UNIVERSITÀ DI PARMA

Università degli Studi di Parma

Dottorato di Ricerca in Tecnologie dell'Informazione

XXXI Ciclo

Fiber-optic Transmissions in Nonlinear Regime: Analysis of the Numerical Error in Split-step Simulations and Modeling of Polarization Dependent Loss

Coordinatore:

Chiar.mo Prof. Marco Locatelli

Tutor:

Dr. Ing. Paolo Serena

Dottorando: *Simone Musetti*

Anni 2015/2018

To my family...

Abstract

One of the main problems in nowadays fiber optical communications is to correctly asses the impact the nonlinearity of the fiber, which is one of the main impairments limiting the signal-to-noise ratio (SNR) of the communication. Over the years the problem has been tackled by means of analytical modeling, experimental investigations, and numerical simulations based on the split-step Fourier method (SSFM). Such approaches present advantages and disadvantages depending on the application. Numerical simulations are a general and precise way to estimate the impact of the nonlinearity since they can be applied to any link configuration. Although a simulation cannot account for all the physical effects occurring in an experiment, it has the advantage to abstract the investigation by isolating the interactions of interest. Such prominent advantages, however, come at the expense of the computational effort to run the algorithm, which could be excessively long in applications where the complexity is an issue. On the contrary, analytical modeling can provide results in faster computational times, but lacks the generality of the numerical simulation due to their inherent simplificative assumptions. In this thesis work, we address such methodologies by facing two different problems arising in nowadays optical systems.

First, we study the accuracy of the SSFM in wideband signal transmissions, showing that the numerical error introduced by the algorithm on the SNR in dB is power-independent and it scales quadratically with the signal bandwidth, in contrast with two common rules used in the literature to set the accuracy of the simulation. We propose a new rule to set the first step of the simulation yielding a constant error on the received SNR in dB for a wide range of signal powers, bandwidths, and fiber dispersions. It can thus be used as a universal method to simulate wideband signal propagations at constant accuracy. The scaling of the computational effort of the SSFM by increasing the signal bandwidth over the C-band is also discussed.

Second, we address the problem of modeling the nonlinearity in presence of random polarization dependent loss (PDL). We show by numerical investigations that the PDL-nonlinearity interaction may change the statistical distribution of the SNR at the receiver in a different way than what expected by current semi-analytical models developed for the linear transmission regime. We then propose for the first time an extension of the Gaussian noise model to estimate nonlinear interference variance with PDL. The new model showed an excellent accuracy in estimating the such interaction with a much smaller complexity overhead compared to SSFM-based simulations.

Contents

List of Figures

List of Tables

List of Acronyms

NLSE Nonlinear Schrödinger equation

GVD group velocity dispersion

DCF disopersion compensating fiber

SNR signal-to-noise ratio

SSFM split-step Fourier method

EDFA Erbium-doped fiber amplifier

SMF single mode fiber

DSF dispersion shifted fiber

NZDSF non-zero dispersion shifted fiber

WDM wavelength division multiplexing

SPM self-phase modulation

XPM cross-phase modulation

FWM four-wave mixing

PDM polarization division multiplexing

CNLSE coupled-nonlinear Schrödinger equation

PMD polarization mode dispersion

SOP state of polarization

ME Manakov equation

XPolM cross-polarization modulation

RP regular perturbation

RP1 first-order regular perturbation

NLI nonlinear interference

ASE amplified spontaneous emission

eRP enhanced regular perturbation

FFT fast Fourier transform

BCH Baker-Campbell-Hausdorff formula

NLP nonlinear phase criterion

CLE constant local error rule

FWM-NLP FWM-triggered nonlinear phase criterion

FWM-CLE FWM-triggered constant local error rule

PDM-16QAM polarization division multiplexing 16-quadrature amplitude modulation

CPU central processing unit

GPU graphics processing unit

AB Adams-Bashforth method

AM Adams-Moulton method

AB3 third-order Adams-Bashforth method

AB4 fourth-order Adams-Bashforth method

AM3 third-order Adams-Moulton method

PDM-QPSK polarization division multiplexing quadrature phase shift keying

GN Gaussian noise model

DU dispersion uncompensated

DM dispersion managed

CPE carrier phase estimation

I/Q in-phase/quadrature components

PSD power spectral density

PDL polarization dependent loss

DGD differential group delay

WSS wavelength selective switch

RCM reverse channel method

NLT nonlinear threshold

PDF probability density function

CDF cumulative density function

Publications

International journals

- **S. Musetti**, P. Serena, and A. Bononi, "On the Accuracy of Split-Step Fourier Simulations for Wideband Nonlinear Optical Communications," to appear on *J. Lightw. Technol.*, DOI: 10.1109/JLT.2018.2877384.

International conferences

- **S. Musetti**, P. Serena, and A. Bononi, "Constant SNR-Error Step-Size Selection Rule for Numerical Simulation of Optical Transmissions," in Proc. ECOC 17, Gothenburg, Sweden, paper P2.SC6.24, 2017
- P. Serena, **S. Musetti**, S. Almonacil, S. Bigo, A. Bononi, P. Jenneve, N. Rossi, P. Ramantanis, "The Gaussian Noise Model Extended to Polarization Dependent Loss and its Application to Outage Probability Estimation ", in Proc. ECOC 18, Rome, Italy, paper Tu4G.4, 2018

National conferences

- **S. Musetti**, P. Serena, "Investigation of Phase Noise Induced by Cross-Channel Nonlinear Interference", in Proc. Fotonica 2016, paper N. B7.7, 6-8 June 2016, Rome, Italy

Patents

- Nokia Solutions and Networks Oy, Device for Correcting the Phase Shift of a Transmission Channel, submitted to European patent application, N. 17306468.4 - 1220, date of filing 25.10.2017

Chapter 1

Introduction

The optical fiber is a nonlinear medium due to the Kerr effect, which changes the fiber refraction index when stimulated by an electric field, causing a nonlinear beating between the frequencies of the field itself. The Kerr effect thus induces distortions on the optical signal whose impact should be correctly predicted to precisely assess an optical communication. In nowadays coherent optical communications, where most of the linear impairment affecting the signals can be digitally equalized at the receiver [1–4], such a nonlinearity is one of the limiting factors of the fiber capacity.

The Kerr nonlinearity is signal dependent, thus it is impacted by all the effects that the signal experiences along the propagation. The prediction and the possible mitigation of the nonlinear distortion is thus extremely difficult, but fundamental to exploit the maximum potential of the optical fiber as a transmission medium. To this aim, the problem is usually tackled by means of *simulation* of the propagation of the signal and analytical *modeling*.

The reference equation describing nonlinear propagation of an electric field into the fiber is the *nonlinear Schrödinger equation* (NLSE). Unfortunately, a closed-form solution such an equation in unavailable for general optical fibers, thus only approximations can be used to emulate the nonlinear channel. Approximations of the NLSE can be obtained by using numerical algorithms to

solve it. Among them, the most widely used algorithm to simulate the transmission in an optical fiber is the split-step Fourier method (SSFM) [5]. Such a method has been considered the benchmark of any optical transmission for over the last forty years. In fact, by numerically simulating the signal propagation, the SSFM allows to precisely define the nonlinear distortion of the signal due to the Kerr nonlinearity and any possible influence of other effects on it.

Like any other numerical method, even the SSFM is affected by the problem of the accuracy of its solution. In order to obtain accurate simulations the price to pay is the computational effort needed to run the algorithm. Among the past years numerous attempts to optimize the computational effort of the SSFM have been made [7–12].

Another approach to the estimation of the impact of the fiber nonlinearity on the transmission is by means of modeling. In a nutshell, modeling the nonlinearity means to find an approximation of the NLSE by some simplificative assumptions to obtain some informations about the nonlinear interference (NLI) with a reduced complexity. The main difference between this approach and the SSFM is that, contrarily to the SSFM, this approach will never converge to the exact solution of the NLSE due to the above-mentioned assumptions. Moreover, the complexity of such models rely on the degree of simplification they use, i.e., the number of assumptions made to describe the system.

Over the last decade, many models for a fast estimation of the variance of the NLI have been proposed [13–21]. Due to their reduced complexity, the use of these model is wide-spreading in the context of optical networking, where a fast estimation of the signal-to-noise ratio (SNR) could be used as a quality of transmission parameter for the signal routing [22]. However, to keep the complexity to an acceptable level, these models usually do not account for random polarization effects, such as polarization mode dispersion (PMD) and polarization dependent loss (PDL). Due to their stochastic nature, these effects could affect the nonlinearity in an unexpected way, causing fluctuations of the SNR. A characterization of the impact of these effects is thus mandatory for

a correct estimation of the nonlinearity and, consequently, of the SNR of the transmission.

The work of this thesis addresses such problems. In the first part we will focus on the problem of the accuracy of the SSFM in a wideband scenario. Until now the resources to run the SSFM efficiently have been in fact more than adequate. However, nowadays optical communications system are pushing the signal bandwidth to very large values, which poses new challenges for their simulation [23,24]. In fact, in this new scenario the computational effort of the SSFM is thus expected to sky-rocket due to the fast variations of the nonlinearity induced by the fast variations of linear effects along fiber propagation. The SSFM could thus results excessively complex to simulate the nonlinearity for huge bandwidths. A fine tuning of the accuracy of the SSFM is thus mandatory, in order to obtain accurate simulations without wasting in computational effort. By an in-depth analysis of the numerical error introduced by the SSFM, we will propose a new parameter to set the accuracy of the simulation. We will show that this choice lets to control the SSFM in a wide range of optical links, thus showing to be particularly indicated to run wideband signal simulations. We will compare our proposal with exisisting rules available in literature, showing its reliability for trustable simulations. Moreover, we will discuss the potential limits of a simulation in terms of computational effort.

The second part of this work will instead focus on the nonlinear modeling in the presence of random polarization effects. As wstated before, in a context of elastic optical networks, the SNR of the transmission can be used as a quality parameter to assess the routing of the signals at the physical level. Unfortunately, with polarization related effects the complexity of performance estimation in nonlinear regime is maximized, since one shoud target estimating the probability density function of the SNR. Extending existing models to account for such stochastic effects with a minimum increment of complexity will thus give a great benefit to the nonlinear modeling. To this goal, we will investigate the interaction between the random PDL and the nonlinearity to understand their implications on the signal. Then, we will propose for the

first time an extension of a popular model to estimate the variance of the nonlinearity, namely the Gaussian noise (GN) model, to include the effect of the random PDL on it. Finally, we will draw our conclusions.

Chapter 2

Theoretical background

The propagation of a wave in a weakly nonlinear transmission medium is described by the nonlinear Schrödinger equation (NLSE). Such an equation can be applied to various field of research such as gravity waves, nonlinear optics and plasma physics.

In fiber optics, the NLSE describes the propagation of the electric field into the optical fiber [25]. Signal propagation into the fiber the signal is unfortunately distorted due to a plethora of effects, both linear and nonlinear, intrinsic of the silica composing the fiber.

The most important linear effect is the fiber attenuation, which describes the gradual loss of the signal during the propagation. This loss of light results in a lower power level of the signal at the end of the propagation with respect to the input signal. The attenuation is typically expressed by the parameter α, in dB/km, which describes the rate of loss with the distance.

Fiber loss is due to many factors, both intrinsic of the material composing the fiber or extrinsic, i.e, due to defects in the fiber manufacturing. Ultraviolet and infrared absorption, as well as the Rayleigh scattering, are the intrinsic causes of fiber loss and cannot be avoided. Extrinsic causes are instead irregularity in the interface between the fiber core and cladding, which lead to energy leakage in fiber cladding, and water absorption, which causes peaks of atten-

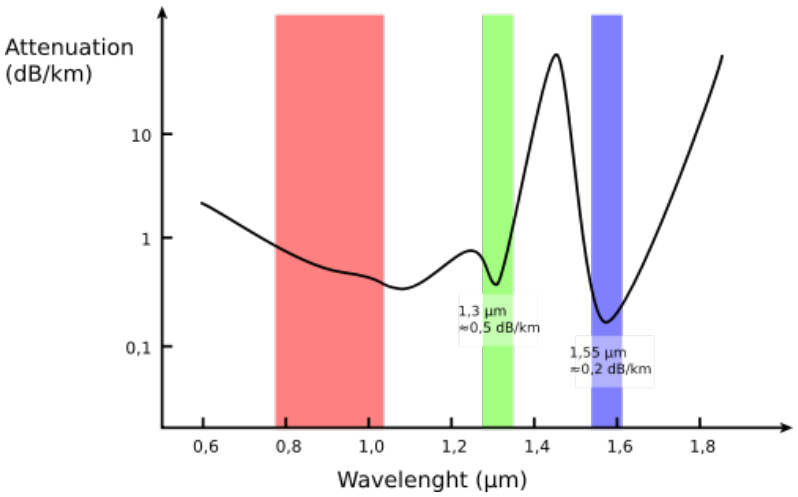

Figure 2.1: Typical wavelength spectrum of the fiber loss, with highlighted the transmissions windows of optical communications. Source: [26].

uation at given wavelengths due to infiltrations of impurity during the fiber manufacturing. Water absorption in particular is extremely important since it set, together with the advances in laser technologies, the wavelength transmission windows for optical communications. Figure 2.1 reports an example of the fiber attenuation spectrum resulting from the sum of all intrinsic and extrinsic causes. In particular, it is worth noting that the attenuation peaks in the figure are caused by water absorption. In the figure are also highlighted the transmissions windows of fiber communications. The current transmission window of nowadays communications is centered on the minimum of the attenuation spectrum in the figure, corresponding to $\lambda = 1550$ nm, granting a typical fiber loss of $\alpha \simeq 0.15 \div 0.25$ dB/km.

The other linear effects of an optical fiber are dispersive effects caused by the frequency dependence of the propagation constant $\beta = \beta(\Omega)$. Due to such a dependency, each frequency of the electrical field experiences a different propagation constant along the propagation, thus they accumulate different phases. This different phase accumulation of each frequency leads to a group

delay, i.e., a delay in the propagation of the pulses composing the optical signal, and group velocity dispersion (GVD), which causes distortion in the shape of the pulses during the propagation. While the group delay is just a temporal shift, thus not distorting the signal, the distortions due to GVD could strongly affect the transmission, creating inter-symbol interference.

Different techniques exist to mitigate such an effect. For instance, in the optical domain dispersion compensating fibers (DCFs) can be placed in the optical link to cancel the GVD accumulation along distance. With the advent of coherent detection, use of DCF to compensate for the GVD has been replaced by digital compensation after the signal detection [1–4].

The electric field propagating along the fiber experiences also nonlinear distortion due to the Kerr effect [25]. Such an effect describes the nonlinear modification of the refractive index of the silica composing the fiber in response to the passage of the electric field itself. The Kerr effect causes a nonlinear beating between different frequencies composing the signal, thus introducing a transfer of energy between them, yielding signal distortions at the receiver. Such a nonlinearity has been one of the main limiting factors of the capacity growth for optical communication since the past two decades [27–30]. Distortions due to Kerr nonlinearity are cubically dependent by the signal power, thus limiting the growth of the signal-to-noise ratio (SNR) with power as it happens in many kinds of communications in linear mediums. The typical dependence of the SNR on signal power for an optical transmission is reported in Fig. 2.2. After an initial linear growth at low power, the SNR decreases due to nonlinearity, taking a typical bell-shaped curve. By limiting the SNR of the transmission, the nonlinearity thus limits the growth of the achievable information rate with power in optical communications[1]. The Kerr nonlinearity captured the attention of the community in the past years, in particular models to characterize and possibly mitigate such a distortion [13–19].

[1]The discussion holds with matched filter detection. However, since the Shannon limit in the nonlinear regime is unknown, it might be possible to overcome such a limitation by using more advanced digital signal processing. For more information, see [31].

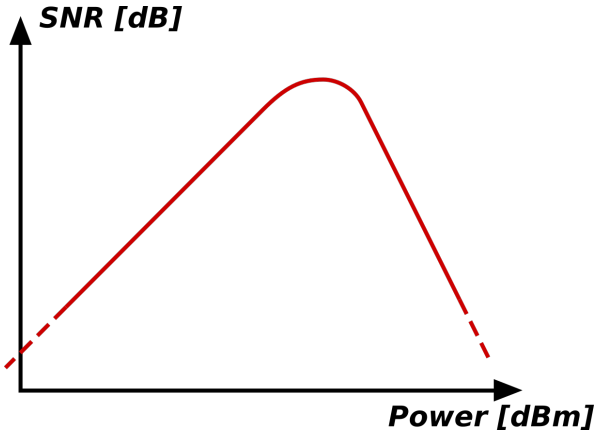

Figure 2.2: Example of a bell-curve of the SNR in optical communications. The decrease of the SNR at high powers is due the impact of Kerr nonlinearity on the transmissions.

The NLSE is the ideal benchmark for the analysis of any optical scenario. Unfortunately, except special cases, such an equation can be solved in closed-form only in presence of linear effects *or* Kerr nonlinearity alone. In a general form, i.e., with both linear and nonlinear effects, the solution of such an equation does not exist in closed form. The only way to get a reasonable approximation is to simulate the signal propagation, i.e., solving the NLSE, through the use of numerical algorithms. Such numerical algorithms come at the cost of the computational effort spent to run them.

Every numerical method to solve the NLSE introduces intrinsically an error, which comes from the approximation used to mimic the NLSE. The accuracy of a simulation describes how much the estimated solution computed by the algorithm is close to the real one. Since the computational effort generally translates in computational time, a compromise between the accuracy and the simulation time must be made. In this framework, the split-step Fourier method (SSFM) [5] founds the reasons for its popularity to solve the NLSE [32]. Over the years many studies of the SSFM has been made in order to find rules

to setup simulation for the optimization of the ratio between accuracy and computational effort [6–12, 33] .

This chapter aim to introduce the reader to the topic of the numerical resolution of the NLSE by the SSFM. First, the NLSE will be theoretically presented, in order to set the basis to understand how the SSFM method operates on it. Then a detailed analysis of the SSFM method to numerically integrate the NLSE will be reported, intended to introduce the reader in the numerical aspects of such a technique and the principles behind it.

2.1 NLSE in optical communications

The aim of the Section is to introduce the reader to the propagation of the signal in the optical fiber and to describe the main effects affecting the signal during the propagation.

2.1.1 Single polarization NLSE

Let $s(z,t)$ the scalar complex envelope of the electric field propagating within an optical fiber, with z running distance and t time. The NLSE at the time t and spatial coordinate z can be expressed in general form as:

$$\frac{\partial s\,(z,t)}{\partial z} = -\frac{\alpha}{2}s - j\beta_0 s - \beta_1\frac{\partial s}{\partial t} + j\frac{\beta_2}{2}\frac{\partial^2 s}{\partial t^2} + \frac{\beta_3}{6}\frac{\partial^3 s}{\partial t^3} - j\gamma\,(s^* s)\,s\,, \qquad (2.1)$$

where * denotes conjugation. Each right-hand side term of the equation describes a single effect manifesting on the signal during the propagation. Hence, α is the attenuation coefficient, $\beta_k,\ k = 0, ..., 3$ the dispersive coefficients, while γ the nonlinear coeffiecint. The dispersive coefficients arise from a Taylor expansion of $\beta\,(\Omega)$ truncated to the third order around a reference pulsation $\Omega_0 = 2\pi f_0$, with f_0 frequency.

Each β_i is thus:

$$\beta_i \triangleq \left.\frac{\partial^i \beta\,(\Omega)}{\partial \Omega^i}\right|_{\Omega=\Omega_0}, \qquad i = 0, 1, 2, 3$$

Please note that Ω refer to an absolute frequency, which is related to the low-pass frequency ω of the complex envelope by $\Omega \triangleq \omega + \Omega_0$.

The terms $-j\beta_0 s$ and $-\beta_1 \frac{\partial s}{\partial t}$ describe, respectively, a phase shift term and the group delay accumulated by the signal during the propagation. These two terms can be omitted in (2.1) by a proper setting of the reference system with no implications on the NLSE solution. Mathematically, it corresponds to consider a signal in a reference system that follows the accumulated phase $\beta_0 z$ in a retarded time frame $\tau \triangleq t - \beta_1 z$, i.e.:

$$a(z,t) \triangleq s(z,\tau)\, e^{j\beta_0 z} e^{-\frac{\alpha}{2}z}\,. \tag{2.2}$$

In (2.2) we also included the attenuation to simplify the notation.

Equation (2.1) in the new reference system is:

$$\frac{\partial a(z,t)}{\partial z} = j\frac{\beta_2}{2}\frac{\partial^2 a}{\partial t^2} + \frac{\beta_3}{6}\frac{\partial^3 a}{\partial t^3} - j\gamma e^{-\alpha z}\,(a^* a)\,a\,. \tag{2.3}$$

In the following the signal will be referred to such a reference system.

β_2 and β_3 describe respectively the GVD, a.k.a. the chromatic dispersion, and the dispersion slope. These two coefficients are usually given by fiber manufacturers as two parameters, D and S, whose units are respectively ps/nm/km and ps/nm^2/km, and are related to β_2 and β_3 by:

$$\beta_2 = -D\frac{\lambda_0^2}{2\pi c}\,, \qquad\qquad \beta_3 = \left(2\lambda_0 D + \lambda_0^2 S\right)\frac{\lambda_0^2}{2\pi c}\,,$$

with c the light speed in vacuum and $\lambda_0 = \frac{c}{f_0}$ the reference wavelength. The reference wavelength adopted in this thesis is $\lambda_0 = 1550$ nm. This is the central frequency of the Erbium-doped fiber amplifiers (EDFA) C-Band, which is the transmission bandwidth used nowadays in optical communications and corresponds to $f_0 = \frac{c}{\lambda_0} \simeq 193.5$ THz. Typical values for D and S are 16.7 ps/nm/km and 0.057 ps/nm^2/km, respectively, for a standard single mode fiber (SMF). Other types of fiber exist with lower values of dispersion, such as dispersion shifted fibers (DSF) and non-zero dispersion shifted fibers (NZDSF), whose values of dispersion[2] are $\simeq 0$ ps/nm/km and < 10 ps/nm/km at $\lambda_0 = 1550$

[2] According to ITU-T G.655 recommendation.

nm. Although these fibers are not best for modern optical transmissions, they were widely used in the past when optical communication links were designed with dispersion management at the optical level, and are still present in many optical links.

The solution of the NLSE in presence of only chromatic dispersion D and dispersion slope S is a complex linear filtering of the input signal. By ignoring the last term in (2.3), the optical field at a generic coordinate z is thus:

$$a(z,t) = h_{\mathrm{GVD}}(z,t) \otimes a(0,t) \ , \qquad \text{only } D+S \tag{2.4}$$

with:

$$\tilde{h}_{\mathrm{GVD}}(z,\omega) = \mathcal{F}\{h_{\mathrm{GVD}}(z,t)\} \triangleq e^{j\left(\frac{\beta_2}{2}\omega^2 + j\frac{\beta_3}{6}\omega^3\right)z} \tag{2.5}$$

the frequency response of the linear filtering, $\mathcal{F}\{.\}$ indicating Fourier transform and $\omega \triangleq 2\pi f$ the angular frequency or pulsation. The impulse response of the GVD is visible in Fig 2.3, where the real and imaginary part of $h_{\mathrm{GVD}}(z,t)$ are plotted. GVD induces inter-symbol interference in the temporal domain by enlarging the signal pulses in time. However, in frequency domain it acts as a unitary transformation, i.e., a phase shift,, hence it does not impact the power of the signal.

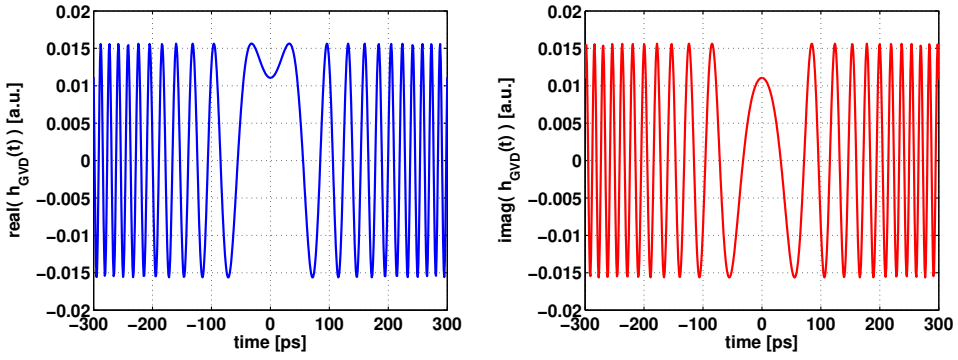

Figure 2.3: Real and imaginary parts of the GVD filter $h_{\mathrm{GVD}}(z,t)$. Here the fiber dispersion is $D = 17$ ps/nm/km and the distance $z = 30$ km.

The last term of (2.3) is the Kerr nonlinearity [25]:

$$-j\gamma e^{-\alpha z}\left(a^{*}a\right)a\,. \tag{2.6}$$

In (2.6) γ is the nonlinear coefficient of the fiber, i.e., the weight of the nonlinear Kerr effect, related to the fiber technological parameters by:

$$\gamma \triangleq \frac{2\pi n_2}{\lambda_0 A_{\text{eff}}} \qquad \left[\frac{1}{\text{W}\cdot\text{km}}\right] \tag{2.7}$$

with n_2 the nonlinear index of the fiber and A_{eff} its effective area. A typical value for SMFis $\gamma = 1.3\left[\frac{1}{\text{W}\cdot\text{km}}\right]$. As stated previously, the term $e^{-\alpha z}$ in (2.6) is the fiber attenuation. Although it is a linear effect, we found more useful to include it in the Kerr nonlinearity to improve the accuracy of the SSFM [11]. The relation $(a^{*}a)\,a$ is the local effect of Kerr nonlinearity.

The solution of the NLSE when only the nonlinearity is applied results in a time-dependent phase shift to the input signal proportional to the power of the signal itself. The optical field at a generic coordinate z is thus:

$$a\left(z,t\right) = a\left(0,t\right)e^{-j\gamma\int_0^z e^{-\alpha\xi}|a(0,t)|^2\mathrm{d}\xi}\,, \qquad \text{only Kerr effect.} \tag{2.8}$$

The integral at the exponent in (2.8) can be closed as:

$$\int_0^z e^{-\alpha\xi}\left|a\left(0,t\right)\right|^2\mathrm{d}\xi = \left|a\left(0,t\right)\right|^2 L_{\text{eff}}\left(z\right)\,,$$

where

$$L_{\text{eff}}\left(z\right) \triangleq \int_0^z e^{-\alpha\xi}\mathrm{d}\xi = \frac{1-e^{-\alpha z}}{\alpha} \tag{2.9}$$

is the effective length of the Kerr nonlinearity [25], i.e., the equivalent propagation distance in a lossless fiber giving the same nonlinear phase shift. Such a length saturates to $L_{\text{eff}} = \frac{1}{\alpha}$ for increasing distance. After that point, all the contributions to the nonlinearity can be considered negligible. For a typical attenuation coefficient $\alpha \simeq 0.2$ dB/km, L_{eff} is around 20 km.

