bertolini, M. Salsi, and S. Bigo, "Demonstration of the benefits brought by PMD in polarization-multiplexed systems," in *European Conference on Optical Communication (ECOC)*, Turin, Italy, Sep. 2010, paper Th10E.4.
- [114] P. Serena, N. Rossi, O. Betran-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, Apr. 2011.
- [115] Y. Yamada, T. Sakamoto, M. Wada, S. Nozoe, Y. Sagae, Y. Yamashita, H. Izumita, K. Nakajima, and H. Tanioka, "Design of High-Density Cable Parameters for Controlling Spatial-Mode Dispersion of Randomly Coupled Multi-Core Fibers," *J. Lightw. Technol.*, vol. 39, no. 4, pp. 1179–1185, Feb. 2021.
- [116] T. Hayashi, Y. Tamura, T. Hasegawa, and T. Taru, "Record-Low Spatial Mode Dispersion and Ultra-Low Loss Coupled Multi-Core Fiber for Ultra-Long-

- Haul Transmission,” *J. Lightw. Technol.*, vol. 35, no. 3, pp. 450–457, Feb. 2017.
- [117] G. Rademacher and K. Petermann, “Nonlinear Gaussian Noise Model for Multimode Fibers with Space-Division Multiplexing,” *J. Lightw. Technol.*, vol. 34, no. 9, pp. 2280–2287, May 2016.
- [118] A. D. Ellis, N. Mac Suibhne, F. C. Garcia Gunning, and S. Sygletos, “Expressions for the nonlinear transmission performance of multi-mode optical fiber,” *Opt. Express*, vol. 21, no. 19, pp. 22 834–22 846, Sep. 2013.
- [119] C. Antonelli, O. Golani, M. Shtaif, and A. Mecozzi, “Nonlinear interference noise in space-division multiplexed transmission through optical fibers,” *Opt. Express*, vol. 25, no. 12, pp. 13 055–13 078, Jun. 2017.
- [120] C. Antonelli, O. Golani, M. Shtaif, and A. Mecozzi, “Propagation effects in few-mode fibers,” in *European Conference on Optical Communication (ECOC)*, Goteborg, Sweden, 2017, paper W1B.1.
- [121] P. Serena, C. Lasagni, A. Bononi, C. Antonelli, and A. Mecozzi, “A Model of the Nonlinear Interference in Space-Division Multiplexed Systems with Arbitrary Modal Dispersion,” in *European Conference on Optical Communication (ECOC)*, Bordeaux, France, 2021, paper Tu2D.3.
- [122] J. P. Gordon and H. Kogelnik, “PMD fundamentals: Polarization mode dispersion in optical fibers,” *In Proc. Natl. Acad. Sci. USA*, vol. 97, no. 9, pp. 4541–4550, 2000.
- [123] M. Karlsson and J. Brentel, “Autocorrelation function of the polarization-mode dispersion vector,” *Opt. Lett.*, vol. 24, no. 14, pp. 939–941, 1999.
- [124] C. Antonelli, A. Mecozzi, M. Shtaif, and P. J. Winzer, “Stokes-space analysis of modal dispersion in fibers with multiple mode transmission,” *Opt. Express*, vol. 20, no. 11, pp. 11 718–11 733, May 2012.

- [125] R. Ryf and C. Antonelli, “Space-division Multiplexing,” in *Handbook of Optical Networks*. Springer International Publishing, 2020.
- [126] T. Hayashi, “Multi-core Fibers for Space Division Multiplexing,” in *Handbook of Optical Fibers*. Singapore: Springer, 2019.
- [127] A. Vannucci and A. Bononi, “Statistical characterization of the Jones matrix of long fibers affected by polarization mode dispersion (PMD),” *J. Lightw. Technol.*, vol. 20, no. 5, pp. 811–821, May 2002.
- [128] T. Hayashi, T. Nagashima, T. Nakanishi, T. Morishima, R. Kawawada, A. Mecozzi, and C. Antonelli, “Field-deployed multi-core fiber testbed,” in *OptoElectronics and Communications Conference (OECC) and International Conference on Photonics in Switching and Computing (PSC)*, Fukuoka, Japan, Jul. 2019, paper PDP.3.

Acknowledgements

We can all agree that a lot has happened in the past few years. In my micro-cosmo, the last three years resulted in the research presented in this thesis. None of this would have been possible without invaluable collaborations, scientific guidance, and an exceptional support system.

First and foremost I want to express my deep gratitude to my supervisor Paolo Serena, for making this PhD possible and for shaping my scientific growth throughout these years. Thank you for always keeping the bar high but never leaving me alone along the way. It was a tremendous pleasure working with you.

I would also like to extend my sincere gratitude to Alberto Bononi. His relentless enthusiasm and expertise were a constant source of inspiration, since my very first day of Communication Engineering classes. Thank you for your precious guidance.

I am grateful to the whole transmission group at Nokia Bell Labs, France, for hosting me. Special thanks to Amirhossein Ghazisaeidi for being my advisor during my visiting period. His passion and inquisitive spirit have fueled my research interest. Thank you for all the support that you gave me.

I would also like to acknowledge the other collaborations behind the work presented in this thesis. Jean-Christophe Antona from Alcatel Submarine Networks, for his great intuitions and comments. The PRIN project FIRST, for providing me the chance to collaborate and interact with such inspiring researchers.

Thanks should also go to my colleagues at the University of Parma. Matteo Lonardi and Simone Musetti, for sharing this journey with me, inside and outside this office. Federica Poli, for never wavering in her support and helping me see the silver lining even in my darkest days.

On a more personal note, I would like to thank my family and friends for keeping my mental health on track, especially in pandemic times.

A special thought goes to Luca, who is now a great expert on stimulated Raman scattering. Thank you for being there each step along the road.