2.1.2 Extension to wavelength division multiplexing

Wavelength division multiplexing (WDM) is typically used to fill the whole available band of the fiber, in order to increase the capacity of the transmission. In this scenario, many independent signals are multiplexed on different carriers. The total WDM signal is thus the sum of all the data streams composing each channel, i.e.:

$$a(z,t) = \sum_{n=1}^{N_{\text{ch}}} a_n(z,t)\, e^{j\omega_n t} \tag{2.10}$$

where N_{ch} is the number of different channels occupying the WDM comb, each modulated on a different carrier placed at frequency $f_n = \frac{\omega_n}{2\pi}$.

To describe the main implications of the presence of many independent channels in the WDM signal, we will consider a *separate field* approach, i.e., a set of different equations, each one describing the propagation of a single channel of the WDM comb. Due to the different frequency of each modulating carrier, all the channels experience different linear effects, i.e., different values of β_0, β_1 and β_2. Each channel experiences thus a different phase, group delay and GVD during the propagation. Each new value of $\beta_{0,n}$, $\beta_{1,n}$ and $\beta_{2,n}$, referred to the n^{th} channel of the WDM comb, can be inferred by a Taylor expansion of $\beta(\omega)$ around the channel frequency f_n, thus:

$$\beta_{i,n} \triangleq \left.\frac{\partial^i \beta(\Omega)}{\partial \Omega^i}\right|_{\Omega=\Omega_0+\omega_n} .$$

By neglecting β_3 for the sake of simplicity, the coefficients $\beta_{i,n}$ are related to the fiber dispersion coefficients evaluated at the reference frequency Ω_0 by:

$$\beta_{0,n} = \frac{\beta_2}{2}\omega_n^2 \tag{2.11}$$

$$\beta_{1,n} = \beta_2 \omega_n \tag{2.12}$$

$$\beta_{2,n} = \beta_2 \tag{2.13}$$

$$\beta_{3,n} = \beta_3 . \tag{2.14}$$

where ω_n is the low-pass central frequency of the channel n. Such relations

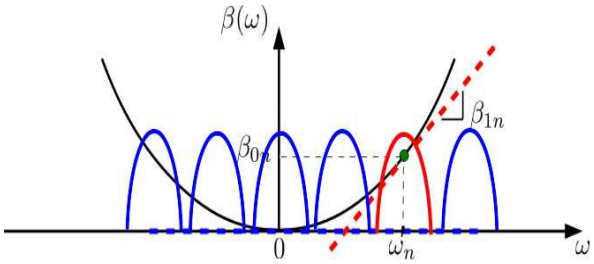

Figure 2.4: The frequency of $\beta(\omega)$ affecting the phase shift $\beta_{0,n}$ and the group delay $\beta_{1,n}$ of the generic n^{th} channel in the WDM signal spectrum centered on the frequency ω_n.

are clarified in Fig. 2.4, where a quadratic dependence on frequency of $\beta(\omega)$ is considered. Please note that by including β_3 in the Taylor series of $\beta(\omega)$ even the GVD $\beta_{2,n}$ changes with frequency. However, the impact of β_3 on all the coefficients is generally negligible unless very large bandwidths are considered. Despite each channel experiences different linear effects, such effects can still be modeled as a linear filtering different for each channel.

The main impact of such different linear effects is on the nonlinearity. In the separate-field point of view, the last term of the NLSE in (2.1) for the generic n^{th} channel[3] is [25]:

$$-j\gamma \sum_{k,l,m} \left(s_m^*(z,t)\, s_l(z,t)\right) s_k(z,t)\ . \tag{2.15}$$

The triple summation in (2.15) accounts only for the channel combinations k,l,m whose corresponding nonlinear contributions falls on the n^{th} channel, placed on the frequency ω_n. In the case of constant waves such a requirement is fulfilled by respecting the energy conservation law, which in this case assumes the following relation:

$$\omega_k + \omega_l - \omega_m - \omega_n = 0\,. \tag{2.16}$$

[3]This separation of nonlinear effects neglects FWM from next neighbor channels, which may be significant for a Nyquist WDM [14].

It is convenient for the discussion to refer to a reference system tracking $\beta_{0,n}$, $\beta_{1,n}$ and α, thus:

$$a_n = s_n\left(z, t-\beta_{1,n}z\right) e^{-\frac{\alpha}{2}z} e^{-j\beta_{0,n}z}\,.$$

The NLSE for the n^{th}channel is thus:

$$\frac{\partial a_n(z,t)}{\partial z} = j\frac{\beta_2}{2}\frac{\partial^2 a_n}{\partial t^2} + \frac{\beta_3}{6}\frac{\partial^3 a_n}{\partial t^3} - j\gamma e^{-\alpha z}\sum_{k,l,m}\left(a_m^* a_l\right) a_k e^{-j\Delta\beta z}\,. \tag{2.17}$$

In this new reference system (2.15) become thus:

$$-j\gamma e^{-\alpha z}\sum_{k,l,m}\left(a_m^* a_l\right) a_k e^{-j\Delta\beta z}\,. \tag{2.18}$$

The coefficient $\Delta\beta$ is the phase matching coefficient [6, 7, 34], defined as:

$$\Delta\beta \triangleq \beta_{0,k} + \beta_{0,l} - \beta_{0,m} - \beta_{0,n}\,, \tag{2.19}$$

which describes the phase mismatch between the channels generating the nonlinearity on the n^{th} channel. It thus weights the nonlinear contribution of each combination of the k, l, m channels.

The nonlinear effects acting on the n^{th} channel can be classified by the combinations of the indexes k, l, m [25]. For example, in the case $k = l = m = n$ the nonlinearity is defined as self-phase modulation (SPM). SPM coincides with the single channel transmission already shown in (2.8).

A similar solution is obtained for the cases where $(l = m) \neq (k = n)$. In this case (2.18) becomes:

$$-j\gamma e^{-\alpha z}\sum_{l\neq n} 2\left(a_l^* a_l\right) a_n \tag{2.20}$$

where the factor 2 account for the two degenerate cases in (2.18), i.e., $\left(a_k^* a_k\right) a_n$ and $\left(a_k^* a_n\right) a_k$. For two generic channels l and n, the solution of (2.20) is :

$$a_n(z,t) = a_n(0,t)\, e^{-j2\gamma\int_0^z e^{-\alpha\xi}|a_l(0,t+d_{ln}\xi)|^2 \mathrm{d}\xi}$$

which describe the nonlinear phase shift that a generic channel a_l induces on a_n. Such a case is thus defined as cross-phase modulation (XPM). The delay

terms $d_{ln}z$ is due to the different group velocities of the two channels, where $d_{ln} \triangleq (\beta_{1,l} - \beta_{1,n})$ is the relative group delay of channel l to channel n. It is worth noting that each channel contribution to the XPM is independent from the other, thus the solution of (2.20) accounting for all channels is:

$$a_n(z,t) = a_n(0,t)\, e^{-j2\gamma \sum_{l \neq n} \int_0^z e^{-\alpha\xi} |a_l(0, t+d_{ln}z)|^2 \mathrm{d}\xi}$$

where the summation extends to all channels but the n^{th} one. It is worth noting that both SPM and XPM are phase-matched, i.e., $\Delta\beta = 0$.

For the general case not already accounted by SPM and XPM, the nonlinearity takes the name of four-wave mixing (FWM). A substantial difference between FWM and SPM, XPM, is that for the FWM the phase matching coefficient is always $\Delta\beta \neq 0$, i.e., the channels generating nonlinearity are mismatched in phase. Depending on the value of $\Delta\beta$, each channel combination impacts differently on the total nonlinearity.

To better understand the FWM dynamics, it will be useful to switch to a more general *unique field* approach, i.e., by considering the WDM signal as a whole, wide-band signal. Under this point of view, the concept of channel does not make sense anymore. However, the separate field approach to the nonlinearity can be extended to each infinitesimal frequency composing the signal, such that summations are substituted by integrals.

Considering thus the WDM signal spectrum as $\tilde{a}(z,f) = \mathcal{F}\{a(z,t)\}$, the NLSE in frequency domain and in integral form is:

$$\begin{aligned} \tilde{a}(z,f) = \tilde{a}(0,f) - j\gamma \int_0^z \iint_{-\infty}^{\infty} e^{-\alpha\xi} e^{j\Delta\beta\xi} \tilde{a}(\xi, f+f_1) \\ \times \left[\tilde{a}^*(\xi, f+f_1+f_2)\tilde{a}(\xi, f+f_2)\right] \mathrm{d}f_1 \mathrm{d}f_2 \mathrm{d}\xi \,. \end{aligned} \tag{2.21}$$

Each combination of f, f_1, f_2 accounts for a single FWM process occurring on the frequency f. Please note that all the dispersive effects are included in $\Delta\beta$.

From (2.19) and by considering (2.16), the phase matching coefficient $\Delta\beta$ can be expressed as:

$$\Delta\beta = (2\pi)^2 \beta_2 (f - f_1)(f - f_2) \,. \tag{2.22}$$

where the frequencies $f_{k,m,n}$ are the central frequencies of the channels.[4]

To better understand the role of $\Delta\beta$ in the FWM generation, it is useful to consider an approximation of (2.21) where the signal $\tilde{a}$ is constant along distance. Such an approximation, as we will discuss later, is the first-order regular perturbation of the nonlinearity [14, 15, 17, 51]. In this case the signal $\tilde{a}$ can be taken out of the spatial integral in (2.21), leaving with:

$$\int_0^z e^{-\alpha\xi} e^{j\Delta\beta\xi} \mathrm{d}\xi = \frac{1 - e^{-(\alpha - j\Delta\beta)z}}{\alpha - j\Delta\beta}\,. \tag{2.23}$$

Equation (2.23) is the kernel [35] of the FWM , which assumes a low-pass behavior in the space identified by $\mathcal{W} \triangleq (f - f_1)(f - f_2)$ [7]. Thus, spaced away channels induce small FWM on the signal a_n. Similarly, high dispersion values helps mitigating FWM. In nowadays transmissions over SMF at high symbol rate and with a large amount of dispersion accumulated by the signal, the impact of FWM on fiber nonlinearity is smaller compared with SPM and XPM, which instead suffer less of the presence of large GVD [36]. The FWM is the most general way to describe the nonlinearity since both SPM and XPM can be considered special cases of it.

2.1.3 The Coupled NLSE and the Manakov equation

In the regime of polarization division multiplexing (PDM) as in nowadays communications the optical signal is described by a 2×1 vector $A \triangleq [\ A_x,\ A_y\ ]^T$, with x, y two orthogonal polarizations carrying two independent information signals. In this scenario the propagation is described by two coupled equations, one per polarization, generally called coupled-NLSE (CNLSE).

In dual polarization the signal propagation experiences novel effects, such as the fiber birefringence. The main causes of birefringence are geometric imperfections, fiber stresses, variations in the index profile, etc. As a result, bire-

[4] In (2.22) we excluded the third order linear effects for the sake of simplicity. However, such an inclusion can be easily done by considering the additive term $\frac{\beta_3}{6}\omega_n^3$ to (2.11) and by expanding (2.19).

fringence induces a different fiber propagation constant $\beta(\omega)$ between polarizations, and, most important, polarization crosstalk. Such effects are generally random with a correlation length of the order of meters.

Birefringence is also frequency-dependent. As a consequence it causes the two signal polarizations to experience different group delay, i.e., polarization mode dispersion (PMD) [37–39]. PMD is a random effect due to the random reasons of birefringence. Its growth along distance can be mitigated by inducing strong crosstalk between the polarizations, leading to a scaling of the differential group delay with $\sqrt{z}$ [40] instead of z as for a polarization maintaining fiber.

The random birefringence of the fiber is usually mimicked by a concatenation of wave-plates, each characterized by a randomly-oriented state of polarization (SOP). The passage of the optical signal in such plates causes its SOP to change randomly during the propagation similarly to a Brownian motion over the Poincaré sphere.

The fiber birefringence also affects the nonlinear effect. A statistical approach to the CNLSE is thus mandatory to correctly assess the signal propagation. However, since the characteristic length of the birefringence is very small compared to the fiber dispersive and nonlinear lengths, its effect on the nonlinearity can be averaged out, leading to the so-called Manakov equation (ME) [41–45]:

$$\frac{\partial A(z,t)}{\partial z} = j\frac{\beta_2}{2}\frac{\partial^2 A}{\partial t^2} + \frac{\beta_3}{6}\frac{\partial^3 A}{\partial t^3} - j\gamma\frac{8}{9}e^{-\alpha z}\left(A^\dagger A\right)A \tag{2.24}$$

where $\dagger$ stands for transpose conjugate and the fraction $\frac{8}{9}$ comes from the averaging process among all the possible SOP of the signal. The derivation of such an equation is outside the scope of this work, thus it will not be reported here. It is worth noting that (2.24) does not include PMD, thus it is an approximation. Such a simplification, however, has been shown to be accurate up to large bandwidths, in the order of THz [46]. All the following work will be thus referred to (2.24).

In a WDM scenario, the PDM signal generates another nonlinear effect

other than the three described in Section 2.1.2. In this case in fact the two degeneracy of XPM, leading to the factor 2 in (2.20), breaks down. The nonlinear operator of the Manakov equation for the n^{th} channel A_n, in polarization division multiplexing and neglecting the FWM, can be expressed as [47]:

$$\mathcal{N}(A_n) = -j\frac{8}{9}\gamma e^{-\alpha z}\left[\underbrace{A_n^{\dagger}A_n}_{(1)} + \sum_{k\neq n}\left(\underbrace{\frac{3}{2}A_k^{\dagger}A_k}_{(2)} + \underbrace{\frac{1}{2}\mathring{A}_k\cdot\bar{\sigma}}_{(3)}\right)\right]A_n \tag{2.25}$$

where $\mathring{A}_k$ is the Stokes representation of A_k, " $\cdot$ " indicates scalar product, and $\bar{\sigma} = [\sigma_1, \sigma_2, \sigma_3]$ is the Pauli spin vector, with $\sigma_k, \quad k = 1, 2, 3$ the Pauli spin matrices [48]. The three terms highlighted in (2.25) are classified as, respectively, SPM (1), XPM (2) and cross-polarization modulation (XPolM) (3). The XPolM describes the nonlinear crosstalk between the two polarizations of the signal. The factors $\frac{3}{2}$ and $\frac{1}{2}$ weighting the XPM and XPolM come from the averaging over all possible SOP of the signal [47]. While SPM and XPM are scalar terms inducing a phase rotation to A_n, XPolM is a matrix term, thus inducing polarization rotation.

2.1.4 Regular perturbation of the nonlinearity

A popular way to approximate the NLSE to obtain a closed form solution of (2.1) is by using the so-called regular perturbation method (RP) [49]. Such an approximation is largely used to deal with the modeling of nonlinearity in the fiber, allowing to study its impact on the system performance alternatively to the SSFM [14–19]. While a detailed mathematical derivation of the RP is outside of the scope of this work, it will be useful to recap the major results of such an approximation to exploit a parallelism found between RP and the SSFM method in solving the NLSE.

We start with the NLSE in (2.3):

$$\frac{\partial a(z,t)}{\partial z} = j\frac{\beta_2}{2}\frac{\partial^2 a}{\partial t^2} - j\gamma e^{-\alpha z}(a^* a)\, a\,. \tag{2.26}$$

where we dropped the dependence on β_3 for simplicity. The following procedure can be however generalized to include it. The solution with only linear effects is $a_0 = h_{\mathrm{GVD}}(z,t) \otimes a(0,t)$, i.e., the convolution with the GVD filter (2.5), also called *unperturbed solution* of the NLSE. If the nonlinearity can be considered a small perturbation to the signal, it is reasonable that the true solution $a(z,t)$ is not too much different to a_0. $a(z,t)$ can be thus expanded in Taylor series over the nonlinear coefficient γ as:

$$a(z,t) = a_0 + \gamma a_1 + \gamma^2 a_2 + ... = \sum_{p=0}^{\infty} \gamma^p a_p \tag{2.27}$$

and then truncated to the first order [17], obtaining:

$$a(z,t) \simeq a_0 + \gamma a_1$$

with a_1 the *perturbed solution*, i.e., the first-order term of the Taylor expansion in γ, thus depending only on a_0. The nonlinearity in this framework is thus an additive contribution, which is usually referred to as nonlinear interference (NLI). In the frequency domain, a_1 can be interpreted as a the sum of all FWM processes generated by the input signal spectrum frequencies [15, 50]. It is worth noting that by truncating the series to a_1 we are neglecting the higher order terms of (2.27), i.e., the higher order terms of the nonlinearity.

The problem is thus to find a_1 as a function of the unperturbed solution a_0. To do so, we substitute (2.27) in (2.26) and we equate all the terms sharing the same power of the coefficient γ. A more detailed explanation of this procedure can be found in [51]. The two equations for a_0 and a_1 result:

$$\frac{\partial a_0(z,t)}{\partial z} = j\frac{\beta_2}{2}\frac{\partial^2 a_0}{\partial t^2} \tag{2.28}$$

$$\frac{\partial a_1(z,t)}{\partial z} = j\frac{\beta_2}{2}\frac{\partial^2 a_1}{\partial t^2} - j\gamma e^{-\alpha z}\left(a_0^* a_0\right) a_0\,. \tag{2.29}$$

By switching to the frequency domain, (2.29) can be solved in closed-form

yielding:

$$\tilde{u}_{\mathrm{RP1}}(z,f) = \tilde{u}(0,f) - j\gamma \int_0^z \iint_{-\infty}^{\infty} e^{-\alpha\xi} e^{j\Delta\beta(f_1,f_2)\xi} \tilde{u}(0,f+f_1) \times \left[\tilde{u}^*(0,f+f_1+f_2)\tilde{u}(0,f+f_2)\right] \mathrm{d}f_1 \mathrm{d}f_2 \mathrm{d}\xi \,. \tag{2.30}$$

where $\tilde{a}_1(z,f) = \tilde{u}_{RP1}(z,f)\, e^{-j\beta_2(2\pi f)^2 z}$. $\tilde{u}$ can be taken out of the spatial integral in (2.30), leading to the FWM kernel (2.23) introduced in Section 2.1.2, which weight each possible frequency combination of $(\tilde{u}_0^* \tilde{u}_0)\, \tilde{u}_0$. For simplicity the integration over the distance in (2.30) extends to only one span of the link. In a generic multi-span link the FWM kernel can be expressed as:

$$\eta(z,\Delta\beta) = \int_0^z e^{-g(\xi)} e^{j\Delta\beta\xi} \mathrm{d}\xi \tag{2.31}$$

where the function:

$$g(\xi) \triangleq e^{\int_0^\xi -\alpha + \sum_k \log(G_k)\delta(x-z_k)\mathrm{d}x} \tag{2.32}$$

is the fiber gain profile along the propagation, with G_k the gain of the generic k^{th} amplifier, here considered lumped in a given coordinate of the link z_k. Such a formulation of (2.32) is a general form to account for links composed of inhomogeneous spans, as the one depicted in Fig. 2.5.

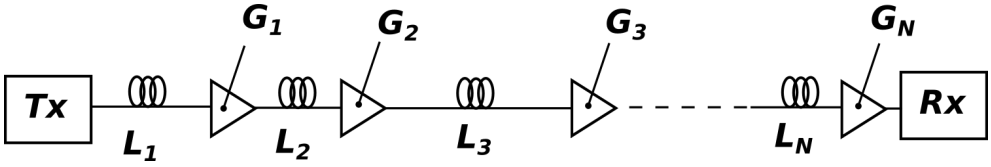

Figure 2.5: Example of optical link with inhomogeneous spans. G_k is the gain of the k^{th} amplifier, generally different span by span due to the different span length L_k.

One notable property coming from the RP1 approach is that if we factor out the signal amplitude from the integral in (2.30), the nonlinear perturbation

scales with $P^{\frac{3}{2}}$, with P power. This way, the optical link is described by a channel in additive noise whose SNR assumes the form[5]:

$$\mathrm{SNR} \triangleq \frac{P}{\sigma^2_{\mathrm{ASE}} + a_{\mathrm{NL}} P^3} \tag{2.33}$$

where:

$$a_{\mathrm{NL}} \triangleq \frac{\mathrm{Var}\left[u_{\mathrm{RP1}} - u\right]}{P^3} \tag{2.34}$$

is the unit-power variance of the NLI [52,53]. Such a behavior of the SNR has been experimentally confirmed in [54–56].

It is worth noting that the RP1 approach is limited by the perturbative assumption of the nonlinearity. When the perturbed and unperturbed solutions are comparable, the higher orders of the Taylor expansion in (2.27) become significant, changing the amount of nonlinearity. Such a requirement is often not respected in practical systems due to the average phase rotation induced by the nonlinearity on the signal. Such a problem can be bypassed by considering a reference system for the perturbed solution tracking such a phase. In a nutshell, it translates in considering the nonlinearity as $(a^* a - P)\, a$, with P a proper power. Such a method is referred to as *enhanced regular perturbation* (eRP) [51].

2.2 The split-step Fourier method

This section describes the ideas behind the SSFM and the proof that connects it to the numerical integration of the NLSE.

SSFM has been first applied to fiber optics in 1973 by Tappert *et al.* [5], and since then it has been widely used as the benchmark for the analysis of any fiber-optic system. It can be virtually used to simulate any optical link, with the only constraint of the computational effort, thus it is a reliable tool for the analysis of optical transmissions at the physical level.

[5] The SNR in (2.33) can be converted in bit error rate under the assumption of circular noise. Such assumption is discussed in Appendix A. In particular, we investigate the implications of possible phase noise that may break the circularity.

SSFM is a pseudo-spectral method to integrate differential equations. Hence, it solves the NLSE by moving back and forth between the time and the frequency by exploiting Fourier transforms. In the context of nonlinear fiber optics the equation to be integrated is the NLSE (or alternatively the ME), where the fiber linear effects and the Kerr nonlinearity in the NLSE are associated to two different operators, $\mathcal{L}$ and $\mathcal{N}$, respectively, acting on the input signal [25]. In this work, $\mathcal{L}$ and $\mathcal{N}$ are defined from (2.24) as:

$$\mathcal{L}(A) \triangleq j\frac{\beta_2}{2}\frac{\partial^2 A}{\partial t^2} + \frac{\beta_3}{6}\frac{\partial^3 A}{\partial t^3} \tag{2.35}$$

$$\mathcal{N}(A) \triangleq -j\gamma\frac{8}{9}e^{-\alpha z}\left(A^{\dagger}A\right)A \tag{2.36}$$

where we recognize, respectively, the GVD and the Kerr nonlinearity. Each operator applied alone has a closed form solution as we showed in the previous section, thus it can be easily computed numerically. In particular, the convolution of the signal with the GVD filter (2.5) can be efficiently performed in the frequency domain by a multiplication, while the nonlinear phase shift described by the nonlinear operator (2.8) can be efficiently done by a multiplication in the time domain. The SSFM takes advantage of the fact that over a distance approaching zero these two operators in the NLSE can be considered independent with a good approximation, i.e., they can be performed sequentially without impacting too much the accuracy of the solution. The switching between the frequency/time domain necessary to perform the linear/nonlinear step in the right domain can be performed by fast Fourier transforms (FFT). Such a continuous switching between the two domains is the main cause of the complexity of the SSFM.

The SSFM applied to a signal $A(0,t)$ over a distance z operates this way:

- it splits the fiber length in "*small*" steps;
- it propagates the signal along distance by cascading in each step the nonlinear and linear operators.

The output signal of each step is used as input of the next one. A basic idea of the SSFM is sketched in Fig. 2.6, which alternates the $\mathcal{N}$ and the $\mathcal{L}$ operators. The opposite alternation is possible with a similar accuracy [12].

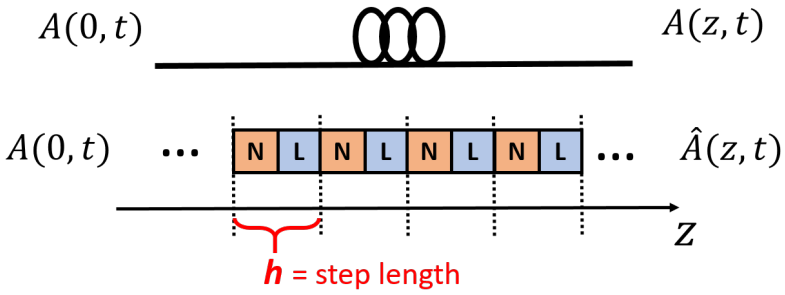

Figure 2.6: Basic SSFM scheme. The nonlinear and linear operators $\mathcal{N}$ and $\mathcal{L}$ are applied consecutively in each step h.

Unfortunately, the SSFM introduces an error in the solution of the NLSE as much as every numerical method. Such an error comes from considering the two operators separately inside the step h, and gets bigger for increasing step lengths. The strong point of the SSFM is that it is a convergent method, i.e., when the step length h approaches zero the SSFM solution $\widehat{A}(L,t)$ tends to the true solution of the NLSE $A(L,t)$, with L the propagation distance, unfortunately with an infinite computational effort.

The SSFM is generally introduced in literature starting from the Baker-Hausdorff idea [57]. Here we follow an alternative approach that highlights the relationships between the SSFM and numerical integration techniques. To relate the SSFM to the NLSE we start from the ME in operator form:

$$\frac{\partial A}{\partial z} = (\mathcal{L}+\mathcal{N})A\,. \tag{2.37}$$

where $\mathcal{L}A = \mathcal{L}(A)$ and $\mathcal{N}A = \mathcal{N}(A)$ are the two operators (2.35) and (2.36), respectively.

For the sake of convenience, we express the signal in the reference system tracking linear effects, i.e., $A(z,t) \triangleq e^{\mathcal{L}z}U(z,t)$, where $e^{\mathcal{L}z}$ can be read as

a shorthand notation for the linear convolution with the scalar GVD filter reported in (2.5). The ME thus become:

$$\frac{\partial U}{\partial z} = e^{-\mathcal{L}z}\mathcal{N}\left(e^{\mathcal{L}z}U\right)$$

which can be expressed in integral form as:

$$U(z,t) = U(0,t) + \int_0^z e^{-\mathcal{L}\xi}\mathcal{N}\left(e^{\mathcal{L}\xi}U(\xi,t)\right)\mathrm{d}\xi \tag{2.38}$$

where the interval $[0,\, z]$ can be read as the step under analysis. It is worth noting that inside the integral in (2.38) both the signal $U(\xi,t)$ and the linear operator $e^{\mathcal{L}\xi}$ depend on distance, thus they are variable inside the step. The key idea of the SSFM is to concentrate the linear effects of the whole step z in a given coordinate ξ_0 inside the step while leaving the optical field $U(\xi,t)$ free to vary. This way, (2.38) is approximated by:

$$\widehat{U}(z,t) \simeq U(0,t) + e^{-\mathcal{L}\xi_0}\int_0^z \mathcal{N}\left(e^{\mathcal{L}\xi_0}\widehat{U}(\xi,t)\right)\mathrm{d}\xi\,. \tag{2.39}$$

By performing the change of variable $E(z,t) \triangleq \widehat{U}(z,t)$ we obtain:

$$E(z,t) = E(0,t) + \int_0^z \mathcal{N}\left(E(\xi,t)\right)\mathrm{d}\xi\,, \tag{2.40}$$

where the equality $E(0,t) = e^{\mathcal{L}\xi_0}U(0,t)$ holds since at the boundary $U(0,t) \equiv \widehat{U}(0,t)$. Equation (2.40) is the ME of a purely nonlinear fiber, thus it can be solved in closed form. The solution of such an equation is:

$$E(z,t) = e^{\int_0^z \mathcal{N}\mathrm{d}\xi}E(0,t)\ .$$

By retrieving back the expression for the optical field in the original reference system, i.e., $E(z,t) = e^{\mathcal{L}(\xi_0 - z)}\widehat{A}(z,t)$, we obtain the following expression for the optical field estimated by the SSFM at the end of the step:

$$\widehat{A}(z,t) \simeq e^{\mathcal{L}(z-\xi_0)}e^{\int_0^z \mathcal{N}\mathrm{d}\xi}e^{\mathcal{L}\xi_0}A(0,t)\ , \tag{2.41}$$

where we recognize in the three exponential functions the cascading of the linear, nonlinear and linear operators $e^{\mathcal{L}\xi_0}$, $e^{\int_0^z \mathcal{N}\mathrm{d}\xi}$ and $e^{\mathcal{L}(z-\xi_0)}$, respectively. Considering for example $\xi_0 = 0$ in (2.41), such equation corresponds to the basic scheme of the SSFM reported in Fig. 2.6 of a single step. We will show in the next section two popular choices for ξ_0, leading respectively to the asymmetric and symmetric step SSFM.

2.2.1 Asymmetric and symmetric step SSFM

The value of ξ_0 in (2.41) can be freely chosen. However, the choice of ξ_0 strongly affects the efficiency of the method. The two most popular choices for ξ_0, are:

$$\widehat{A}(z,t) \simeq \begin{cases} e^{\mathcal{L}z} e^{\int_0^z \mathcal{N}\mathrm{d}\xi} A(0,t) & \xi_0 = 0, \text{ Asymm. step SSFM} \\ e^{\mathcal{L}\frac{z}{2}} e^{\int_0^z \mathcal{N}\mathrm{d}\xi} e^{\mathcal{L}\frac{z}{2}} A(0,t) & \xi_0 = \frac{z}{2}, \text{ Symm. step SSFM} \end{cases} \tag{2.42}$$

which consists respectively into applying the nonlinear operator for the whole length of the step at the beginning or in the middle of a purely linear fiber. Because of the way nonlinear and linear steps are applied the two choice are called asymmetric and symmetric SSFM, respectively. The symmetric SSFM

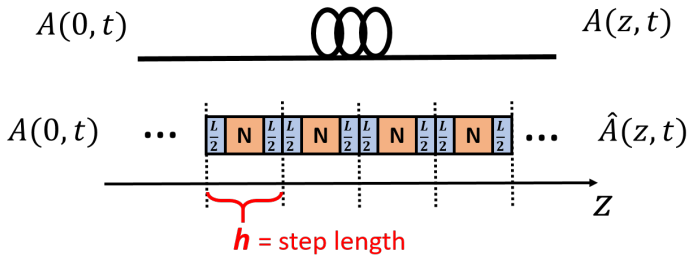

Figure 2.7: Scheme of the symmetric step SSFM. The nonlinear step $\mathcal{N}$ is sandwiched between two linear steps of half length each.

scheme is sketched in Fig. 2.7. Despite the minor difference between the asymmetric and symmetric step, the scaling of the local error due to the SSFM approximation is different, as we will show in the next Chapter.

2.2.2 Parallelism between RP1 and SSFM

In this section, we report some interesting parallelism found between the RP1 approximation of the NLSE and the SSFM.

Consider the SSFM and the RP1 solution of the NLSE in (2.39) and (2.30), respectively, reported here for clarity both with scalar field, for a propagation step from 0 to z:

$$u_{\mathrm{SSFM}}(z,t) = u(0,t) + \int_0^z e^{-\mathcal{L}\xi_0}\mathcal{N}\left(e^{\mathcal{L}\xi_0}u(\xi,t)\right)\mathrm{d}\xi\,, \qquad \text{SSFM} \qquad (2.43)$$

$$u_{\mathrm{RP1}}(z,t) = u_0 - j\gamma\int_0^z e^{-\alpha\xi}e^{j\Delta\beta\xi}\left(u_0^* u_0\right)u_0\mathrm{d}\xi\,, \qquad \text{RP1} \qquad (2.44)$$

where (2.44) is the equivalent of (2.30) in the time domain. In operator notation, the RP1 solution is alternatively written as:

$$u_{\mathrm{RP1}}(z,t) = u(0,t) + \int_0^z e^{-\mathcal{L}\xi}\mathcal{N}\left(e^{\mathcal{L}\xi}u(0,t)\right)\mathrm{d}\xi\,. \qquad (2.45)$$

By comparing (2.43) and (2.45) we can see that SSFM and RP1 follow two opposite approaches in approximating the NLSE.

The SSFM samples in the integral the spatially varying linear operator $e^{\mathcal{L}\xi} = e^{\mathcal{L}\xi_0}$ at a given spatial coordinate inside the step while leaving the optical field $u(\xi,t)$ free to vary. Such an approximation correctly integrate the nonlinearity in the step, thus SSFM considers all the orders of the nonlinearity in (2.27), but over a "virtual zero-dispersion fiber". Hence, the approximation is reliable when the GVD induce a small perturbation on the signal inside the step. On the contrary the RP1 samples the optical field $u(\xi,t) = u(0,t)$ at the beginning of the step, thus by correctly integrating only the oscillating linear operator $e^{\mathcal{L}\xi}$. Such an approximation is thus reliable when the nonlinearity generated in the step is small.

These two opposite approaches found similarities to two popular methods

for the numerical integration of oscillating functions in the form:

$$\int_0^1 f(x)\, e^{j\omega g(x)} \mathrm{d}x\,.$$

Such methods are the Gaussian quadrature and the Filon's method [58], respectively. Both rules perform a discretization of x in steps to approximate the integral, but while the in the Gaussian method the sampling frequency must follow the fast oscillation of $e^{j\omega g}$ to keep the error small, the Filon's method only samples the slowly varying function $f(x)$ and analytically integrates $e^{j\omega g}$ over each step. Note that in the framework of the NLSE the fast function is the linear operator $e^{\mathcal{L}x}$ while the slow one is the optical field u.

Chapter 3

Numerical error of the Split-step Fourier Method

In this chapter, we will analyze in depth the numerical error introduced by the SSFM. A new framework for the analysis of such an error on the SNR will be presented, which is more suitable for nowadays optical communications based on coherent detection. We will also discuss several properties of the SSFM error coming from this approach. The parallelism between the SSFM method and well–known rules for the numerical integration will be used through the discussion to help understanding the reason of the SSFM accuracy.

The most common ways to control the SSFM error will be then analyzed under the new point of view. We will show a novel parameter to set the simulation accuracy independently from signal power, bandwidth and fiber dispersion. The new parameter will be numerically analyzed and compared with alternative rules available in literature in terms of both scaling of the error and computational effort. Then, a comparison between SSFM and other alternative methods to solve the NLSE will be discussed. Finally, the proposed techniques will be used to investigate the nonlinearity in a case of study.

3.1 SSFM error on SNR

The accuracy of a simulation should be referred to the target parameter on which performance is evaluated. Typically, the error of the SSFM is considered an additive contribution on the true solution, i.e.:

$$e_{\mathrm{SSFM}} \triangleq \widehat{A}(z,t) - A(z,t) \tag{3.1}$$

where $\widehat{A}(z,t)$ is the numerical solution of the NLSE and $A(z,t)$ is the unknown true solution. Since such an error is a random process in time t, the accuracy is often translated in a more practical form as [8–12]:

$$\frac{\|e_{\mathrm{SSFM}}\|}{\|A(z,t)\|} = \frac{\left\|\widehat{A}(z,t) - A(z,t)\right\|}{\|A(z,t)\|} \tag{3.2}$$

where $\|.\|$ indicates a signal norm. The normalization in (3.2) relates thus the numerical error to the signal amplitude, which is useful to control such an error over the pulse shape of the signal.

However, in coherent optical communications, the performance of an optical system is typically expressed in terms of the SNR. The SNR can be usually directly converted to bit error rate (BER), for instance under the assumption of additive Gaussian noise. In this framework, the accuracy of the simulation is therefore the accuracy in the SNR estimation. Moreover, since the SNR is generally expressed in decibels, a reliable SSFM simulation should be such that the SNR is estimated with a bounded accuracy error in dB, as sketched in Fig. 3.1 by the constant error bars. The dashed line indicates the SNR considering just the numerical error of the simulation, that should have a fixed gap in a dB scale from the "true" SNR in order to have a fixed accuracy. We proceed now to analyze the numerical error under this novel point of view.

3.1.1 Perturbative approach to the SSFM error

The error introduced by the SSFM comes from sampling the space-dependent operator $\mathcal{L}\xi$ in (2.38) at a given coordinate $z = z_0$. Such an error should

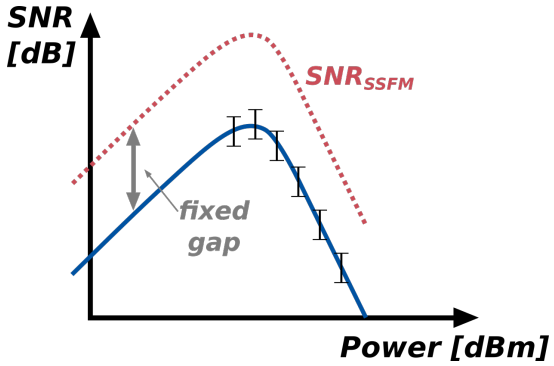

Figure 3.1: Visual representation of an ideal simulation setup on the SNR of the communication. The error bars on the estimated curve should remain constant whatever the transmitted signal power. The dashed curve in b) ($\mathrm{SNR}_{\mathrm{SSFM}}$) represents the SNR due only to the numerical error, that should have a fixed gap from the estimated SNR curve.

thus depend on how different is the oscillating function $e^{\mathcal{L}\xi}$ inside the step with respect to the fixed value $e^{\mathcal{L}\xi_0}$. Although the SSFM does not need a perturbative assumption to solve the NLSE as the RP1 approximation does, it is useful to use some of the assumptions of the RP1 for the NLI to obtain some properties on e_{SSFM}. This is not a limitation since the error of a good SSFM simulation usually is small. First of all, we consider a perturbative assumption of the Kerr nonlinearity [14,16,17]. In this framework, the error of SSFM e_{SSFM}, which is generated by a "*bad*" simulation of the NLI, will be a perturbation to the signal too. The channel can be thus considered as an additive noise channel with two different noise sources: one the NLI and one the SSFM error, as depicted in Fig. 3.2. By assuming the NLI and the SSFM error uncorrelated for simplicity, a new formulation of the SNR can be defined from (2.33) in the

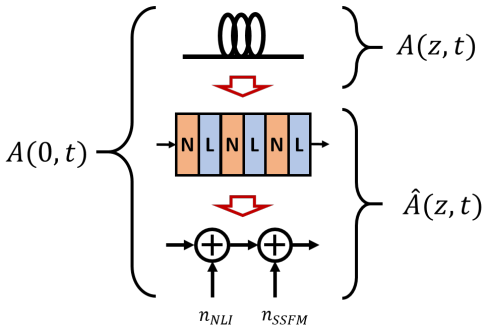

Figure 3.2: Progressive abstraction of the fiber channel first by the SSFM approximation and then by the perturbative assumption of e_{SSFM}.

framework of SSFM as[1]:

$$\widehat{\mathrm{SNR}} \triangleq \frac{P}{\sigma^2_{\mathrm{ASE}} + \sigma^2_{\mathrm{NLI}} + \sigma^2_{\mathrm{SSFM}}}$$

where P is the signal power, while σ^2_{NLI} and σ^2_{SSFM} are respectively the variances of NLI [14] and SSFM error under the perturbative assumptions, with uncorrelation between NLI, SSFM error and the propagating signal. σ^2_{ASE} is the variance of the ASE introduced by EDFAs, which is also uncorrelated with the signal and all the other noises. The true SNR under perturbative assumptions is (2.33):

$$\mathrm{SNR} \triangleq \frac{P}{\sigma^2_{\mathrm{ASE}} + \sigma^2_{\mathrm{NLI}}}$$

which corresponds to $\widehat{\mathrm{SNR}}\left(\sigma^2_{\mathrm{SSFM}} = 0\right)$. It is worth noting that in the framework of SSFM the obtained SNR estimation is always smaller than the true one due to σ^2_{SSFM}.

[1]Such an expression of the SNR describe completely the system only in the case of additive circular noises, as in the assumption of the GN model. In some optical links the circularity of the NLI can be argued, since the NLI could manifest partially as phase noise. However, as we demonstrate in Appendix A, the eventual non-circularity of the NLI (and by extension of e_{SSFM}) does not affect its total variance, thus this SNR formula can still be used.

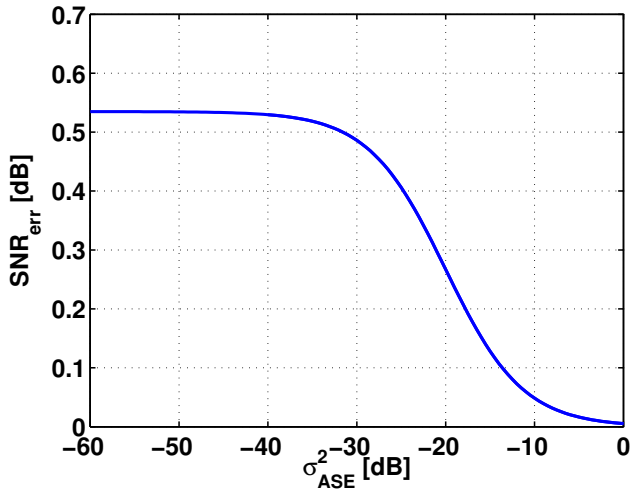

Figure 3.3: Example of SNR error as a function of the ASE noise variance σ^2_{ASE}. In this example $\sigma^2_{\mathrm{NLI}} = -20.28$ dB and $\sigma^2_{\mathrm{SSFM}} = -29.1086$ dB.

By referring now to Fig. 3.1 and remembering that the SNR is expressed in dB, we can define the error on the SNR as:

$$\mathrm{SNR}^{\mathrm{err}} \triangleq \frac{\mathrm{SNR}}{\widehat{\mathrm{SNR}}} \triangleq 1 + \frac{\sigma^2_{\mathrm{SSFM}}}{\sigma^2_{\mathrm{ASE}} + \sigma^2_{\mathrm{NLI}}}\,. \tag{3.3}$$

Please note that SNR and $\widehat{\mathrm{SNR}}$ in (3.3) are expressed in linear scale, thus in dB the SNR error is the usual

$$\mathrm{SNR}^{\mathrm{err}}_{dB} = \mathrm{SNR}_{dB} - \widehat{\mathrm{SNR}}_{dB}$$

Equation (3.3) is a monotonic decreasing function of ASE noise, as depicted in Fig. 3.3. For increasing values of σ^2_{ASE}, after $\sigma^2_{\mathrm{ASE}} > \sigma^2_{\mathrm{NLI}}$ the SNR error approaches zero dB. The SNR error in dB is thus bounded between:

$$0 \leq \left(\frac{\mathrm{SNR}}{\widehat{\mathrm{SNR}}}\right)_{\mathrm{dB}} \leq \left(1 + \frac{\sigma^2_{\mathrm{SSFM}}}{\sigma^2_{\mathrm{NLI}}}\right)_{\mathrm{dB}}\,. \tag{3.4}$$

Unfortunately, we cannot control the error, hence a good simulation should bound the worst error case, i.e., the upper bound in (3.4). Such a bound is reached in absence of ASE noise.

Moreover, by considering the perturbative assumption we can neglect the SSFM error due to higher order contributions to the nonlinearity. This way σ^2_{SSFM} can be treated as σ^2_{NLI}, hence it is expected to scale with the cube of the power, as we showed in Section 2.1.4 for σ^2_{NLI}. We thus find convenient to work with normalized variances:

$$a_{\mathrm{NL}} \triangleq \frac{\sigma^2_{\mathrm{NLI}}}{P^3}\,, \qquad\qquad a_{\mathrm{SSFM}} \triangleq \frac{\sigma^2_{\mathrm{SSFM}}}{P^3} \tag{3.5}$$

with a_{NL} and a_{SSFM} the unit-power variances of the NLI [52] and the SSFM error, respectively. It has been proved that a_{NL} is constant with the power P as long as the perturbative assumption holds. By the same arguments it can be claimed that even a_{SSFM} is power independent. Although this is just an assumption, we will show later that when the NLI is small, i.e., for typical values of SNR, the perturbative assumption on NLI grants a_{SSFM} to be constant with the input signal power.

Focusing thus on the SNR error upper bound in (3.4) and considering (3.5), under the perturbative assumption the SNR error can thus be expressed as:

$$\frac{\mathrm{SNR}}{\widehat{\mathrm{SNR}}} \triangleq 1 + \frac{a_{\mathrm{SSFM}}}{a_{\mathrm{NL}}} \tag{3.6}$$

whatever the power of the transmitted signal. In the limit of absence of numerical error, i.e., $a_{\mathrm{SSFM}} = 0$, the SNR error $\frac{\mathrm{SNR}}{\widehat{\mathrm{SNR}}} = 1$, which means that the solution of the NLSE estimated by the SSFM coincide with the true one.

The main conclusion is thus that the SNR error of the simulation in (3.6), under perturbative assumption of e_{SSFM}, is independent of the signal power P. The power-independent parameter $\frac{a_{\mathrm{SSFM}}}{a_{\mathrm{NL}}}$ is thus the reference parameter to setup a SSFM simulation, and it should remain constant when varying all the other simulation parameters to grant a constant error on the SNR of the simulation. However, typical rules in literature do not follow such a rule [7–10].

3.2 SSFM error scaling rules

The accuracy of the SSFM depends on the discretization of the spatial coordinate within the fiber. Since SSFM is a convergent algorithm, tiny steps yield accurate simulations at the expense of computational complexity. It is therefore extremely important to know how the SSFM error scales with the step size, which is the subject of this Section.

3.2.1 The Baker-Campbell-Hausdorff formula

The fundamental formula commonly used to show the scaling properties of the SSFM error is given by the Baker-Campbell-Hausdorff formula (BCH) [57], which provides a series expansion of the product of two non-commutative operators. Applied to the NLSE, the two non-commutative operators are clearly $\mathcal{L}$ and $\mathcal{N}$. For the sake of simplicity, consider the BCH for the asymmetric step SSFM [25]:

$$e^{\mathcal{N}h}e^{\mathcal{L}h} = e^{(\mathcal{L}+\mathcal{N})h+\frac{h^2}{2}[\mathcal{N},\mathcal{L}]+\mathcal{O}\left(h^3\right)} \tag{3.7}$$

with h the step length and $[\,\mathcal{L}\,,\mathcal{N}\,] \triangleq (\mathcal{L}\mathcal{N}-\mathcal{N}\mathcal{L})$ the commutator between the two operators $\mathcal{L}$ and $\mathcal{N}$. The $\mathcal{O}\left(h^3\right)$ indicates that other terms are present in (3.7), but for such terms the scaling with h is faster and thus they are expected to be negligible for short steps. The left-hand side of (3.7) is the SSFM solution of the NLSE (2.42) of the asymmetric step SSFM applied to step h while in the first term in the right-hand side we recognize the formal solution of the NLSE of (2.37) over the step length h, i.e.,

$$A\,(h,t) = e^{(\mathcal{L}+\mathcal{N})h}A\,(0,t)\ .$$

All the remaining terms are thus the numerical error. The dominant error term in (3.7) depends quadratically on the step, thus the local error of the asymmetric step SSFM thus scales as $\mathcal{O}\left(h^2\right)$. Equation (2.7) can be easily proved by expanding each exponential up to second order in h. With similar arguments, it can be proved that the local error of the symmetric splitting $e^{\frac{\mathcal{L}}{2}h}e^{\mathcal{N}h}e^{\frac{\mathcal{L}}{2}h}$ scales with $\mathcal{O}\left(h^3\right)$.

Such scaling rules focus on the local behavior of the SSFM, but what really matters is the global accumulated error. Since the local error scales as $\mathcal{O}\left(h^2\right)$ and $\mathcal{O}\left(h^3\right)$ for the asymmetric and symmetric step SSFM, respectively, in the special case of constant step size common wisdom is that the global error accumulated in the propagation scales as $\mathcal{O}\left(h\right)$ and $\mathcal{O}\left(h^2\right)$ for the asymmetric and symmetric step, respectively [8–10, 12]. For that reason, the asymmetric and symmetric steps SSFM are referred to as first-order and second-order methods, respectively [12]. We will show and comment in the next Section that such claim fail at large step-size.

Other more complex versions of the SSFM exist in the literature [12] that increase the order of the error scaling, but they are not considered in this work due to the high complexity they need to implement in the step.

3.2.2 SSFM error as numerical integration error

A different approach to explain the scaling of the local and global error of the SSFM with the step size is presented now, which connects the SSFM to a numerical integration. Let us focus on the ME in integral form (2.38) over a single step h:

$$U\left(h,t\right)=U\left(0,t\right)+\int_{0}^{h}e^{-\mathcal{L}\xi}\mathcal{N}\left(e^{\mathcal{L}\xi}U\left(\xi,t\right)\right)\mathrm{d}\xi$$

where we remember $U(z,t)=e^{-\mathcal{L}z}A(z,t)$ is the optical field A but in a reference system tracking the linear effects.

The function $e^{\mathcal{L}\xi}$ is an oscillating function. In Section 2.1.2 we have seen that such a function plays a fundamental role in mitigating the FWM processes generating nonlinearity on the signal. Such a function in the SSFM is sampled to a given value ξ_0 in each step h and the nonlinearity of the whole step is weighted by such a value. Not surprisingly thus the error of the SSFM is generated by such an approximation, and manifests as an overestimation of the nonlinearity in the step.

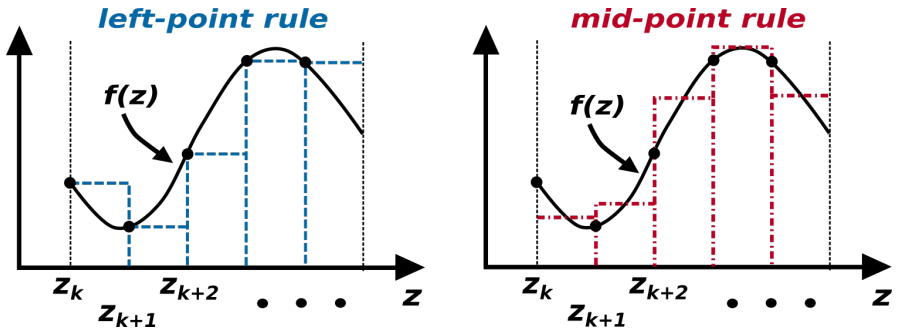

Figure 3.4: Sketch of the left-point rule (left) and mid-point rule (right) for numerical integration. In the SSFM the function $f(z)$ is $e^{-\mathcal{L}\xi}\mathcal{N}\left(e^{\mathcal{L}\xi}U\right)$.

To obtain the asymmetric and symmetric step SSFM such a "weight function" is sampled at $\xi_0 = 0$ and $\xi_0 = \frac{h}{2}$, respectively, and then integrated over the nonlinear accumulation given by $\mathcal{N}(U)$. There is an interesting parallelism with this procedure and two different techniques to numerically integrate a function $f(z)$, namely the left-point rule and the mid-point rule [60]. These two rules are sketched in Fig. 3.4, where $f(z)$ is the function under integration (thus representing $e^{-\mathcal{L}\xi}\mathcal{N}\left(e^{\mathcal{L}\xi}U\right)$ in (2.38) at a given temporal instant) and z_k , z_{k+1} , ... are the simulations steps.

The local error of the numerical quadrature rules comes from the error of the Lagrange interpolation, which can be inferred by the remainder of the Taylor's Theorem applied to the function f. For the left- and mid-point rules, such an error is proportional to the first and second derivative of f in a given coordinate ξ in the step, respectively. Such a coordinate, unfortunately, is unknown, thus the error is a random variable [60].

Following this intuition, the global error of the SSFM simulation e_G can be defined as the sum of the local errors of the above-mentioned integration rules as:

$$e_G = \begin{cases} \sum_{k=1}^{N} \frac{1}{2} h^2 f'(\xi_k) & \text{Asymm. step SSFM} \\ \sum_{k=1}^{N} \frac{1}{24} h^3 f''(\nu_k) & \text{Symm. step SSFM} \end{cases} \tag{3.8}$$

where N is the number of steps and f', f'' are the first and second derivatives of the integrand function $f(z)$ at the unknown spatial coordinates ξ_k and ν_k inside the k^{th} step. Note that ξ_k and ν_k for the two rules are generally different coordinates even in the same step, and depend on the function f and the step length h. For $h \to 0$, i.e., $N \to \infty$, the summations in (3.8) can be approximated to an integral as:

$$\sum_k h f^{(n)}(\xi_k) \sim \int_0^L f^{(n)}(\xi)\,\mathrm{d}\xi = f^{(n-1)}(L) - f^{(n-1)}(0) \tag{3.9}$$

which is independent of h, thus constant over variations of the step length. The new expression of the global error is thus:

$$e_G = \begin{cases} \frac{1}{2}h\mathcal{K} & \text{Asymm. step SSFM} \\ \frac{1}{24}h^2\mathcal{Q} & \text{Symm. step SSFM} \end{cases}, \tag{3.10}$$

with $\mathcal{K} \triangleq f(L) - f(0)$ and $\mathcal{Q} \triangleq f'(L) - f'(0)$ from (3.9). Note that $\mathcal{K}$ and $\mathcal{Q}$ are random variables since both f and f' at coordinate L are unknown constants depending only on the function f. By looking at (3.10) it is clearly visible that in this scenario with small step-size the scaling of the global error with the step length h as $\mathcal{O}(h)$ and $\mathcal{O}\left(h^2\right)$ for the asymmetric and symmetric step SSFM, thus confirming the approach of the previous section.

3.2.3 SSFM error as Monte Carlo integration

The validity of the global error scaling rule (3.10) strongly depends on the assumption (3.9). Such an assumption is true only for $h \to 0$, thus for values of $f^{(n)}(\xi_k)$ strongly correlated between consecutive steps. When the step h become larger, the function f (and consequently its derivatives) is likely to experience many fluctuations within a single step due to FWM, hence the local errors of both the left and mid-point rule are expected to be independent random processes between consecutive steps. Moreover, under ergodic assumptions, the local errors are also identically distributed.

In this case the integration of the function $f(z)$ is the sum of uncorrelated samples, thus it resembles a Monte Carlo integration rather than a numerical quadrature.

In Monte Carlo integration in fact the integral of a generic function $f(z)$ over the integration domain L is approximated by:

$$I = \int_0^L f(\xi)\,\mathrm{d}\xi \;\sim\; \widehat{I} = L\left(\frac{1}{N}\sum_{k=1}^{N} f(x_k)\right),$$

i.e., is given by the arithmetic mean of independent samples of the function $f(x_k)$ over the integration domain, multiplied by the size the integration domain itself. For increasing number of samples, i.e., of simulation steps, the results of the Monte Carlo integration tends to the product of the integration domain L and the mean value of the function $< f >$, which is by definition equal to the integral of $f(z)$ over the domain L.

The error of the Monte Carlo integration is by definition:

$$e \triangleq I - \widehat{I} = I - \frac{L}{N}\sum_{k=1}^{N} f_k\,.$$

If the samples f_k are independent random variables, the variance of the error is:

$$\sigma^2_{e,\mathrm{MC}} = \mathrm{Var}\left[\frac{L}{N}\sum_{k=1}^{N} f_k\right] = \frac{L^2}{N}\sigma_k^2 = Lh\sigma_k^2$$

where $h = \frac{L}{N}$. Hence, contrary to what observed in the previous Section, in this scenario the variance of the SSFM error is expected to scale linearly with the step length h.

3.3 SSFM with variable step-size

In all the previous sections we considered for simplicity the step-size of the SSFM algorithm constant along the propagation. However, previous works

[7–11] showed that for variable step size along distance the performance of the method improves in terms of computational effort while keeping the same accuracy. Two popular rules to update the step-size along propagation will be reported in this section, namely the *nonlinear phase criterion* (NLP) [8] and the *constant local error rule* (CLE) [9, 10]. We propose a different analysis of these rules, based on the separation between the setup of the first step of the simulation and the step-updating rule, i.e., the criterion used to enlarge the step length along the propagation. Finally, we will propose a new criterion to set the first step of the propagation to target a given numerical error on the SNR of the simulation.

3.3.1 first step and step-updating rule

Typically, the step length depends on the physical parameters of the signal and the fiber [6–10]. Since the need for the SSFM is motivated by the presence of the Kerr nonlinearity, the step length should follow the variation of such nonlinearity along distance to keep the error on the step under control. The dependence of the step length by such parameters is usually expressed explicitly for each step along distance [6, 8–10]. However, many parameters are usually considered constant along distance, e.g. the fiber dispersion and the span length in point-to-point transmissions. Thus, by considering an iterative step updating rule, i.e., where each step depends on the previous ones, the dependence of the step-length by system parameters can be explicitly stated just in the first step and "transferred" to the other steps of the simulation by the iterative procedure.

The idea here is thus to separate the analysis of each rule in two sub-problems:

- the setup of the first step of the simulation, and
- the actual step-updating rule, i.e., the scaling of the step-size along the propagation distance.

The main intuition behind this approach is that once the first step is set with a certain criterion, the step-updating rule will keep the error under control along

the propagation. An important consequence of such an intuition is that the error of the whole simulation can be controlled just by wisely setting the first step once the step-updating rule is chosen.

3.3.2 The nonlinear phase criterion

The nonlinear phase criterion has been deeply analyzed in [8]. It is based on the physical intuition that the local error induced by the SSFM is proportional to the strength of the nonlinearity inside the step. Such a nonlinearity manifests typically as a SPM, i.e., a phase rotation depending on the instantaneous power of the signal as reported in (3.5). The NLP bounds the maximum amount of nonlinear phase shift $\Delta\phi$ tolerable inside a step as a way to bound the error of the SSFM. Since the power of the signal decreases along distance exponentially due to the fiber attenuation, such a rule enlarges the step accordingly.

In the framework of the ME, $\Delta\phi$ can be expressed as:

$$\Delta\phi \triangleq \frac{8}{9}\gamma P_{\max} L_{eff}(h)$$

where $P_{\max} = \max_t (P(t))$ is the peak power of the signal at the beginning of the step and $L_{\text{eff}}(h)$ is the effective length of the step h, defined as in (2.9), which takes into account the decrement of the signal power along the step due to the fiber attenuation α. Such a rule thus stretches the step length along the propagation in order to compensate for the power loss of the signal, as sketched in Fig. 3.5. By inverting the above formula and exploiting (2.9), we obtain the step-size at the generic coordinate z_k as:

$$h_k = \begin{cases} \frac{1}{\alpha}\log\left(\frac{\gamma P_k}{\gamma P_k - \alpha\Delta\phi}\right) & \gamma P_k > \alpha\Delta\phi \\ L - z_k & \text{else} \end{cases} \tag{3.11}$$

where P_k is the peak power of the signal at the input of the step h_k. The term $L - z_k$ in (3.11) limits the step length to reach the end of the fiber in a single step when power is too low. It physically means that for $\gamma P_k < \alpha\Delta\phi$ the accumulated nonlinearity from z_k to the end of the fiber is not enough to

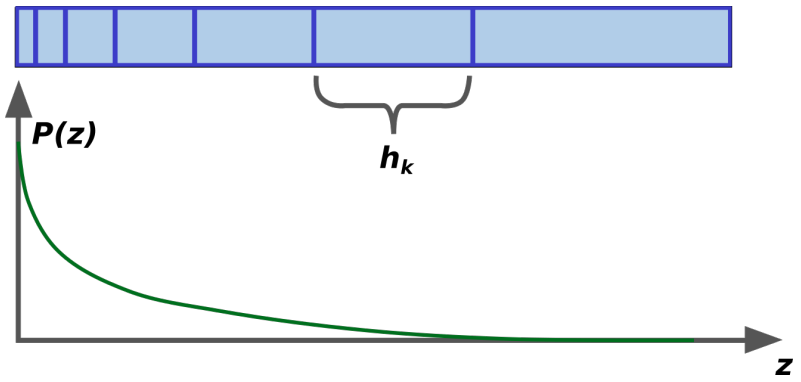

Figure 3.5: Sketch of the nonlinear phase criterion step-size rule. The step-size of the simulation is stretched along distance to compensate for the fiber loss.

reach the upper bound $\Delta\phi$, thus the rest of the fiber can be done in one single step.

The first step of the simulation can be inferred from (3.11) by simply substituting P_k with P_1, i.e., the peak power of the input signal, as:

$$h_1 = \begin{cases} \frac{1}{\alpha}\log\left(\frac{\gamma P_1}{\gamma P_1 - \alpha\Delta\phi}\right) & \gamma P_k > \alpha\Delta\phi \\ L & \text{else} \end{cases} \tag{3.12}$$

The step-updating rule can instead be inferred by matching $\Delta\phi$, which is a constant parameter through the propagation, for two consecutive steps h_k and h_{k+1}, obtaining:

$$L_{\text{eff}}\left(h_{k+1}\right) = \frac{P_k}{P_{k+1}} L_{\text{eff}}\left(h_k\right) .$$

It is worth noting that P_k and P_{k+1} are different because of both attenuation and GVD. However, if we neglect the peak power fluctuation due to GVD, the following rule is obtained:

$$L_{\text{eff}}\left(h_{k+1}\right) = L_{\text{eff}}\left(h_k\right) e^{\alpha h_k} . \tag{3.13}$$

Such a step-updating rule coincides with the step-updating rule of the logarithmic step-size distribution proposed by Bosco *et al.* in [7].

3.3.3 The constant local error rule

The constant local error rule (CLE) has been proposed by Zhang *et al.* for both asymmetric and symmetric step SSFM [9,10] as a step-size rule granting a constant global error for scaling fiber and signal parameters. The CLE has been derived by exploiting the scaling properties of the local error, i.e., the error in a single step, to give a constant error in all the steps of the simulation. The local error in the step h in the asymmetric step SSFM is estimated by evaluating the difference of the signal pulse width passed in the two opposite SSFM schemes $e^{\mathcal{N}h}e^{\mathcal{L}h}$ and $e^{\mathcal{L}h}e^{\mathcal{N}h}$, which coincide with the first error term in (3.2.1). The local error of the symmetric step is similarly derived by considering $e^{\frac{\mathcal{N}}{2}h}e^{\mathcal{L}h}e^{\frac{\mathcal{N}}{2}h}$ and $e^{\frac{\mathcal{L}}{2}h}e^{\mathcal{N}h}e^{\frac{\mathcal{L}}{2}h}$. For a detailed derivation of such a procedure, we remand to [10].

From [9,10] the local error $\Delta\zeta$ is:

$$\Delta\zeta = h_k^2 P_k \gamma \left(D\Delta f \Delta\lambda\right) \qquad \text{Asymm. step SSFM} \tag{3.14}$$

$$\Delta\zeta = h_k^3 P_k \gamma \left(D\Delta f \Delta\lambda\right)^2 \qquad \text{Symm. step SSFM} \tag{3.15}$$

where D is the dispersion parameter. Δf ($\Delta\lambda$) is the signal bandwidth expressed respectively in the frequency (wavelength) domain. Please note that $\Delta\zeta$ is kept identical for all the steps.

The global error of the simulation e_G is the sum of all the local errors. Thus, despite the local error is kept constant along the propagation in the fiber, for variable signal properties it must be scaled together with the step length in order to obtain a constant global error. Hence, a trustworthy simulation cannot be set by fixing $\Delta\zeta$ once and forever. By exploiting the scaling law of the local error in (3.14), (3.15), as well as the consequent variation of the number of step of the simulation, a formula for the first step of the simulation for a fixed global error can be derived.

First, as in [10], we suppose that the global error e_{G} is proportional to the product between $\Delta\zeta$ and the number of steps of the simulation, i.e.:

$$e_{\mathrm{G}} = \Delta\zeta N_{\mathrm{steps}}\,. \tag{3.16}$$

We will show later that such an assumption hold only for small simulation steps. Since the step is variable $N_{\text{steps}} \neq L/h_1$ with L the fiber length. However, we assume that all the steps are small enough such that by shrinking the first step h_1 by a factor k, the resulting number of steps will consequently increase of the same factor k. By such an assumption the number of steps can be defined as:

$$N_{\text{steps}} \triangleq \frac{L}{\delta h_1} \tag{3.17}$$

with L the fiber length and $\delta < 1$ an almost constant coefficient accounting for the variable step along the propagation.

Since the local error $\Delta\zeta$ is supposed identical for each step, by substituting (3.14) for the first step and (3.17) in (3.16) we have:

$$e_G = h_1^q P_1 \gamma \left(DB_{\text{WDM}}\Delta\lambda\right)^{q-1} \frac{L}{\delta h_1}\,. \tag{3.18}$$

where we expressed the signal bandwidth Δf as B_{WDM}. The constant $q = 2$ or 3 depending on the step type, i.e., the asymmetric or symmetric step SSFM, respectively. Expressing $D = -\frac{2\pi c}{\lambda_0^2}\beta_2$ and $\Delta\lambda = \frac{2\pi c}{\lambda_0^2}B_{\text{WDM}}$ and by rearranging (3.18), the following formula for the first step can be inferred:

$$h_1 = \frac{\Psi_G}{(\gamma P_1)^{\frac{1}{q-1}} \left|\beta_2\right| B_{\text{WDM}}^2} \tag{3.19}$$

where Ψ_G is a constant proportional to the global error, i.e., $\Psi_G \propto e_G$. It is worth noting that P_1 is the peak power of the optical signal, thus it depends on the signal bandwidth B_{WDM}. For instance, in a WDM comb of evenly spaced channels with equal power, P_1 scales almost linearly with B_{WDM}.

The step updating rule can be inferred instead by matching the constant $\Delta\zeta$ between two consecutive steps h_k and h_{k+1}, i.e:

$$h_{1k}^q P_k \gamma \left(DB_{\text{WDM}}\Delta\lambda\right)^{q-1} = h_{k+1}^q P_{k+1} \gamma \left(DB_{\text{WDM}}\Delta\lambda\right)^{q-1}$$

Considering the dispersion D and the signal bandwidth B_{WDM} constant during the propagation, the terms inside brackets can be simplified in the equation,

obtaining the following step updating rule for the CLE:

$$h_{k+1} = h_k \left(\frac{P_k}{P_{k+1}} \right)^{\frac{1}{q}} \simeq h_k e^{\frac{\alpha}{q} h_k} \tag{3.20}$$

where the last approximation ignores the peak power fluctuations due to GVD in the step.

It is worth noting that both NLP and CLE stretch the step exponentially along distance, but with a different rate, as depicted in Fig. 3.6.

A small consideration can be done about the the impact of the approximation $\frac{P_k}{P_{k+1}} \simeq e^{\frac{\alpha}{m} h_k}$ for both NLP and CLE, with $m = 1, 2, 3$. The reason of taking into account the peak power of the signal in the local error formulas is to bound the maximum numerical error in time in each step, which is intuitively located where the signal amplitude is maximized. The growth of the step length is thus irregular along distance. However, by using the approximation $e^{\frac{\alpha}{m} h_k}$ we do not expect a much different growth along distance since the peak power is accounted in the first step.

This intuition is confirmed in Fig. 3.6, which shows an example of the scaling of the step-size for the NLP and the CLE versus the propagated distance. The solid curve is the NLP under the approximation $\frac{P_k}{P_{k+1}} \simeq e^{\alpha h k}$ for a starting step h_1 equal to the original NLP. Although this is just an example, we found the same excellent match in many other setups.

It is worth noting that NLP does not make any difference between asymmetric or symmetric step SSFM, contrary to CLE. This comes from the fact that the NLP is a criterion based on the physical properties of the signal in the nonlinear step only, thus it is not aware of the linear/nonlinear interactions within a step. Contrarily, the CLE exploits somehow such interactions by considering the scaling rules of the local error versus the step length.

3.3.4 The constant local error rule as a numerical integration

Since SSFM can be considered as a numerical integration (see Section 3.2.2), the same step updating rule as (3.20) can be inferred by exploiting the scaling

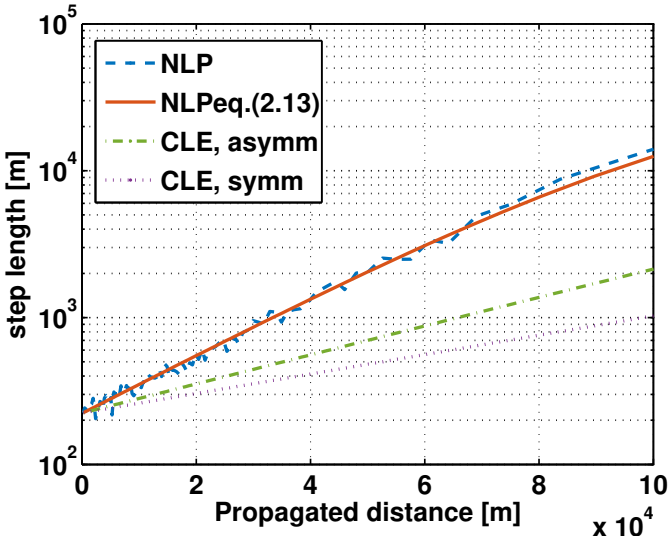

Figure 3.6: example of variable step-updating values of the considered step rules for a given case of study. y-axis on logarithmic scale to highlight the exponential growth of the step along distance. NLP eq. (3.13), is the NLP by neglecting the peak power fluctuations along distance.

properties of the numerical error in the numerical integration theory. The left-point rule of numerical integration, i.e., the technique akin the asymmetric step SSFM, approximates the integral of the function $f(x)$ as:

$$\int_a^b f(x)\,\mathrm{d}x \simeq f(a)(b-a)$$

with $h = b - a$ is the step. The error of such an approximation can be inferred by the remainder of the Taylor series of f truncated at the first order as:

$$\Delta\zeta = \frac{1}{2}h^2 f'(\xi_k) \tag{3.21}$$

with h the step size, f' the first derivative of the integrand function f and ξ_k an unknown coordinate inside the step $[z\,,\, z+h]$. In the SSFM scenario, the function f consists in the integrand function of (2.38), which is a two-variable function of space and time $f(z,t)$. By considering such a function as $f(z,t) = e^{-\alpha z} g(z,t)$ in a generic step between $[z_k,\, z_{k+1}]$, the integral to approximate is:

$$\int_{z_k}^{z_{k+1}} e^{-\alpha\xi} g(\xi,t)\,\mathrm{d}\xi\,.$$

A first approximation can be made by considering the step short enough to have in the integral $e^{-\alpha\xi} \simeq e^{-\alpha z_k}$, $\xi \in [z_k,\, z_{k+1}]$. This way, the local error in (3.21) becomes:

$$\Delta\zeta(t) = \frac{1}{2}h_k^2 e^{-\alpha z_k} g'(\xi_k(t), t)$$

where $h_k = z_{k+1} - z_k$ is the step-size and $\xi_k(t)$ is an unknown coordinate inside the step generally different for each t. Due to the dependence by t of ξ_k even the error is time dependent.

Although $\xi_k(t)$ is unknown for each t, by averaging the error over time the following is obtained:

$$\Delta\zeta = \langle\Delta\zeta(t)\rangle = \frac{1}{2}h_k^2 e^{-\alpha z_k}\mathbb{E}\left[g'(\xi_k(t), t)\right]\,.$$

where the substitution of time averaging with ensemble averaging is motivated by ergodic assumptions. If enough GVD has been accumulated we can assume that the random realizations of $g^{'}$ are spatially independent, thus by matching the error of two consecutive steps we get:

$$h_{k+1}^2 = h_k^2 e^{-\alpha(z_k - z_{k+1})} = h_k e^{\frac{\alpha}{2} h_k}$$

which is exactly (3.14) for the asymmetric step SSFM. The same procedure can be applied to the mid-point rule, whose local error is $\frac{1}{24} h^3 f''(\xi)$, obtaining instead the symmetric step SSFM.

3.4 A new proposal to set the first step length

The two step-size rules described above can be used to set up the SSFM simulations for any variable parameter. The dependence of the SSFM numerical setup from these parameters has to be carefully taken into account to avoid inaccurate results, as well as over-accurate simulations, thus wasting computational effort. It has been shown in [6, 7] and demonstrated in the previous sections that the SSFM overestimates the nonlinearity inside a step.

In fact, referring to Section 3.2.2, it has been shown that the numerical error of the SSFM comes from the approximation to a constant value of the oscillating linear function $e^{\mathcal{L}\xi} \equiv e^{j\Delta\beta\xi}$. Such a function, multiplied with the attenuation $e^{-\alpha\xi}$, weights the accumulation of the FWM in the step as showed in (2.21). The angular frequency of such a function is the phase matching coefficient $\Delta\beta$ (2.22), that we report here for clarity:

$$\Delta\beta = (2\pi)^2 \beta_2 (f - f_1)(f - f_2) \ .$$

The oscillations of $e^{j\Delta\beta\xi}$ thus depend linearly on the GVD of the fiber β_2 as well as the spacing $(f - f_1)(f - f_2)$ of the frequencies involved in the FWM process. The idea is thus to limit the oscillations of the function $e^{j\Delta\beta\xi}$ in the step h by limiting the step length. The maximum frequency of the oscillation

happens between the two side-frequencies of the WDM signal, i.e.,

$$\max_{f,f_1,f_2} \left[(f - f_1)(f - f_2) \right] = B_{\text{WDM}}^2$$

with B_{WDM} the WDM signal bandwidth. According to the idea that the error along propagation can be kept under control by the chosen step-updating rule such as (3.13) and (3.20), a new proposal for the choice of the first step of the simulation is thus:

$$h_1 = \frac{\Phi_{\text{FWM}}}{|\beta_2| \left(2\pi B_{\text{WDM}}\right)^2} \tag{3.22}$$

where Φ_{FWM} is the maximum phase shift of the complex exponential function $e^{j\Delta\beta z}$ allowed by the user over the first step length h_1. Such a phase is kept constant over variations of system parameters. It is worth noting that such a choice shrinks the first step linearly with the GVD parameter of the fiber as well as quadratically with the signal bandwidth.

Such a choice for the first step (3.22) has the following properties:

- It is independent of the signal power P, in agreement with the power independence of the SNR error shown in (3.6);
- it accounts for GVD fluctuations, thus be applied to any type of fiber with different values of GVD;
- it is bandwidth-aware, thus it is reliable even for variations of the WDM signal bandwidth.

Given the first step choice as (3.22), all the successive steps can be set by using any iterative step-updating rule of choice, e.g., the two inferred by the NLP (FWM-NLP) and the CLE (FWM-CLE) in Section 3.3.

3.4.1 Comparison with the NLP and the CLE

In this Section the main differences between the proposal for the first step of the simulation (3.22) and the other two inferred by the NLP and CLE, i.e., (3.12) and (3.19), will be briefly analyzed.

In Section 3.1 we claimed that the error on the SNR of the simulation under perturbative assumption of the nonlinearity is power independent. Such a criterion is respected by the proposal (3.22) by fixing Φ_{FWM} in the first step. The NLP and CLE, are instead power dependent. Equation (3.12), which is the first step inferred by the NLP, depends on the signal power directly from P and indirectly from the number of the channels. In fact, the signal power is also a function of the signal bandwidth. For instance, by considering a WDM comb with evenly spaced channel and each channel at fixed symbol rate and a fixed launch power, the total power P_{tot} of the signal is:

$$P_{\mathrm{tot}} \simeq N_{\mathrm{ch}} P_{\mathrm{ch}}$$

with N_{ch} and P_{ch} respectively the number of channels composing the WDM signal and the channel power. For closely-spaced signals, $N_{\mathrm{ch}} \simeq \frac{B_{\mathrm{WDM}}}{R}$ with R symbol rate, thus

$$P_{\mathrm{tot}} \simeq \frac{B_{\mathrm{WDM}}}{R} P_{\mathrm{ch}} \,. \tag{3.23}$$

It is worth noting that the dependence of the NLP on the signal bandwidth is linear, as visible in (3.23), contrary to the quadratic dependence on B_{WDM} of (3.22). Moreover, since NLP is based only on the physical properties of the nonlinear step, it is completely unaware of the dispersive effect due to GVD.

The first step inferred from the CLE, i.e. equation (3.19), is instead aware of the signal bandwidth and the fiber dispersion since it depends on both B_{WDM}^2 and β_2. However, it is still dependent on the signal power, thus in contrast with the claim of power independence of the SNR error in (3.6). Moreover, although it seems to have the same bandwidth dependence on signal bandwidth as (3.22), i.e., a quadratic scale with it, in reality even the signal power is bandwidth dependent according to (3.23). Therefore, the final scaling with bandwidth differs from a quadratic scaling.

3.5 Numerical results

In this section, the previous step-size setup are numerically analyzed. In particular, we prove the power independence of the numerical error of the SNR under perturbative assumptions, as well as the scaling of the error with other system parameters such as the dispersion of the fiber and the signal bandwidth.

3.5.1 Simulation setup

The accuracy of the SSFM has been investigated in a wide range of scenarios. The tested signal was composed of a WDM comb of PDM-16 quadrature amplitude modulation (QAM) channels spaced by $\Delta f = 50$ GHz. Each channel had symbol rate $R = 49$ GBaud, with root-raised cosine supporting pulses with roll-off 0.01. The number of channels was varied between 1 and 101 to check the SSFM accuracy at different WDM bandwidths. The analysis has thus reached a maximum simulated bandwidth of $B = 5.05$ THz, i.e., about the whole C-band. The optical link used fibers of length 100 km, attenuation 0.2 dB/km, nonlinear coefficient $\gamma = 1.3$ $\left[\frac{1}{\text{mW}\cdot\text{km}}\right]$ and variable dispersion depending on the setup under investigation. The fiber dispersion was varied between $17/8 = 2.125$ and 17 ps/nm/km with a third order dispersion of 0.057 ps/nm^2/km. All dispersive effects were always fully recovered at the receiver side. We overlooked PMD [46]. The number of spans was 20 unless few cases where we varied them. The fiber loss was perfectly recovered by noiseless optical amplifiers span-by-span. This choice represent a worst case for SSFM accuracy as stated in (3.4). The signal propagation has been simulated in MATLAB with the Optilux toolbox by applying the SSFM to the Manakov equation. The transmitted sequence length was set longer than the maximum walk-off length over the whole link between the two border channels of the WDM comb, with a minimum value of 4096 symbols. Each symbol was discretized with enough samples to correctly reproduce at least the bandwidth enlargement of the first-order FWM without aliasing. A more detailed explanation of the choice of these two parameters can be found in Appendix 3.6.5.

For efficient FFTs, both parameters have been rounded off to the closest larger integer with factorization containing only powers of 2,3,5.

The receiver detected the central channel of the WDM comb at $\lambda = 1550$ nm. In order to minimize the impact of the detection on the SSFM accuracy, we considered a basic receiver optimized for linear transmission, hence based on a cascade of a matched filter followed by a one tap least squares butterfly equalizer able to recover the average polarization/phase mismatch.

We define the SSFM error variance σ^2_{SSFM} at the end of the link as:

$$\sigma^2_{\mathrm{SSFM}} \triangleq \mathrm{var}\left[\widehat{A} - A_{\mathrm{acc}}\right]$$

with $\widehat{A}$ the current SSFM estimation under investigation and A_{acc} our most accurate SSFM solution. Since it was impossible to know the exact solution of the NLSE, we found A_{acc} by running simulations at increasing accuracy until observing saturation of $\widehat{\mathrm{SNR}}$. Saturation was declared to occur when the difference of $\widehat{\mathrm{SNR}}$ between two consecutive runs was less than 0.0005 dB. From A_{acc} we also estimated the unit-power NLI variance a_{NL} (3.5) from the constellation clouds of the received samples. a_{NL} is defined as the variance:

$$a_{\mathrm{NL}} = \mathrm{Var}\left[\frac{A_{\mathrm{acc}}e^{-j\varphi} - A_{\mathrm{tx}}}{P\sqrt{P}}\right]$$

with A_{tx} the transmitted signal and φ the average phase rotation induced by the link. The accuracy of the variable step-size rules under analysis was investigated by varying the first step size h_1 over a logarithmic grid, i.e., by reducing h_1 by a factor $\sqrt{2}$ between two consecutive runs.

The new rule to setup the first step of the SSFM simulation proposed in Section 3.4, based on fixing Φ_{FWM}, is applied to the two step-updating rules reported in Section 3.2, i.e., (3.13) and (3.20), and compared to the same step-updating rules with the first step selected as the original criteria (3.12) and (3.19).

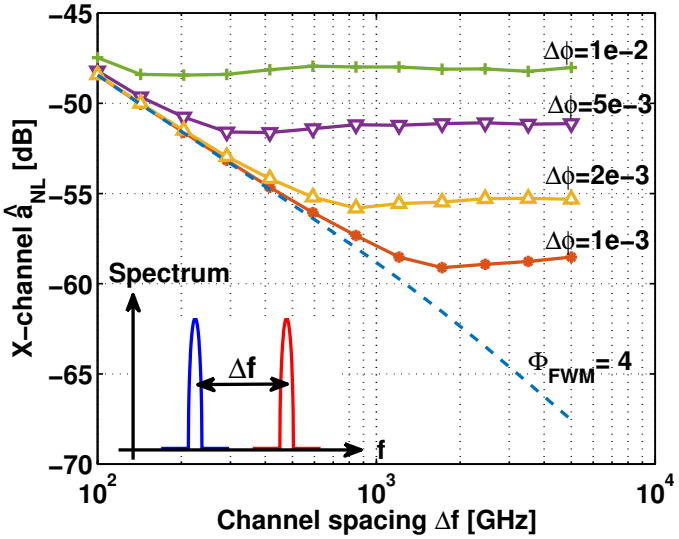

Figure 3.7: Estimated cross-channel $\hat{a}_{\mathrm{NL}}$ (3.5) of a 2-channel simulation at variable channel spacing. Solid lines: original NLP at various values of $\Delta\phi$; dashed-line: accurate SSFM with FWM-NLP at $\Phi_{\mathrm{FWM}} = 4$ rad.

3.5.2 Two-channel propagation: an interesting example

A first evidence of the importance of relating the step size to the signal properties is reported in Fig. 3.7. The figure shows the estimated unit-power NLI variance $\hat{a}_{\mathrm{NL}}$, affected by numerical noise, for a 2-channel WDM at variable channel spacing. The NLI variance refers to cross-channel interference only, being estimated by removing self-channel distortion from the detected signal. Power was 0 dBm while the link was just one SMF span. In this simplified scenario $\hat{a}_{\mathrm{NL}}$ is expected to be inversely proportional to the channel spacing [61]. However, this is not the case for the original NLP (solid curves) which saturates after a given channel spacing, depending on the value $\Delta\phi$. This is a numerical artifact that can be removed by accounting for the signal bandwidth in the step-size selection, e.g., by starting NLP with our FWM-aware first step choice, as we did for the dashed line in Fig. 3.7 with $\Phi_{\mathrm{FWM}} = 4$ rad.

We proceed now to show the accuracy dependence on signal and fiber parameters of fully loaded WDM signals.

3.5.3 SSFM error vs power

We started investigating the dependence of the SSFM accuracy with the signal power. The analysis of the system performance versus the signal power is extremely important due to the limitation to the achievable information rate with power due to Kerr nonlinearity [31]. Such a curve is also fundamental in determining the nonlinear threshold P_{NLT} [52,53], i.e., the power corresponding to the maximum SNR reachable by the system, thus the best power for transmissions.

Following equation (3.6), the fundamental parameter to obtain a constant error on the SNR is the ratio $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$, with a_{SSFM} and a_{NL} respectively the unit-power variance of the SSFM error and the NLI, defined in (3.5). Although the SNR error is the key performance estimator, it may be more useful to focus on $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$ which is more sensitive to small variations of SSFM noise. The accuracy of the SSFM will be thus expressed in terms of such a ratio. To obtain a feeling with this new accuracy parameter, Fig. 3.8 depicts the relation between the SNR error $\mathrm{SNR}/\widehat{\mathrm{SNR}}$ and the ratio $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$, both expressed in dB. In any case, whenever possible the SNR error value will be reported together with $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$.

The dependence of the ratio $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$ with power is reported in Figs. 3.9-3.10 for a 20×100 km SMF link ($D = 17$ ps/nm/km) simulated with the symmetric step SSFM. The signal was composed of 27 channels spaced as reported in Section 3.5.1. By equation (3.6) small values of $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$ indicate an accurate simulation, as visible in the plots by looking at the right vertical axis reporting the corresponding SNR relative error.

To explore different accuracy regions, we tested three different values of the first step per span h_1. The value of h_1 has been set at $P = 0$ dBm, and then it has been changed with power accordingly to the rule under investigation. The three reference values for the first step were 20, 40 and 400 m for respectively

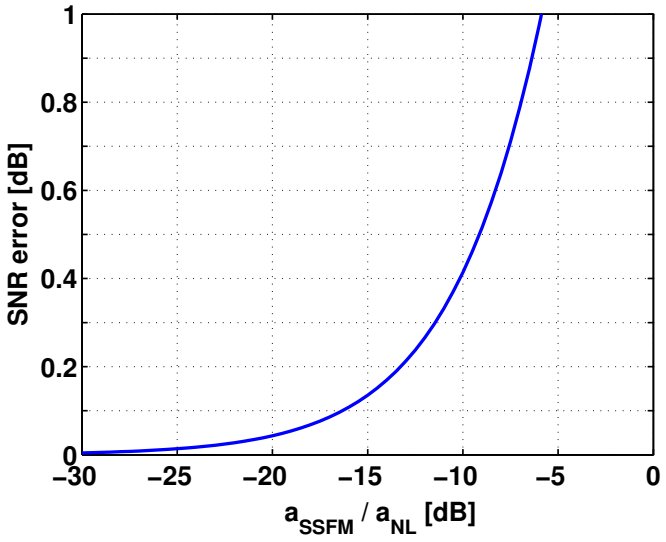

Figure 3.8: Relation between the SNR error and the ratio $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$ as defined in (3.6).

Fig. 3.9top, bottom and Fig. 3.10. For different powers we either chose h_1 according to the original indications of the NLP, eq. (3.12), the CLE, eq (3.19), or by keeping it constant as suggested by the FWM-aware proposal (3.22) based on Φ_{FWM}. The proposal has been either applied to the step-updating rule inferred by the NLP, i.e., the FWM-NLP, or the one inferred by the CLE, i.e., the FWM-CLE, reported respectively in equations (3.13) and (3.20).

From the figures one can note that choosing the first step h_1 with the proposed rule (3.22) yields an almost constant ratio $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$, thus a constant SNR error, whatever the power is used for both the step-updating rules considered, as discussed in Section 3.1. Comparing the three plots it can be noted that the power independence of the error holds for all the h_1 considered.

The numerical error of the original NLP and the CLE methods instead is not constant, decreasing for increasing power. The reasons for such a behavior can be explained by looking at Fig. 3.11 where it is reported the first step

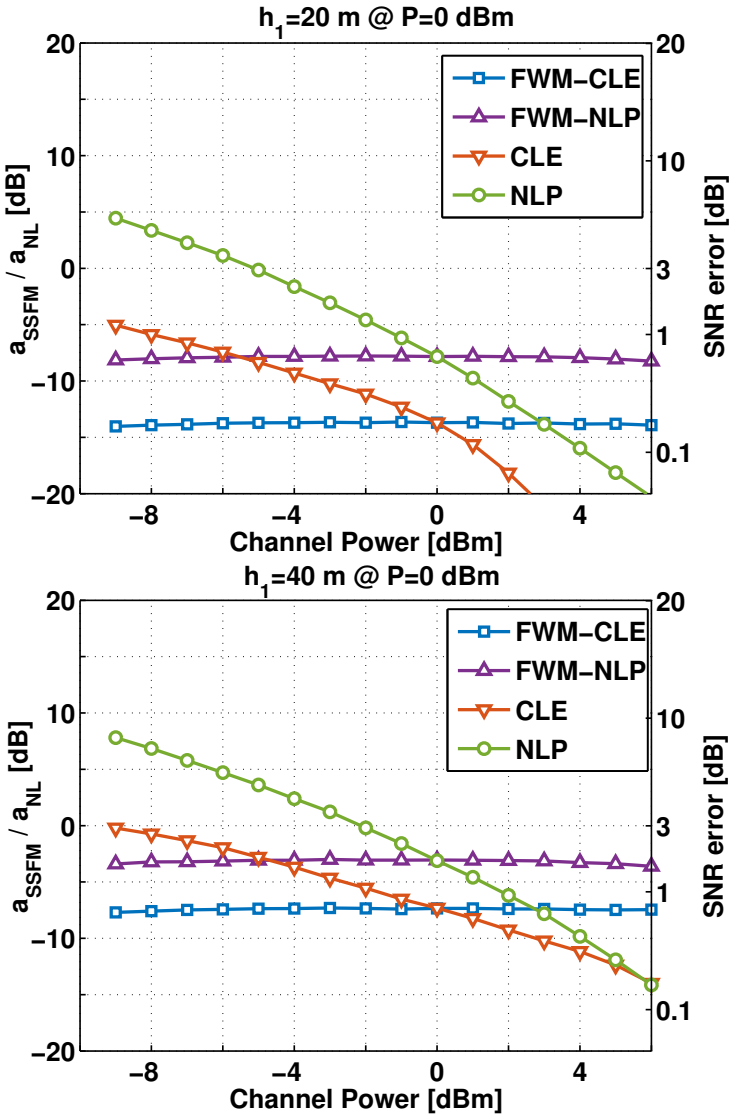

Figure 3.9: SSFM accuracy $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$ vs channel power. The right vertical axis also reports the corresponding SNR relative error. We compare the original nonlinear phase criterion (NLP) and the original constant local error method (CLE) with the FWM-aware extensions FWM-CLE and FWM-NLP. The first step was set at $P = 0$ dBm to (top) $h_1 = 20$ m, and (bottom) $h_1 = 40$ m and scaled by varying signal power according to the method under analysis. 27 channels WDM signal, 20×100 km SMF link. SSFM with symmetric step.

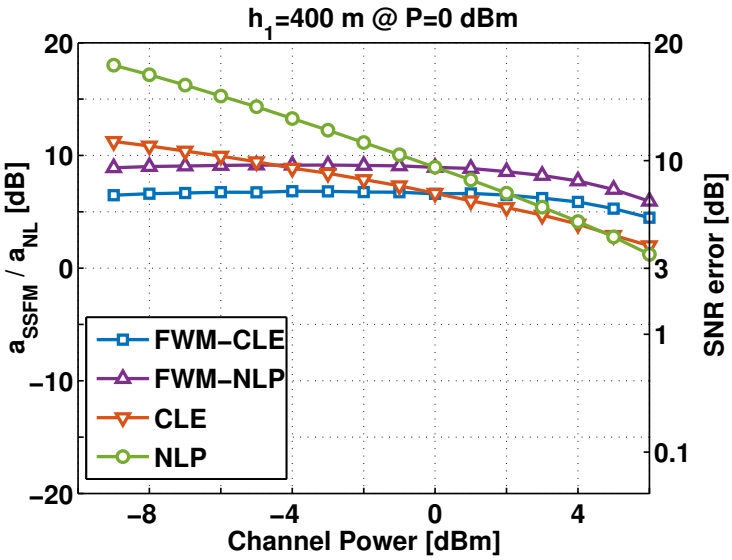

Figure 3.10: SSFM accuracy $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$ vs channel power. The right vertical axis also reports the corresponding SNR relative error. We compare the original nonlinear phase criterion (NLP) and the original constant local error method (CLE) with the FWM-aware extensions FWM-CLE and FWM-NLP. The first step was set at $P = 0$ dBm to (top) $h_1 = 400$ m and scaled by varying signal power according to the method under analysis. 27 channels WDM signal, 20×100 km SMF link. SSFM with symmetric step.

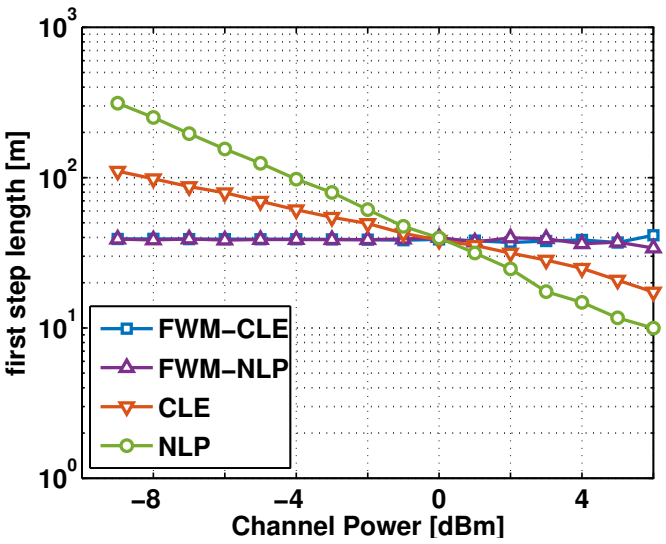

Figure 3.11: First step h_1 of the simulation versus signal power. 27 channels WDM signal, 20×100 km SMF link. SSFM with symmetric step.

length versus the channel power for a first step $h_1 = 40$ at $P = 0$ dBm. Contrary to the FWM-NLP and FWM-CLE rules, where the first step remains constant, the original criteria NLP and CLE shorten the step for increasing power accordingly to (3.12) and (3.19), thus wasting computational effort for increasing power.

For big first step-size, e.q. as in Fig. 3.10, even FWM-NLP and FWM-CLE show power-dependent accuracy. To explain such a behavior, we report in Fig. 3.12 the SNR due only to the numerical error, defined as $\mathrm{SNR}_{\mathrm{SSFM}} \triangleq \frac{P}{\sigma^2_{\mathrm{SSFM}}}$. It can be seen that at high power $\mathrm{SNR}_{\mathrm{SSFM}}$ falls to values lower than 8 dB. Under the assumption of additive Gaussian noise, such values correspond to a Q-factor around 2 dB. The perturbative assumption of the numerical noise does not hold anymore, thus resulting in a failure of the proposal first step rule. However, it is worth noting that such a rule is conservative since the error is decreasing for increasing power.

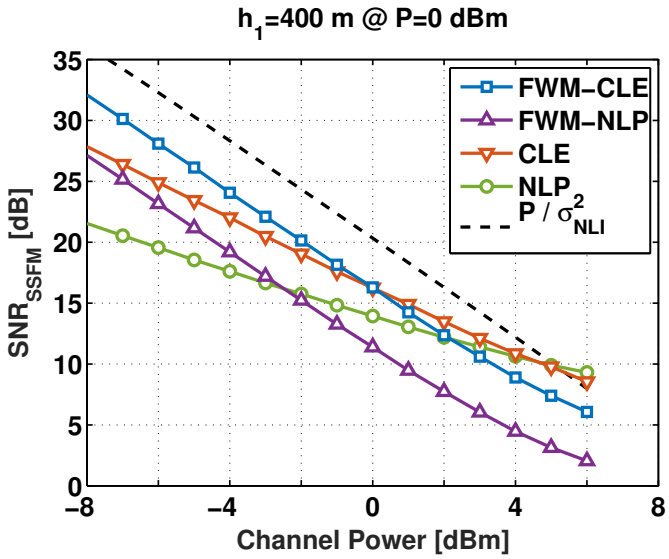

Figure 3.12: $\mathrm{SNR}_{\mathrm{SSFM}}$ versus channel power for a first step of the simulation $h_1 = 400$ m @$P = 0$ dBm. Setup and step-size criteria identical to Fig. 3.10. The additional dashed curve $P\,/\,\sigma^2_{\mathrm{NLI}}$ is the SNR obtained with the most accurate simulation, used here as reference.

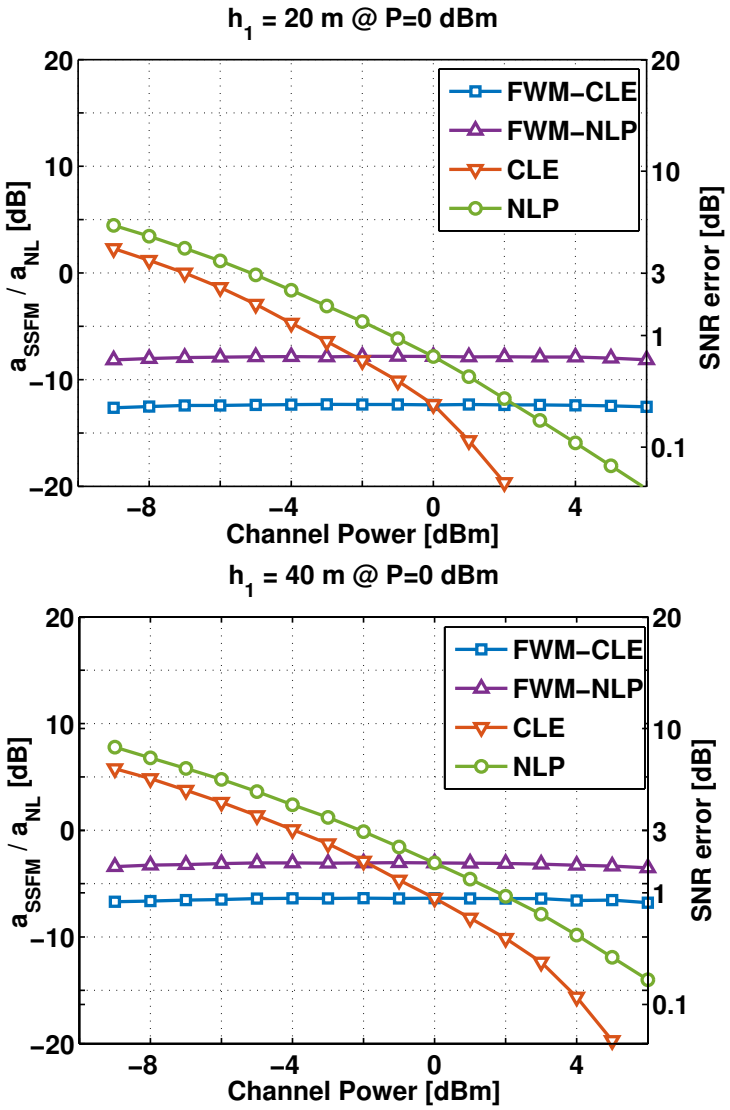

Figure 3.13: SSFM accuracy $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$ vs channel power. The right vertical axis also reports the corresponding SNR relative error. We compare the original nonlinear phase criterion (NLP) and the original constant local error method (CLE) with the FWM-aware extensions FWM-CLE and FWM-NLP. The first step was set at $P = 0$ dBm to (top) $h_1 = 400$ m and scaled by varying signal power according to the method under analysis. 27 channels WDM signal, 20×100 km SMF link. SSFM with asymmetric step.

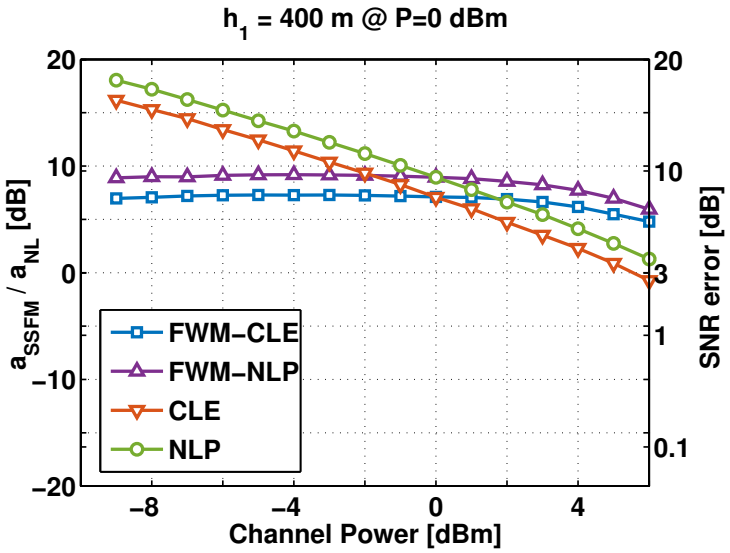

Figure 3.14: SSFM accuracy $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$ vs channel power. The right vertical axis also reports the corresponding SNR relative error. We compare the original nonlinear phase criterion (NLP) and the original constant local error method (CLE) with the FWM-aware extensions FWM-CLE and FWM-NLP. The first step was set at $P = 0$ dBm to (top) $h_1 = 400$ m and scaled by varying signal power according to the method under analysis. 27 channels WDM signal, 20×100 km SMF link. SSFM with asymmetric step.

The same conclusions are still valid for the SSFM with asymmetric step, as visible in Figs. 3.13-3.14. It is worth noting that with the asymmetric step the CLE scales the first step with power in the same way as the NLP, i.e., as $\frac{1}{P}$ according to (3.19). However, please remember that the NLP and CLE step-updating rules are different, thus they result in a different accuracy for the same first step length.

3.5.4 SSFM error vs fiber dispersion

A key parameter influencing the accuracy of the SSFM is the fiber dispersion. In order to evaluate how the SSFM error scales with such a parameter the ratio $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$ at variable Φ_{FWM} has been analyzed. We investigated here a 101 channel WDM signal at fixed launch power per channel $P = 0$ dBm. We first report the ratio $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$ versus fiber dispersion in Fig. 3.15 for both the symmetric step SSFM (top) and the asymmetric step SSFM (bottom). The two step-updating rules considered here are FWM-NLP and FWM-CLE. Note that decreasing Φ_{FWM} means decreasing the step size, i.e., improving the accuracy of the simulation. Each curve refers to a different value of fiber dispersion tested in the range $D = [17/8, 17/4, 17]$ ps/nm/km. The symmetric step SSFM shown in Fig. 3.15(top) yields overlapping curves for both step-size updating rules by varying dispersion, thus they have the same error. Such a result confirms that fixing a value of Φ_{FWM} ensures the same accuracy at all considered dispersions. It is worth noting that fixing Φ_{FWM} consists in varying the step-size inversely with the dispersion, thus increasing the number of simulation steps (i.e., the complexity) linearly with it. The proposed FWM-aware choice to set the first step (3.22) is thus able to adapt to the fiber dispersion. The same results has been obtained for the asymmetric step SSFM, as showed in Fig. 3.15(bottom).

An interesting aspect to highlight is that the choice of the step-updating rule impacts the value of the ratio $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$, i.e., the error of the simulation, but it does not impact the independence of the error with the fiber dispersion. In fact, both step-updating rules almost overlap by varying dispersion, thus confirming the idea proposed in Section 3.2 that the error of the simulation

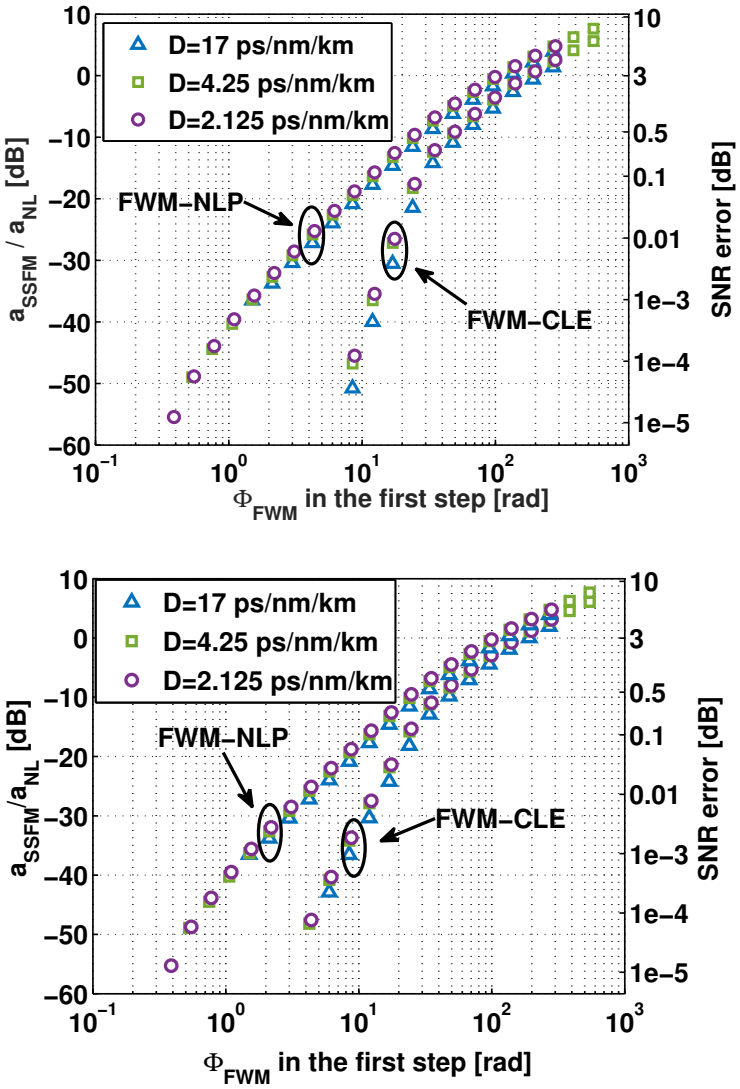

Figure 3.15: (top) $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$ as a function of Φ_{FWM} in the first step computed with symmetric step SSFM. (bottom) Same as top plot but with asymmetric step SSFM. 101 channel WDM signal ($B_{\mathrm{WDM}} = 5.05$ THz) at channel power $P = 0$ dBm. 20×100 km link with variable dispersion. Symmetric SSFM. The corresponding SNR error is shown in the right vertical axis.

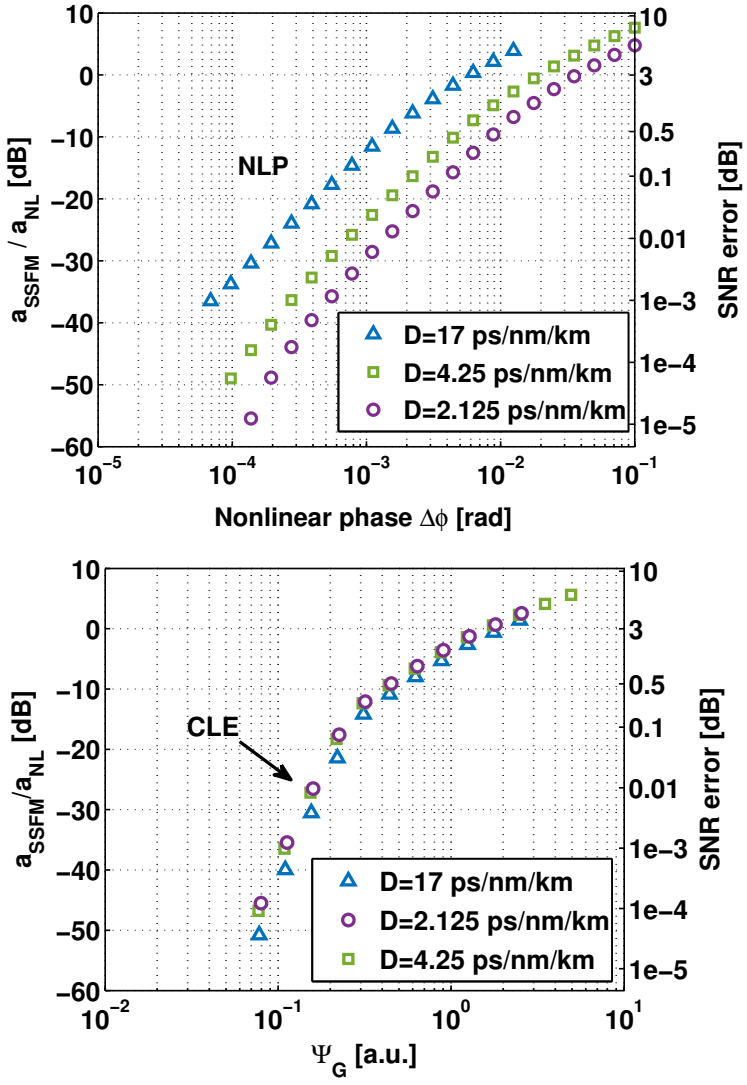

Figure 3.16: (top) $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$ of the SSFM propagation set as the original NLP, thus a function of $\Delta\phi$ (3.12). (bottom) Same as top plot but with CLE, as a function of Ψ_G (3.19). 101 channel WDM signal ($B_{\mathrm{WDM}} = 5.05$ THz) at channel power $P = 0$ dBm. 20×100 km link with variable dispersion. Symmetric step SSFM. The corresponding SNR error is shown in the right vertical axis.

can be controlled only by setting the first step once a given step-updating rule is chosen.

Moreover, it is worth noting that since error of the asymmetric and symmetric step SSFM scale differently with the step length, as reported in Section 3.1, the scaling of the ratio $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$with Φ_{FWM} is expected to be different. In fact, as reported in Section 3.2.1, the SSFM global error scales as $\mathcal{O}(h)$ and $\mathcal{O}\left(h^2\right)$ for the asymmetric and symmetric step, respectively. The error variance is thus expected to scale as $\mathcal{O}\left(h^2\right)$ and $\mathcal{O}\left(h^4\right)$, respectively. Such observation however is not followed for the FWM-NLP in Fig. 3.15(top) and (bottom) as they show the same scaling of $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$ in asymmetric and symmetric step SSFM equal to $\mathcal{O}\left(h^2\right)$. Such a observation will be addressed later in Section 3.6 in a computational analysis of the considered rules.

Figure 3.16 shows the scaling of the error as a function of the two fundamental parameters depicted in (3.12) and (3.19), i.e., the maximum nonlinear phase shift in the step $\Delta\phi$ for the NLP and the constant Ψ_G for the CLE. As expected, the NLP with the first step set by its original parameter shows a variable error with dispersion. Contrarily, the CLE accounts for fiber dispersion, as depicted in (3.19). It can be noted that both (3.19) and (3.22) have the same dependence on β_2, thus not surprisingly even the CLE in its original setup is able to track the variations of fiber dispersion, as visible in Fig. 3.16(bottom). However, a constant SSFM error for variable dispersion does not mean a constant SSFM error for variable GVD, since the scaling properties with the bandwidth are different than what guessed by the CLE. The next Section investigates such a problem in more detail.

3.5.5 SSFM error vs bandwidth

The dependence of the accuracy on the signal bandwidth is another important aspect to explore. Nowadays communications are pushing the WDM signal bandwidths up to the whole C-band and beyond [23, 24], opening new opportunities to increase the achievable information rate. The SSFM simulations should keep up to such a new scenario, thus a study of the SSFM error for

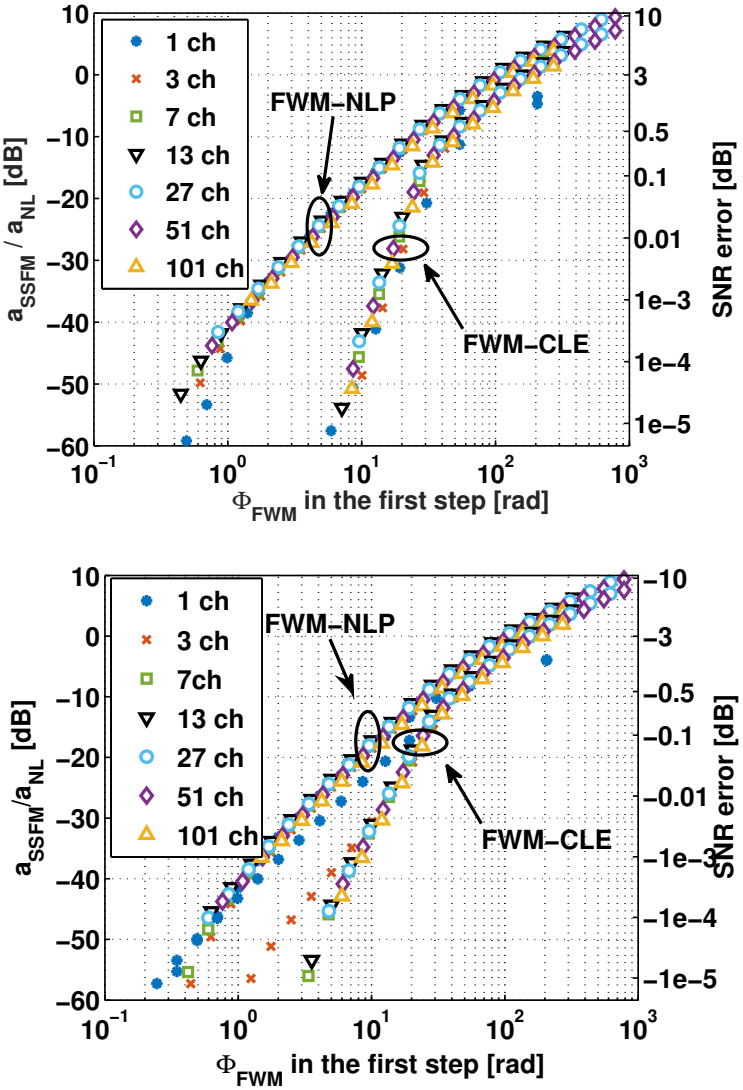

Figure 3.17: $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$ as a function of Φ_{FWM} in the first step for the two different step-updating rules indicated in the graph. Variable number of WDM channels from 1 to 101 ($B_{\mathrm{WDM}} = 5.05$ THz) and channel power $P = 0$ dBm. 20×100 km SMF link. (top) Symmetric step SSFM; (bottom) asymmetric step SSFM. The right axis shows the corresponding SNR error.

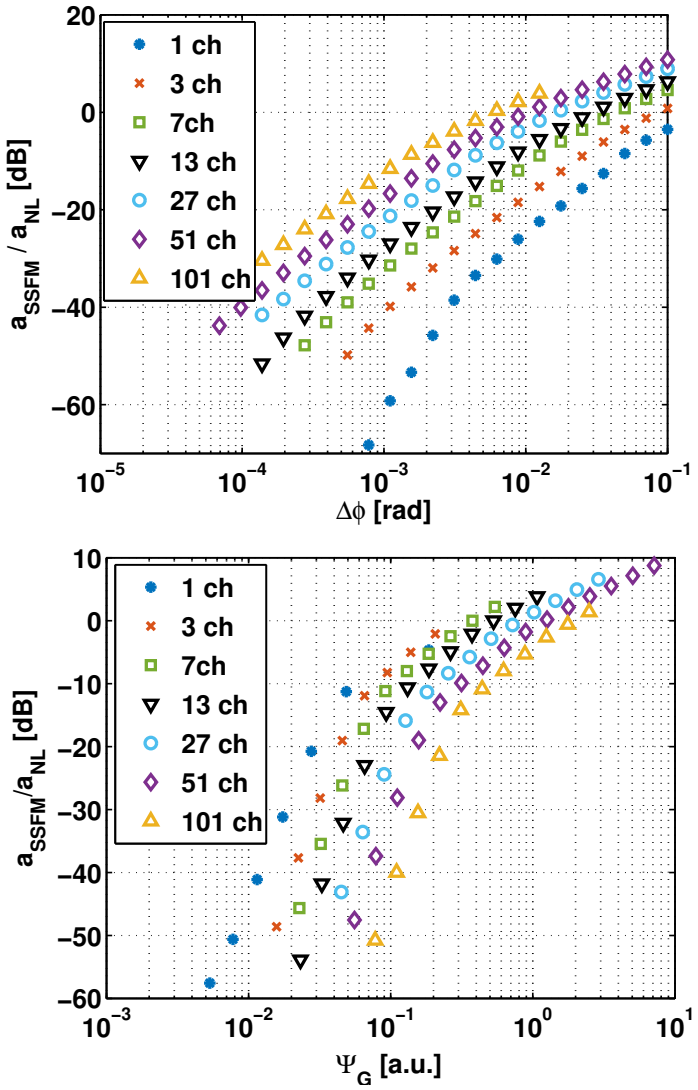

Figure 3.18: $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$ as a function of $\Delta\phi$ (top) and Ψ_G (bottom). Variable number of WDM channels from 1 to 101 ($B_{\mathrm{WDM}} = 5.05$ THz) and channel power $P = 0$ dBm. 20×100 km SMF link. Symmetric step SSFM.

increasing bandwidth of fully loaded WDM signals is mandatory.

Figure 3.17(top) depicts the dependence of the SSFM error with the number of channels of the WDM comb. The SSFM error, expressed in terms of the ratio $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$, is plotted versus the maximum FWM phase shift Φ_{FWM} in the first step for both the FWM-NLP and FWM-CLE for the symmetric step SSFM. Each curve in the plot refers to a WDM signal with a different number of channels, starting from a single-channel transmission and up to 101 channels, i.e., a full C-band WDM system, corresponding to a total signal bandwidth $B_{\mathrm{WDM}} = 5.05$ THz. All the curves with more than 3 channels overlap for both FWM-NLP and FWM-CLE, highlighting the insensitivity of Φ_{FWM} to the total system bandwidth. Figure 3.17(bottom) reports similar curves but for the asymmetric step SSFM, with similar conclusions.

The two original step-size rules, i.e., the NLP and the CLE, used with their own first step criterion, fail to maintain the same error for variable bandwidth. Such a result is depicted in Fig. 3.18 for the symmetric step SSFM and can be explained by a different dependence of the first step size on the signal bandwidth of these two rules with respect to Φ_{FWM}. The NLP in fact depends on the bandwidth only through the dependence of total signal power with the number of channels of the signal. For evenly-spaced WDM signals, such a relation is almost linear with the number of channel. The CLE, as reported in (3.19), accounts for the signal bandwidth by the quadratic relation Δf^2, which in this scenario corresponds to B^2_{WDM}. However, the first step of the CLE scales with power as much as the NLP, hence overall it shows a scaling with B_{WDM} different from B^2_{WDM}. For this reason, the CLE in its original setup is not able to track the error variations due to the increment of the number of channel of the WDM signal. It is worth noting that while the dependence of the CLE from bandwidth is faster than quadratic, resulting thus in a conservative rule for increasing bandwidth, the NLP scales the first step only linearly with B_{WDM}. The NLP is thus less accurate for increasing signal bandwidth, as visible in Fig. 3.18(top) for a fixed value of the parameter $\Delta\phi$.

3.5.6 SSFM error vs distance

The dependence of SSFM accuracy on propagation distance is shown in Fig. 3.19a-b). Here we plotted the ratio $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$ versus the number of spans at a fixed value of FWM phase shift in the first step, chosen equal to $\Phi_{\mathrm{FWM}} = 20$ rad, which yields a practical value of SNR error for the FWM-CLE of almost 0.01 dB (see Fig. 3.17). It can be observed that $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$ in Fig. 3.19 decreases for increasing number of spans at all bandwidths under investigation, reaching saturation after roughly ten spans for a number of channels bigger than 3. The one span case is thus a worst case for accuracy. For very small bandwidths, i.e., for 1 and 3 channel signals, reported in Fig. 3.19b), the decrement of $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$ along spans is even more pronounced.

The reason for such a decrease is that a_{NL} grows faster with distance than a_{SSFM}. To address this claim the accumulation of a_{NL} and a_{SSFM} along the spans of the link is plotted in Fig. 3.20. The accumulation ψ is defined as:

$$\psi \triangleq \frac{a_i\,(n)}{a_i\,(1)}\,,$$

where the unit-power variance $a_i,\ \ i \in \{\mathrm{NL}, \mathrm{SSFM}\}$ (see (3.5)), after n spans is normalized to the same variance after 1 span. Here a_{NL} shows a growth in distance faster than linear for any number of channels considered, with a more pronounced slope in the first spans of the link. Such a growth is well-known in literature, and it is due both to the spatial correlation between the nonlinearity generated at different coordinates of the link [14, 17], and to the dependence of the nonlinearity by the modulation format of the signal, which enhances the dependence on the the higher-order statistics of the data symbols in the first spans [18, 19, 21]. Contrarily, a_{SSFM} shows a slower growth along distance, approaching a linear growth after a few spans for every number of channels greater than 3. Such a behavior depends on the step-updating rule, as visible in Figs. 3.21-3.22. Here the same variation with the number of spans of the ratio $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$ for a given Φ_{FWM} and the single variance accumulation ψ are reported for the FWM-NLP. To have SNR errors comparable with the

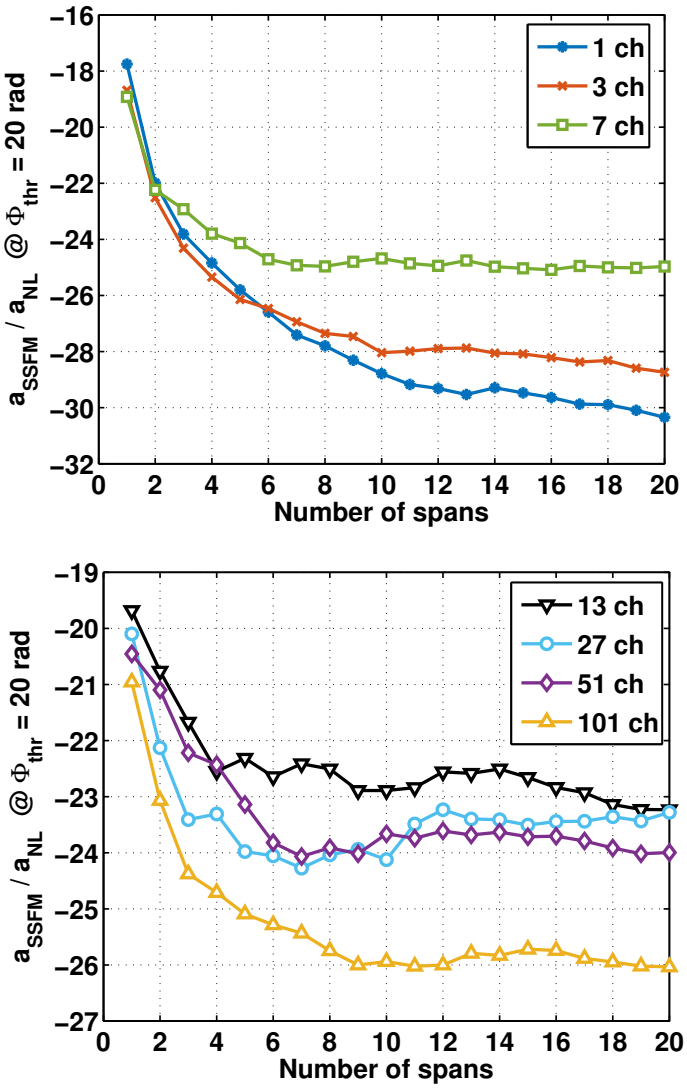

Figure 3.19: $a_{\mathrm{NL}}/a_{\mathrm{SSFM}}$ versus number of spans ($\times$100 km) at fixed $\Phi_{\mathrm{FWM}} = 20$ rad. Number of channels for the WDM signal ranges from 1 to 101. Propagation over SMF. Symmetric SSFM with FWM-CLE.

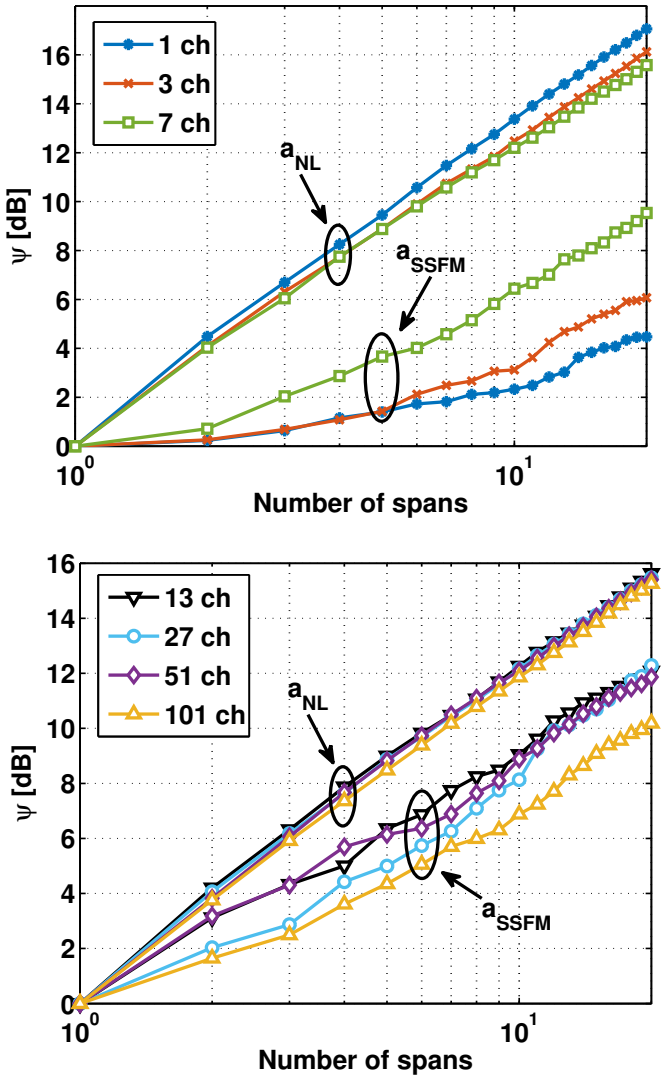

Figure 3.20: Unit-power variance accumulation ψ versus number of spans ($\times$100 km) at fixed $\Phi_{\mathrm{FWM}} = 20$ rad. Channel number range from 1 to 101 between the two plots. Propagation over SMF. Symmetric SSFM with FWM-CLE.

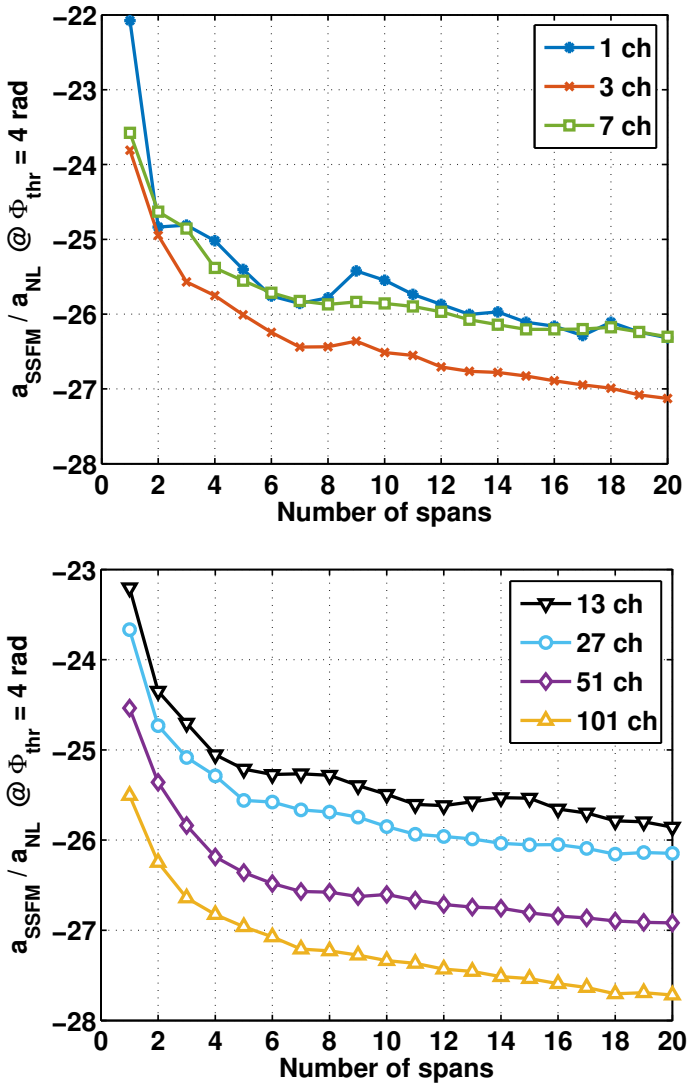

Figure 3.21: $a_{\mathrm{NL}}/a_{\mathrm{SSFM}}$ versus number of spans ($\times$100 km) at fixed $\Phi_{\mathrm{FWM}} = 20$ rad. Number of channels for the WDM signal ranges from 1 to 101. Propagation over SMF. Symmetric SSFM with FWM-CLE.

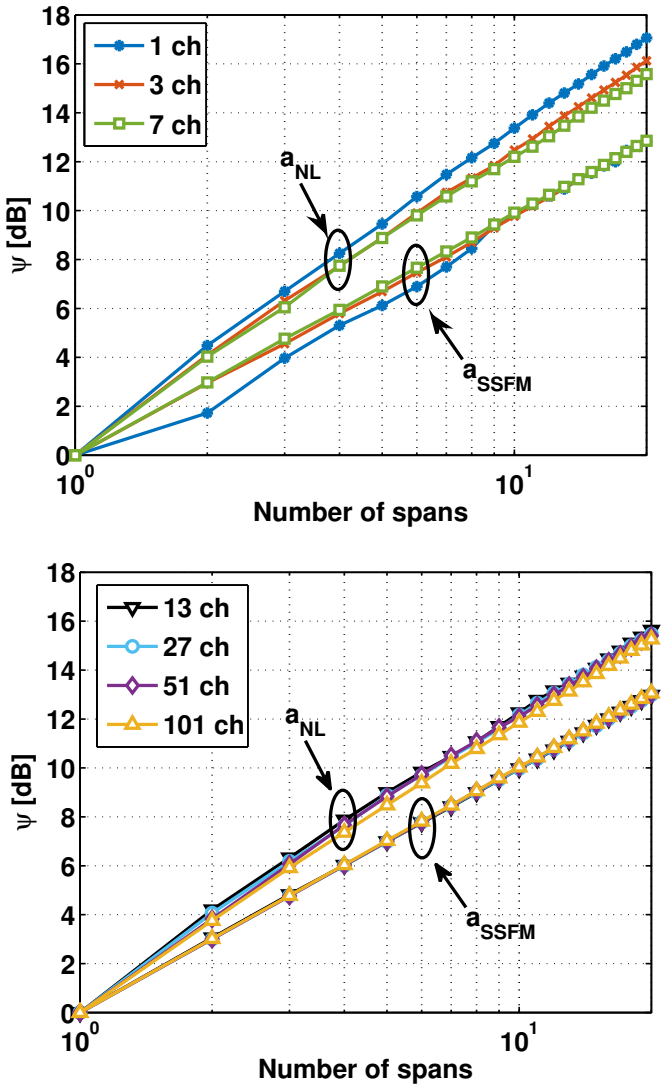

Figure 3.22: Unit-power variance accumulation ψ versus number of spans ($\times$100 km) at fixed $\Phi_{\mathrm{FWM}} = 20$ rad. Channel number range from 1 to 101 between the two plots. Propagation over SMF. Symmetric SSFM with FWM-CLE.

FWM-CLE this time we used $\Phi_{\mathrm{FWM}} = 4$ rad. Even with the FWM-NLP the accuracy grows with the spans. However, by using this rule the accumulation of a_{SSFM} is basically linear for all the spans composing the link, as depicted in Fig. 3.22, while a_{NL} still grows faster than linear. However, such a difference between the two rules is of minor importance. The main conclusion of the above investigation is that the first span is a worst case for SSFM accuracy.

3.6 Computational analysis

In this Section we analyze the step-size rules considered in Section 3.3 in terms of computational effort.

3.6.1 Global error scaling

We start by discussing the scaling of the global simulation error presented in the previous Section. According to the consideration made in Section 3.2.1, the asymmetric and symmetric step SSFM are respectively a first and second order method, i.e., their global error scales respectively as $\mathcal{O}(h)$ and $\mathcal{O}\left(h^2\right)$. The error variance should thus scale as $\mathcal{O}\left(h^2\right)$ and $\mathcal{O}\left(h^4\right)$. However, such an intuition is generally not followed in the figures presented in the previous Section. In Section 3.2.3 we theoretically discussed the validity of such scaling assumptions, showing that at large step length they could not be valid anymore.

To address this point, a wide range of accuracy has been studied, starting from very small steps. The chosen setup was like the one described in Section 3.5.1, but considering just a WDM signal composed of 7 channels propagating over a single SMF span. Such a setup has been chosen in order to perform high-accuracy simulations in a reasonable amount of time. The scaling of the ratio $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$ versus Φ_{FWM} in the first step for this light setup is depicted in Fig. 3.23 for four different cases of study: the FWM-NLP and FWM-CLE, each one implemented both with asymmetric and symmetric step SSFM. The symmetric step SSFM has been implemented as in [11] by combining the linear operators of two consecutive steps in order to give a complexity comparable

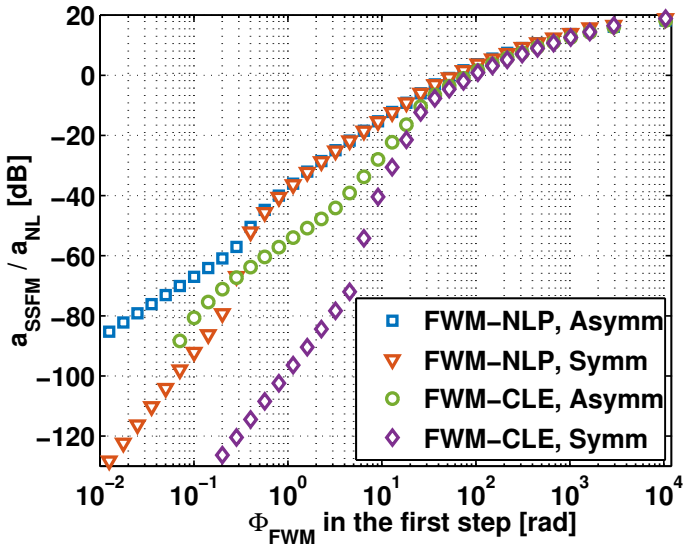

Figure 3.23: $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$ versus Φ_{FWM} for a simulation of a 9-channel WDM signal over a single SMF span. The accuracy of the simulation has been pushed to very high level in order to study the validity of the global simulation error scaling rules.

with the asymmetric step SSFM. As discussed in Section 3.2.3, the common intuition of asymmetric and symmetric step SSFM being first and second order methods is valid only for very small steps, where the approximation (3.9) holds. In such a region, the variance of the error scales as 20 dB/decade for the asymmetric step SSFM and 40 dB/decade for the symmetric step SSFM, which corresponds to the above-mentioned scaling by $\mathcal{O}\left(h^2\right)$ and $\mathcal{O}\left(h^4\right)$.

However, such regions are usually too much accurate and out of interest for the common SSFM user. As an example, for the FWM-CLE with asymmetric step reported in Fig, 3.23 such a scaling is valid for $a_{\mathrm{SSFM}}/a_{\mathrm{NL}} < -40$ dB, which corresponds to an SNR error of ~ 0.0005 dB. For the other rules such a region starts for even lower values of $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$.

By increasing the step length, i.e., Φ_{FWM}, such a scaling is not valid anymore, and all the methods tend to scale with a slope of $\sim$ 10 dB/decade, in agreement with the hypothesis of a Monte Carlo integration depicted in Section 3.2.3.

Each step-updating rule approaches such a slope differently. This comes from the different rate each rule uses to enlarge the step along the propagation. By referring to both (3.13) and (3.20), here reported for clarity:

$$\begin{aligned} L_{\mathrm{eff}}\left(h_{k+1}\right) &= L_{\mathrm{eff}}\left(h_k\right) e^{\alpha h_k} && \text{NLP} \\ h_{k+1} &= h_k e^{\frac{\alpha}{q} h_k} && \text{CLE} \end{aligned}$$

with $q = 2$ or 3 depending if asymmetric or symmetric step SSFM is considered, it can be noted that the FWM-NLP enlarges the step length more with respect to the FWM-CLE, since the stretching is not weighted by the coefficient q. Thus, for a given equal first step between the two rules, the condition (3.9) in Section 3.2.2 is better satisfied across the whole fiber by the FWM-CLE than the FWM-NLP since the steps along the propagation increase slower (see Fig. 3.6 as a reference). In the FWM-NLP instead the longer steps, i.e., the steps near the end of the fiber, are more likely to break such a condition, thus contributing to the global error as the Monte Carlo rule. For this reason the FWM-CLE is more sensible to the variations of the first step length in terms

of the scaling of the global error, as visible in Fig. 3.23 by the higher drop of the error variance for decreasing Φ_{FWM}. Moreover, since the FWM-CLE call for two different step-updating rules, depending if asymmetric or symmetric step SSFM is used, the two methods results in two different error scaling for increasing step length. Not surprisingly, the error of the FWM-NLP scales identically with both asymmetric and symmetric step SSFM for increasing step length, since such a rule does not distinguish between the two computational methods of the SSFM step.

We conclude thus that the common belief of the asymmetric and symmetric step SSFM being a first and second order method, respectively, does not hold for practical values of simulation accuracy. For practical accuracy, all the rules considered show a different scaling with the step length. A notable exception is the FWM-NLP, which keeps the same step-updating rule regardless if asymmetric or symmetric step SSFM is used to compute the single step solution. In this case, outside the region where (3.9) holds, the asymmetric and symmetric step SSFM scale identically.

3.6.2 Dependence on number of steps

The analysis made in Section 3.5 highlighted the dependency of the simulation accuracy by the system parameters for all the rules considered. However, such an analysis does not give any hints about the computational effort of each step-updating rule. We address such a problem in this Section.

The computational effort of the SSFM obviously depends on the implementation. A good coder should use his ability to minimize the number of instructions of the code implementing the SSFM. However, the SSFM is basically an iterative procedure: the computed solution of the NLSE inside a given step will be used as the input of the next one to proceed along the propagation. Inside each step a minimum set of operations is necessary, such as FFT and complex multiplications. Since such operations are identical for each step, a good measure of the complexity of the simulation, i.e., the computational effort, is thus the number of steps.

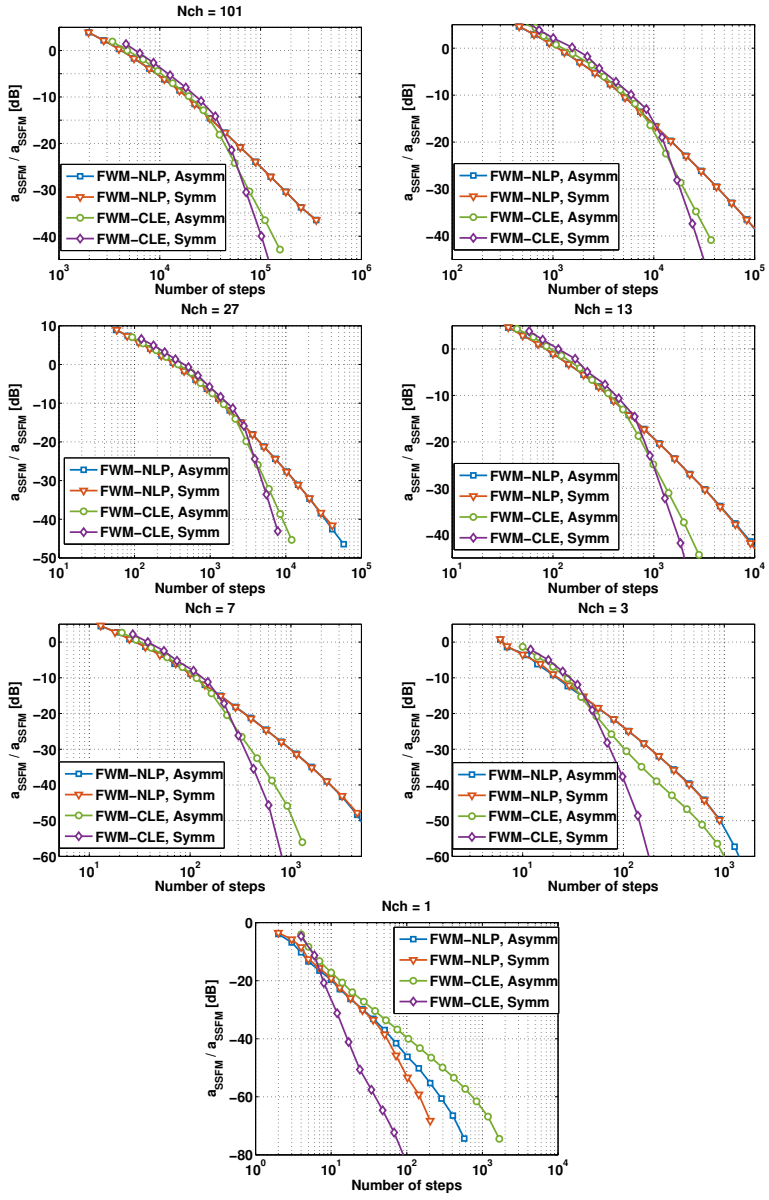

Figure 3.24: Simulation accuracy in terms of $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$ versus the number of steps in a single SMF span. Each plot refers to a different number of WDM channel, i,e, to a different bandwidth occupation, from 1 to 101 channels (5.05 THz of max bandwidth). An error on the SNR of 0.01 dB corresponds to a value of $\frac{a_{\mathrm{SSFM}}}{a_{\mathrm{NL}}} \simeq -25$ dB.

The scaling of the simulation accuracy with the number of steps is reported in Fig. 3.24 for the bandwidths investigated in Section 3.5, i.e., ranging from a single channel transmission up to a WDM filling the whole C-band. As usual, the accuracy is evaluated in terms of the ratio $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$. This time the propagation distance is set to just one span, in order to focus on the worst case of accuracy as discussed in Section 3.5.6. For all bandwidths, with the exceptions of 1 and 3 channel transmission, the scaling of the ratio $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$ is similar. For increasing number of steps, i.e., for decreasing Φ_{FWM}, FWM-CLE (3.20) shows a better accuracy with respect to FWM-NLP (3.13), for either asymmetric and symmetric step SSFM. Such a rule gives thus the best compromise between accuracy and computational effort. For $a_{\mathrm{SSFM}}/a_{\mathrm{NL}} \lesssim 30$ dB the symmetric FWM-CLE slightly outperforms the asymmetric one, needing fewer steps to achieve the same accuracy.

For smaller number of steps, i.e., for increasing Φ_{FWM}, the accuracy of all rules scales similarly with the number of steps in each bandwidth considered, in agreement with the discussion about the Monte Carlo integration made in Section 3.2.3. In this region, we found out that the FWM-NLP requires fewer steps than the FWM-CLE at a given accuracy.

To quantify such a gain, Fig. 3.25 shows the complexity increment of the propagation by using the FWM-NLP with respect to the FWM-CLE versus the ratio $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$. Such an increment is defined as:

$$\mu \triangleq \frac{N_{\mathrm{NLP}} - N_{\mathrm{CLE}}}{N_{\mathrm{CLE}}} \tag{3.24}$$

where N_{NLP} and N_{CLE} are respectively the number of step needed by the FWM-NLP and FWM-CLE to obtain a given accuracy of the simulation. At low values of $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$, i.e., for high accuracy, FWM-CLE outperforms FWM-NLP, while FWM-NLP might be better at low accuracy.

An exception to the above conclusion has been found for very small bandwidths, i.e., for 1 and 3 channel transmission, with asymmetric step SSFM, as visible in Fig. 3.24 and Fig. 3.25.

One can be interested now in knowing how the computational effort of the

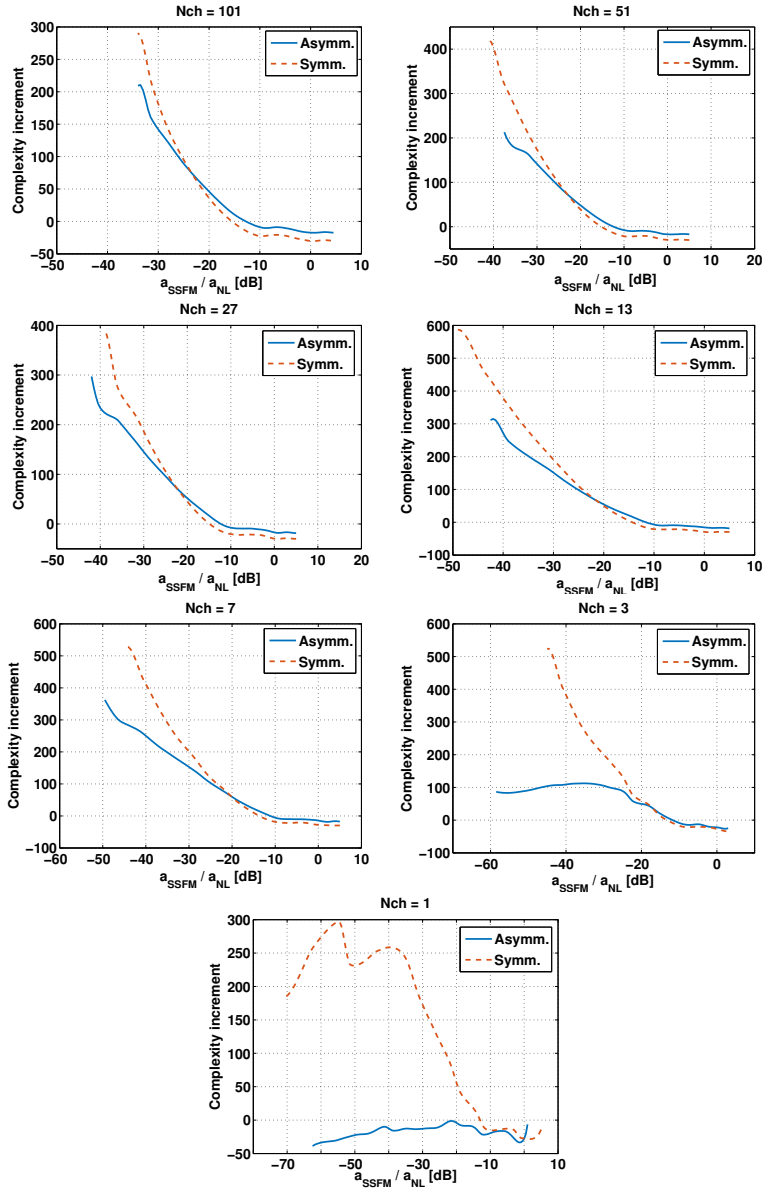

Figure 3.25: Complexity increment μ of the FWM-NLP with respect to the FWM-CLE, defined as (3.24), versus the ratio $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$ for the setups considered in Fig. 3.24, i.e., a single span propagation with variable bandwidth. The FWM-NLP is more convenient than the FWM-CLE for $\mu < 0$.

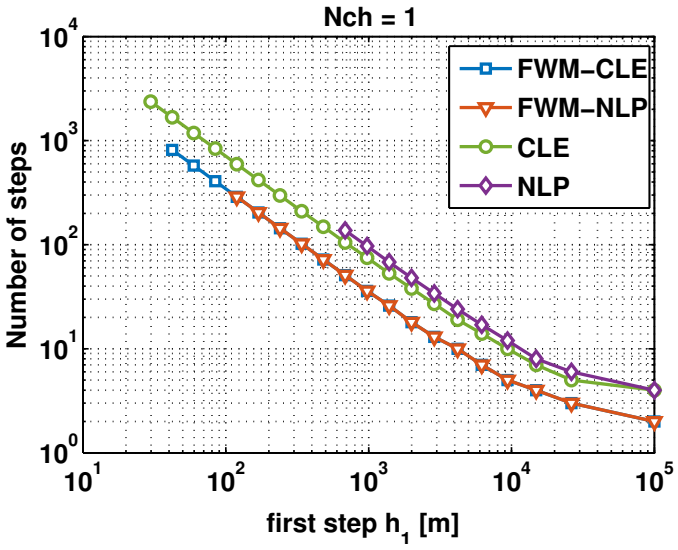

Figure 3.26: Number of steps in a single SMF span vs the corresponding first step length h_1. Single channel transmission.

simulation scales with the system parameters. We have seen in Section 3.5 that the accuracy of the simulation for a given step-updating rule depends only on the first step, whose length is set by the system parameters. Formally, due to the variable step length, the dependence on system parameters of the number of steps is different from the one of the first step. However, in most of our simulation the number of steps and the first step length are linearly dependent for not too large values of h_1, as depicted in Fig. 3.26, thus their dependence on system parameters should be the same. This is confirmed in Fig. 3.27, where the number of steps in the span is plotted versus the dispersion D (top) and the number of WDM channels (bottom) for a given value of simulation accuracy. The chosen accuracy corresponds to $a_{\mathrm{SSFM}}/a_{\mathrm{NL}} = -25$ dB after 20 spans, which calls for an SNR error around 0.01 dB, thus a practical value for the SNR estimation of a transmission. Both FWM-NLP and FWM-CLE are analyzed, each for both the asymmetric and symmetric step SSFM. It is worth

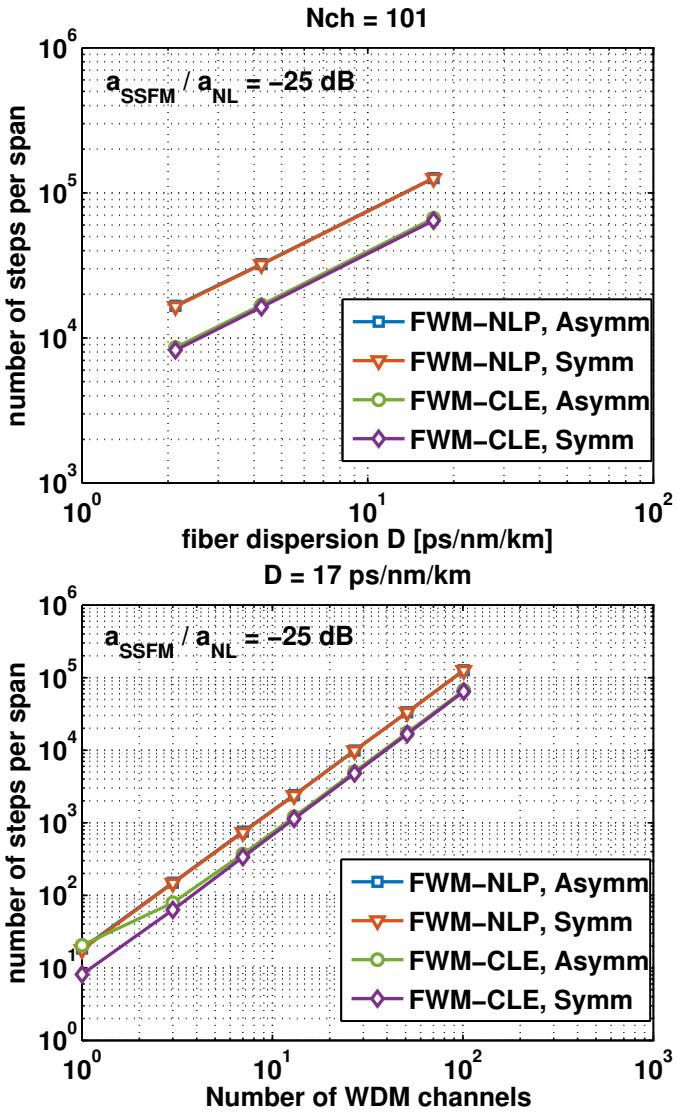

Figure 3.27: Number of steps per span of the simulation versus dispersion D (top) and number of WDM channels (bottom) for a fixed accuracy.

noting that at SNR error of 0.01 dB both asymmetric and symmetric step SSFM show a comparable complexity at the considered accuracy. As expected, the scaling of the computational effort reported in Fig. 3.27 is linear with the fiber dispersion D and quadratic with the bandwidth occupation.

Such a result of the computational effort scaling with bandwidth and dispersion highlights a problem that can be of concern for wideband transmissions. By increasing the signal bandwidth the complexity of the simulation scales quadratically, thus by doubling the bandwidth occupation of a WDM signal the computational effort to keep the same simulation accuracy is expected to be four times higher. Such an increment is the one expected by increasing the maximum simulation bandwidth from the C-band of the EDFA, which is around 5 THz, up to the 10 THz of the C+L band [23,24,62] as the research is doing nowadays. Up to such bandwidths, the time needed for an accurate SSFM simulation will become extremely high. In such a scenario a fine-tuning of the accuracy of the SSFM is thus mandatory, in order to obtain reliable results and at the same time without wasting any computational effort to perform over-accurate simulations.

3.6.3 How to setup an accurate simulation

We have showed in this chapter how to control the SSFM numerical error of the SNR of a simulation. We recap now the major results.

Table 3.1 shows the scaling of the first step of the simulation h_1 with the proposed method (FWM-NLP and FWM-CLE) based on FWM and the two rules known in literature analyzed in this work, i.e., the NLP and the CLE. Our proposal grants a constant accuracy on the SNR of the simulation for any variation of the signal power P, bandwidth B_{WDM} and fiber dispersion D. From the table we can see that the SNR error is power-independent, thus in contrast with the NLP and CLE which vary h_1 with power. It is worth noting that both NLP and CLE shrink h_1 for increasing power, resulting in more accurate simulations at the expense of the computational effort. The CLE tracks correctly the variations of fiber dispersion by shrinking h_1 linearly

	NLP	CLE	FWM-NLP	FWM-CLE
Acc. param.	$\Delta\phi$ [rad]	Ψ_G [a.u.]	Φ_{FWM} [rad]	
$h_1(P)$	$\mathcal{O}\left(P^{-1}\right)$	$\mathcal{O}\left(P^{-\frac{1}{q-1}}\right)$	$\mathcal{O}(1)$	
$h_1(D)$	$\mathcal{O}(1)$	$\mathcal{O}\left(D^{-1}\right)$	$\mathcal{O}\left(D^{-1}\right)$	
$h_1(B_{\mathrm{WDM}})$	$\simeq \mathcal{O}\left(B_{\mathrm{WDM}}^{-1}\right)$	$\mathcal{O}\left(B_{\mathrm{WDM}}^{-(2+\epsilon)}\right)$	$\mathcal{O}\left(B_{\mathrm{WDM}}^{-2}\right)$	

Table 3.1: Scaling of the first step h_1 with system parameters for the three different rules analyzed to set up the simulation accuracy. $q = 2, 3$ depending if asymmetric or symmetric step, and $\epsilon > 0$.

with D, identically to our method, while the NLP keeps it constant since it is unaware of such a parameter.

Both NLP and CLE shrink h_1 with the signal bandwidth B_{WDM} differently from our proposal, resulting thus in variation of the simulation accuracy with this parameter. However, the CLE shrinks the h_1 faster-than-quadratically, resulting thus in over-accurate simulations for increasing bandwidths. It is worth noting that such a faster-than-quadratic scale is due to the power dependence of the first step in such a rule. The NLP, instead, shrinks h_1 linearly with B_{WDM}, resulting in less accuracy for increasing bandwidth. The value of $\Delta\phi$, i.e., the accuracy parameter for this rule, should be thus reduced accordingly to follow such accuracy reduction.

By Fig. 3.18a) we can see that for a signal bandwidth of 5 THz and an SNR error around 0.01 dB, the NLP calls for a $\Delta\phi \simeq 1.5 \cdot 10^{-4}$ rad, a value usually smaller than what is used in the literature.

Due to its ability to track the variations of the system parameters, our FWM-aware proposal is thus an ideal parameter to set up the simulation. Table 3.2 shows the values of Φ_{FWM} to be used to target an SNR error of the simulation around 0.01 dB, corresponding to a $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$ ratio of -25 dB,

	FWM-NLP	FWM-CLE	NLP	CLE
Symm. step	4	20	system dependent	system dependent
Asmm. step	4	15	system dependent	system dependent

Table 3.2: Values of Φ_{FWM} [rad] granting a simulation error on the SNR around 0.01 dB ($a_{\mathrm{SSFM}}/a_{\mathrm{NL}} \simeq -25$ dB) for the step-updating rules and step types analyzed in this work.

for both step-updating rules (FWM-NLP and FWM-CLE), and the splitting scheme, i.e., asymmetric or symmetric step. It is worth noting that Table 3.2 does not highlight the computational effort, thus it does not give any hint on which of the four method is best. Such an analysis can be done by looking at Fig. 3.24 in Section 3.6.2, which we recap here briefly. At this accuracy, the FWM-CLE outperforms the FWM-NLP, calling for almost half the number of steps. Symmetric and asymmetric step in this case call for almost the same number of steps, thus they can be used indifferently. However, for more accurate simulations, i.e., for smaller Φ_{FWM}, the symmetric FWM-CLE outperforms even the asymmetric one, thus it is the optimal choice for very accurate simulations.

The FWM-NLP finds its application in cases where keeping a light computational effort is mandatory, even at the expense of the accuracy. In this case, corresponding to high values of Φ_{FWM}, i.e., big steps, the FWM-NLP outperforms the FWM-CLE as visible in Fig. 3.25 in Section 3.6.2.

3.6.4 Inclusion of the ASE noise

In all the previous results we considered the worst case for the SSFM accuracy in absence of ASE noise. However, in typical communications, the optimal transmission power implies a non-negligible amount of such a noise in the received SNR.

As introduced in Section 3.1.1, including the ASE noise in the SSFM accuracy means to relate the variance σ^2_{SSFM} to the sum $\sigma^2_{\mathrm{ASE}} + \sigma^2_{\mathrm{NLI}}$ instead of σ^2_{NLI} alone. This way, (3.6) transforms in:

$$\frac{\mathrm{SNR}}{\widehat{\mathrm{SNR}}} = 1 + \frac{a_{\mathrm{SSFM}}}{\frac{\sigma^2_{\mathrm{ASE}}}{P^3} + a_{\mathrm{NL}}}, \tag{3.25}$$

thus resulting into an increment of the SNR accuracy. Precisely quantify in advance such an impact could be a difficult task, since σ^2_{ASE} cannot be a-priori related to the nonlinear variance a_{NL} of the link, which is unknown before running the simulation. However, based on the results in the previous sections, some useful consideration can be made. In fact, it is known that at the nonlinear threshold (NLT) P_{NLT}, i.e., the power granting the maximum SNR of the transmission, the ASE variance is $\sigma^2_{\mathrm{ASE}} = 2\sigma^2_{\mathrm{NLI}}$ [53]. Equation (3.25) transform thus in:

$$\left.\frac{\mathrm{SNR}}{\widehat{\mathrm{SNR}}}\right|_{@P_{\mathrm{NLT}}} = 1 + \frac{a_{\mathrm{SSFM}}}{3a_{\mathrm{NL}}}, \tag{3.26}$$

thus gaining a factor 3 in the accuracy parameter $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$ showed in the previous sections, which corresponds to shift down the curves in Fig. 3.24 by 4.77 dB.

To better quantify such a gain in the simulation accuracy we compared in Fig. 3.28 the SNR error due to the SSFM versus the number of steps in the worst case without ASE noise and at P_{NLT}. The WDM signal was composed of 101 channels after 1 SMF span of propagation. The SNR error in absence of ASE has been estimated from simulations while the one at P_{NLT} by applying (3.26) to the results of Fig. 3.24. From Fig. 3.28 we can see that a gain in terms of number of steps for a given SNR error is indeed achieved for both the FWM-CLE and FWM-NLP, although the gain of FWM-CLE is smaller than FWM-NLP. Note that with the FWM-CLE criterion the reduction in number of steps is lower than in the FWM-NLP method. This is due to the higher slope of aNL vs number of steps of the FWM-CLE at Number of steps > 30000, as visible in Fig 3.24.

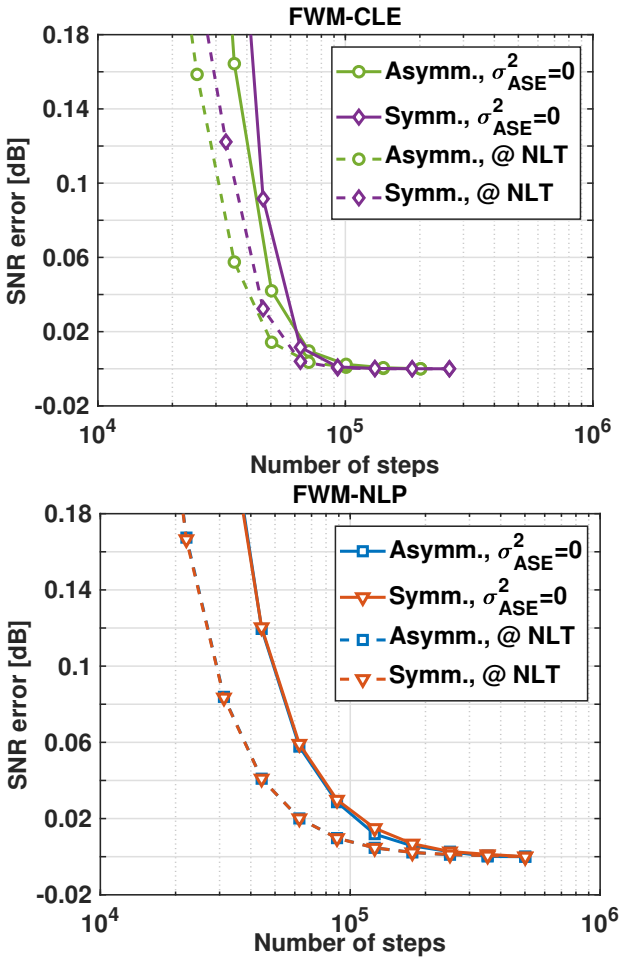

Figure 3.28: SNR error due to SSFM versus number of simulation steps for the FWM-CLE (top) and the FWM-NLP (bottom). Solid curves are obtained in absence of ASE noise while dashed curves refer to the SNR error at the nonlinear threshold P_{NLT}, estimated from $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$ in absence of ASE by (3.26). 101 channel WDM signal. Single SMF span propagation.

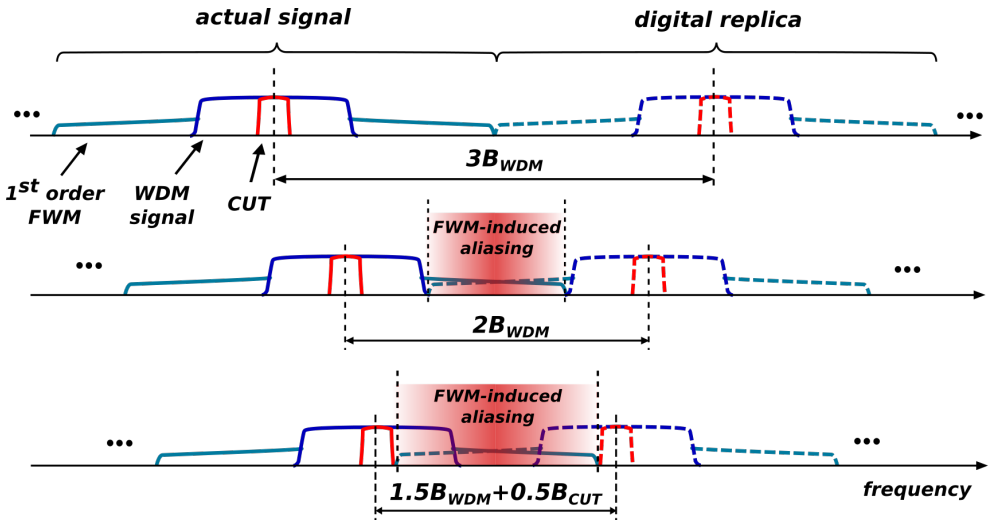

Figure 3.29: Sketch of the FWM-aliasing problem due to discretization. The first-order FWM does not induce aliasing for a simulation bandwidth $B_{\mathrm{Nt}} \geq 3B_{\mathrm{WDM}}$. B_{CUT} is the bandwidth of the channel under test.

In summary, if one is intereseted in estimating the SNR up to the nonlinear threshold, it is possible to increase the ratio assfm/anl analyzed in the previous sections by a factor 3, and hence increase the accuracy parameter according to its relation showed in Fig 3.17 After collecting all SNR by varying the power, the NLT can be estimated, which is expected to be within the target error. However, all the results with power $>$ NLT should be rejected since they should not satisfy the target error.

3.6.5 Signal discretization

In this Section we provide general rules to accurately define a digital signal in SSFM simulations for optical transmissions [33].

3.6.5.1 Numerical bandwidth

Signals have to be sampled in a numerical simulation to avoid spectral aliasing. The correct number of samples N_{t} per symbol on a digital signal is given by the Nyquist-Shannon theorem and must satisfy:

$$B_{\mathrm{Nt}} \geq B$$

where B_{Nt} is the bandwidth called by the discretization and B is the bilateral bandwidth of the signal. In the context of WDM optical propagations, such a condition translates in:

$$B_{\mathrm{Nt}} \triangleq N_{\mathrm{t}} R \geq \kappa \cdot B_{\mathrm{WDM}}$$

where R is the channel symbol rate and B_{WDM} the bandwidth of the WDM signal entering the fiber. $\kappa \geq 1$ is an expansion factor to account for the signal bandwidth enlargement due to FWM. Such a factor can be set by the user depending on how much FWM one want to correctly reproduce. For example, to avoid aliasing on the first-order FWM, the simulation bandwidth should be set to $B_{\mathrm{Nt}} \geq 3B_{\mathrm{WDM}}$ as sketched in Fig. 3.29. By defining $\mathrm{FWM}_{\mathrm{out}}$, $\mathrm{FWM}_{\mathrm{in}}$ and $\mathrm{FWM}_{\mathrm{CUT}}$ the first order FWM falling outside the WDM bandwidth B_{WDM}, inside the WDM bandwidth and inside the bandwidth of the channel under test (CUT), Table 3.3 reports the minimum value of B_{Nt} needed to avoid aliasing on each FWM component.

Table 3.3: Numerical bandwidth B_{Nt} setup to correctly reproduce first order FWM.

B_{Nt}	$\mathrm{FWM}_{\mathrm{out}}$	$\mathrm{FWM}_{\mathrm{in}}$	$\mathrm{FWM}_{\mathrm{CUT}}$
$\geq 3B_{\mathrm{WDM}}$	✓	✓	✓
$\geq 2B_{\mathrm{WDM}}$		✓	✓
$\geq \frac{3}{2}B_{\mathrm{WDM}} + \frac{1}{2}B_{\mathrm{CUT}}$			✓

3.6.5.2 Sequence length

While propagating a WDM signal along an optical link, each channel experiences a specific group delay. Since digital sequences are intrinsically periodic due to discrete Fourier transform operations, if the walk-off is longer than the sequence length temporal-aliasing occurs, which can lead to artificial correlations in circular convolutions. To completely avoid this numerical artifact the sequence length N_{seq} should be longer than the maximum walk-off between the side frequencies of the WDM signal, i.e.:

$$N_{\text{seq}} = |D_{\text{cum}}|\, B_{\text{WDM}} \frac{\lambda^2}{c} R \cdot 10^{-3} \quad \text{[symbols]}$$

where B_{WDM} [GHz] is the WDM bandwidth, R [Gbaud] the channel symbol rate, λ [nm] the central wavelength of the WDM comb, c [m/s] the speed of light, D_{cum} [ps/nm] the peak-to-peak accumulated dispersion along the link. In particular, for dispersion uncompensated links D_{cum} is the dispersion accumulated from input to output.

As a reference, the 101-channel curves of Fig. 3.17 in Section 3.5.5, for a bandwidth occupation of $B_{\text{WDM}} = 5$ THz, call for a sequence of $N_{\text{seq}} > 66709$ symbols.

3.6.6 Complexity issues in pratical implementations

This Subsection addresses the implementation of the SSFM and its relation to the simulation time. The SSFM is a pseudo-spectral method, which in a nutshell means that it obtains the solution at the end of each step by switching back and forth between the time and the frequency domains. In fact, the solution of the NLSE with applied only the linear operator (2.4) or the nonlinear operator (2.8) are a linear filtering and an exponential operation, respectively, thus efficiently performed respectively in the frequency and in the time domain. The switching between the two domains is done by performing FFTs on the digital signal.

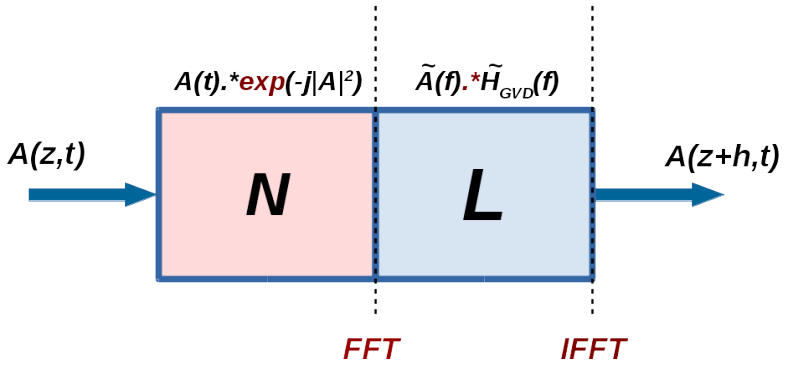

Figure 3.30: Sketch of the implementation of the SSFM with asymmetric step, with highlighted the fundamental operations.

While the complexity of the code implementing the algorithm may vary depending on the coder's ability, there are some unavoidable fundamental operations to implement the SSFM, that are highlighted in Fig. 3.30 for the asymmetric step SSFM (2.42). Such operations are:

- **FFT/IFFT**: used to switch between frequency/time domain. There are at least one pair of FFT/IFFT inside a single step when implementing the asymmetric step. In case of symmetric step, the number of FFT/IFFT pairs doubles due to the double switching between domains inside a single step. However, by combining the linear operators $\mathcal{L}$ of two consecutive steps (See [11]), the complexity of the symmetric step is identical to the asymmetric one plus a minor overhead.

- **Exponential (*exp*)**: used to perform the phase shift due to the Kerr effect in the nonlinear operator $\mathcal{N}$.

- **Point-wise multiplication (*.**)**: used to perform the linear filtering of the fiber linear effects, indicated by $\tilde{H}_{GVD}(f)$ in Fig. 3.30, as well as the application of the Kerr phase shift on the signal in the nonlinear step.

When the SSFM algorithm runs on a central processing unit (CPU), the

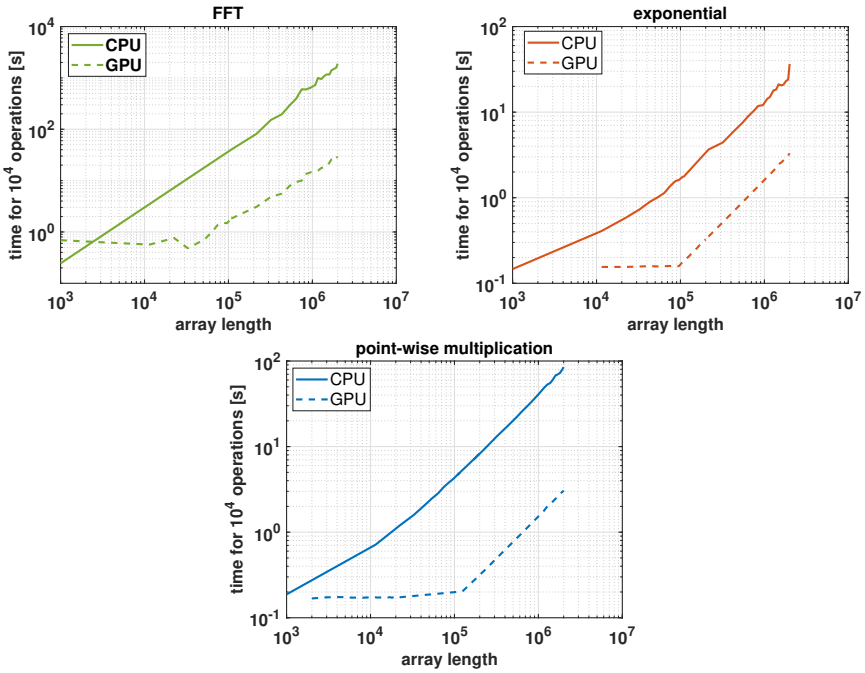

Figure 3.31: Computational time to perform the fundamental operations of the SSFM (FFT, exponential, point-wise multiplication) versus the vector length N in Matlab environment. CPU: 24-cores Intel(R) Xeon(R) CPU E5-2650 v4 running at 2.20GHz. GPU: GeForce GTX 1080 Ti. The gain of performing operation in GPU with respect to CPU is around a factor 15 for arrays longer than 10^5 elements.

complexity of the algorithm is dominated by the number of FFT/IFFT pairs, as depicted in Fig. 3.31 by the solid curves. The figure shows the computational time to perform 10^4 of the above-mentioned operations versus the length of the complex array on which the operation is performed. The CPU used for this test was a 24-cores Intel(R) Xeon(R) CPU E5-2650 v4 running at 2.20GHz with 130 GB of RAM, while the GPU was a GeForce GTX 1080 Ti. Matlab uses the FFTW algorithm [59] to perform FFTs/IFFTs, with a complexity scaling in the order of $\mathcal{O}(N\log(N))$ in CPU, with N the length of the array under FFT. Both the exponential and the multiplication have instead a linear complexity in N.

An effective way to speed up the algorithm which is becoming more and more popular is to run it through a graphics processing unit (GPU). To allow the GPU computation of the code, one must put the array over which the operation are performed, i.e., the complex array describing the signal, inside the GPU, which acts as a volatile memory during the execution. This allows to drastically reduce the computational time of the fundamental operations of the SSFM, as showed in Fig, 3.31 by the dashed curves. We can see that for array length $> 10^5$ the computational time is reduced for all three operations by at least a factor 10, with the main gain in the FFT reaching a factor 60 for $N = 2 \cdot 10^6$. It is worth noting that for smaller N the GPU time seems to remain constant whatever the array length, thus for very small lengths the CPU could still outperform the GPU in time.

It is worth noting that the GPU memory is limited to a fixed value, thus limiting the maximum length of the array describing the signal. The GPU memory used in this work was around 12 Gbytes. Considering that a complex scalar is stored in Matlab with 16 bytes, and a dual polarization signal, the maximum memory occupation for the optical signal in our simulations was around 500 MB. Considering also an arbitrary factor of 10 for any other eventual user-defined auxiliary variable to be stored in the GPU with the signal, as well as any possible burst of memory increment due to GPU operations, we find that our simulations used nearly half of the available memory of the GPU,

thus not exploiting it at full potential.

3.7 Alternative integration methods

Many methods for numerical integration are known in the literature [60]. As seen in the previous chapter, the SSFM in its two most popular implementations, i.e., the asymmetric and the symmetric step SSFM, can be associated with the left-point and mid-point rules for numerical integration. Following this parallelism, in this section we examine and compare in terms of complexity alternative methods to numerically integrate the Manakov equation.

Two families of methods will be analyzed in this section: the *Adams-Bashforth* methods (AB) and the *Adams-Moulton* methods (AM). Both algorithms are linear multistep methods, meaning that they use the already computed solution of previous steps in the computation of the the next step to increase the order of accuracy.

The Runge-Kutta method, which is very popular to numerically solve ordinary differential equations, has been not considered here since its higher number of function evaluations of the integrand function to compute the next step solution yields a computational advantage only at very low accuracy, not of interest for SSFM simulations, especially in wideband regime where complexity is of particular concern.

3.7.1 The Adams-Bashforth methods

For a complete and exhaustive derivation of the methods we refer to [63]. The AB methods update the solution by using a polynomial interpolation of the solution found in the previous steps. The next step solution is then computed by multiplying the step length h with a linear combination of such "weighted past solutions". The coefficients, as well as the error scaling order, depend on the order of the polynomial used to find such coefficients as well as on the number of previous steps used to find the next step solution. In this study the third-order (AB3) and fourth-order (AB4) methods are considered. The

next step solution $U(z_{k+1}, t)$ for these two methods applied to the ME are, respectively:

$$U_{\mathrm{AB3}}(z_{k+1}, t) = U_{\mathrm{AB3}}(z_k, t) + \frac{h}{12}\left[23f(z_k, t) - 16f(z_{k-1}, t) + 5f(z_{k-2}, t)\right]$$
$$U_{\mathrm{AB4}}(z_{k+1}, t) = U_{\mathrm{AB4}}(z_k, t) + \frac{h}{24}\left[55f(z_k, t) - 59f(z_{k-1}, t) + 37f(z_{k-2}, t) - 9f(z_{k-3}, t)\right]$$

where $f(z_{k-i}, t) = e^{-\mathcal{L}z_{k-i}}\mathcal{N}\left(e^{\mathcal{L}z_{k-i}}U_{\mathrm{AB}}(z_{k-i}, t)\right)$ is the integrand function of the nonlinearity in (2.38) calculated in the step $k-i$, with i ranging from 0 to $q-1$, q the order of the method.

One of the problems of a generic q^{th}-order AB is that the first $q-1$ steps cannot be obtained by the method, since not enough previous steps solution have been computed yet. Other methods are thus needed to obtain such points. Moreover, it is important that the local error on such points scales as the local error of the AB method to ensure a correct scaling of the global one. For this purpose, as usual in the literature, we used a Runge-Kutta method of the same order as AB3 and AB4 to obtain respectively the first two and three steps of the simulation. This adds a minor overhead to the overall complexity.

3.7.2 The Adams-Moulton method

The AM methods are a variant of AB which in numerical integration grant a more stable solution than AB methods. The two families are very similar. The main difference is that AM methods are implicit methods, hence they use a prediction on the next step solution in addiction to the previous ones. Such a prediction must be computed by using a method granting the same or higher local accuracy. In this study we used an AB method of the same order as the forward AB one. The next step solution for the 3^{th}-order AM method used in this study (AM3) is:

$$U_{\mathrm{AM3}}(z_{k+1}, t) = U_{\mathrm{AM3}}(z_k, t) + \frac{h}{12}\left[5f(z_{k+1}, t) + 8f(z_k, t) - 1f(z_{k-1}, t)\right]$$

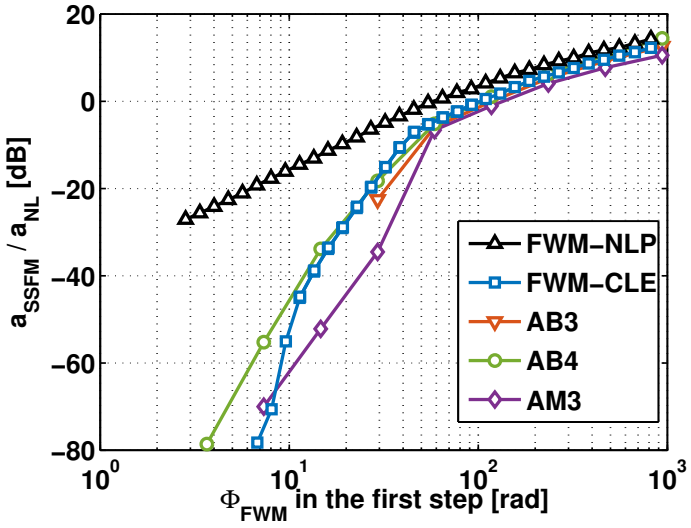

Figure 3.32: $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$ ratio as a function of Φ_{FWM} in the first step. Single span propagation over SMF.

with $f\left(z_{k+1},t\right)=e^{-\mathcal{L}z_{k+1}}\mathcal{N}\left(e^{\mathcal{L}z_{k+1}}U_{\mathrm{AB3}}\left(z_{k+1},t\right)\right)$.

The drawback of such a method is that the predictive step calls for doubling the computational effort.

3.7.3 Numerical results

The investigation has been conducted in the same way described in Section 3.5.1, thus by decreasing Φ_{FWM} in the first step until reaching saturation of the SNR. The setup here was identical to Section 3.5.1, but with the following differences:

- single span propagation over SMF ($D=17$ ps/nm/km);
- 9 channel WDM signal with symbol rate $R=32$ Gbaud, evenly spaced by $\Delta f=37.5$ GHz;
- PDM-quadrature phase shift keying modulation (PDM-QPSK) instead of PDM-16QAM.

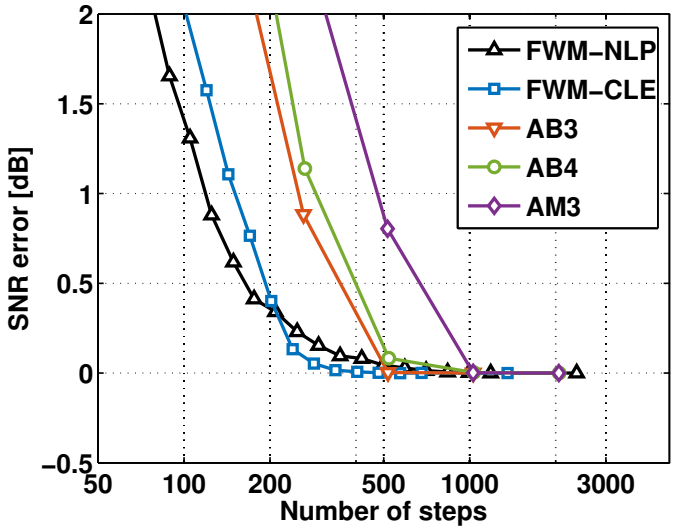

Figure 3.33: SNR error as a function of the number of steps. Single span propagation over SMF.

Figure 3.32 depicts the scaling of the ratio $a_{\mathrm{SSFM}}/a_{\mathrm{NL}}$ versus Φ_{FWM} in the first step of the simulation. Since in Section 3.5.6 we showed the first span to be a worst case for accuracy, we propagated the signal only on one SMF span. The three alternative methods considered, i.e., AB3, AB4 and AM3 have been compared with the FWM-NLP and FWM-CLE computed both with symmetric step SSFM. As visible in the figure, the AB3 and AB4 perform almost identically to the FWM-CLE, with the same drop of the error at small Φ_{FWM}. With the AM3 such a drop of numerical error is even more pronounced before decreasing with a smaller slope for decreasing Φ_{FWM}. It is worth noting that the same considerations made in Section 3.2.3 about the different error scaling for big steps applies even to these alternative methods.

However, Fig. 3.32 does not provide any information about the complexity of the new methods analyzed. Such a topic is addressed in Fig. 3.33 in terms of SNR error versus the number of function evaluations necessary to propagate

the signal to the end of the span. Again, we express the complexity in terms of number of steps, since the AB and AM updating add a minor computational cost. The figure shows that the minimum number of function evaluations is reached by the SSFM for both the FWM-NLP and FWM-CLE. This is because the SSFM, contrary to AB and AM, has the property to increase the step length along the propagation, with big savings in terms of the number of steps. AB and AM, instead, are forced to keep a constant step along distance, resulting in a bigger complexity despite showing similar accuracy at equal first step.

3.8 Conclusions

In this Chapter we investigated in-depth the numerical error of the SSFM simulations for optical transmissions in the framework of the perturbative solution of the ME, thus treating the SSFM error as a distributed additive interference, and we analyzed two common rules know in the literature used to set the accuracy of a simulation. We pushed our simulations up to a bandwidth of 5 Thz, i.e., covering the whole C-band of the EDFA used nowadays for optical transmissions. We showed that the error of the SNR in a simulation is signal-power independent, and scales quadratically with the signal bandwidth and linearly with the fiber dispersion. Such a scaling of the error is in contrast with the rules known today in the literature, which are power dependent. In particular, the nonlinear phase criterion [8], based only on the signal power, misses the inflation of the SSFM error due to the increment of the signal bandwidth and fiber dispersion, while the constant local error rule [9,10], despite being bandwidth and dispersion-aware, still depends on the signal power. We proposed a new parameter to set the first step of the simulation, based on FWM, which instead is able to track such variations. The new method to set the first step can be used in combination with one of the step-updating rules known in literature to grant a constant error on the SNR independently of signal power, bandwidth and fiber dispersion.

Moreover, we compared to the two step-updating rules that can be inferred

by the nonlinear phase criterion and the constant local error rule in terms of computational effort, for a wide range of simulation accuracy. We showed that the assumptions at the basis of the constant error rule hold only for high accuracy, i.e., for very small steps, thus the optimality of such a rule is limited to that scenario. Contrarily, for big simulations steps we found more convenient the step-updating rule inferred by the nonlinear phase criterion. As a future work, the analysis can be extended to SSFM simulations without FWM.

We then analyzed other methods to integrate the NLSE alternative to the SSFM, namely the Adams-Bashforth and the Adams-Moulton methods, showing that in any of the considered cases the SSFM is still more accurate for a given computational effort.

Chapter 4

Polarization dependent loss in nonlinear regime

In this chapter we present a numerical investigation of the impact of polarization dependent loss (PDL) on the SNR of the transmission, focusing on the interaction between the PDL and the Kerr nonlinearity of the fiber. Section 4.1 will introduce the reader to the PDL and its related problems; Section 4.2 will describe theoretically both the PDL and a popular model to predict its impact on the signal to noise ratio. Finally, a study of the SNR statistics induced by PDL in nonlinear transmission will be presented and discussed.

4.1 Introduction

One of the properties of an optical signal is its polarization. Such a property describes the spatial orientation of the electric field propagating in the fiber. By the early work of Poole *et al.* [37,38], the electric field propagating in the fiber can be expressed as linear combination of two signals, each one describing one of two orthogonal polarizations of the electric field. The spatial orientation of the two polarizations is defined by the SOP [37]. In nowadays communications the polarization of the optical field is exploited to increase the capacity

of the optical channel by transmitting signals in the regime of polarization division multiplexing, where two different and typically independent streams of information symbols are modulated in the two orthogonal polarizations of the optical field.

During the propagation in an optical fiber, the optical signals typically experience polarization effects like PMD and PDL. PMD is related to the frequency dependence of the fiber birefringence, which causes the two polarizations of the signal to experience a random differential group delay (DGD), i.e., a misalignment in time of the two signal polarizations at the receiver side, and a depolarization of the signal. The temporal fluctuations of such effect are slow in time compared to the symbol timing, such that, despite their random nature, they can be recovered at the receiver side by means of digital equalization [1,2].

Polarization dependent loss manifests instead in a different attenuation of the two orthogonal polarizations. Such an attenuation is mainly lumped into optical components, like optical amplifiers and wavelength selective switches (WSS) used to route the WDM channels along a network [64]. Typically any optical link includes many components, thus PDL accumulates along distance. Each PDL element is randomly-oriented, consequently even the total accumulation of PDL in the link is random [65–68]. The resulting polarization dependent loss, when expressed in dB, converges to Maxwellian statistics when a large number of PDL elements is considered in the link [65, 66]. With few PDL elements the statistics are more complicated [67, 68], even if they can be computed analytically as in [69]. PDL affects both the signal and the noise on the line. While the PDL affecting the signal can be recovered by using digital equalization, its effects on the system performances cannot be completely mitigated due to the interaction between PDL and the optical noise accumulated in the line. Such a beating between PDL and noise causes in fact unavoidable random fluctuations on the SNR at the receiver side of the communication [70–72]. Such fluctuations can eventually lead to the outage of the system, thus it is necessary to define some margins on the SNR to bound the

outage probability [73–75].

Typically, the impact of the PDL on the SNR of the transmission has been evaluated by considering the ASE noise as the only noise source of the link [70–75], thus neglecting the Kerr nonlinearity of the fiber. However, in nowadays optical communications the paradigm is to maximize the SNR of the transmission by setting the signal power to the nonlinear threshold [52, 53], i.e., the power achieving the maximum reachable SNR in the system. At such a power, the variance of the nonlinear interference is half the ASE noise variance, thus its impact on the system performances is not negligible. Due to the different nature of the two noises, the interaction between the PDL and the nonlinear noise may change the SNR statistics.

Recently, studies on the interaction between PDL and fiber nonlinearity has been performed [76–81], showing that by increasing the transmission power the SNR shows different statistics from the case considering only the ASE noise. However, the first few investigations lack of generality since in [76] only a single element of PDL has been considered, while in [77] all PDL elements were aligned together. More general studies with randomly-oriented PDL elements have been later conducted [78–81], showing different results depending on the link studied. A more detailed analysis is thus mandatory to confirm the results. Moreover, a more in-depth investigation is needed in order to better understand the dynamics of the PDL-NLI interactions. The following work is devoted to such an analysis.

Since PDL randomly varies in time [67, 68], a statistical approach must be performed to study the variability of the PDL-affected SNR. Many different realizations of PDL should be analyzed in order to explore the whole space of possible SNR realizations and reach thus low values of outage probability. Such investigations result in high computational effort when the PDL is investigated with split-step Fourier simulations. To face such a problem different models have been proposed for a fast estimation of the impact of PDL on the SNR of the transmission [74, 75]. The common characteristic of these models is that they consider the accumulation of PDL on a link affected by ASE only, thus

completely neglecting the impact of fiber nonlinearity. Such models are thus expected to present significant deviations from the real SNR statistics when fiber nonlinearity is of concern.

The work reported in this chapter has been done during the Author's internship at Nokia Bell Labs, Paris-Saclay, France, in the 6 months period between June 2017 and December 2017.

4.2 Theoretical background of the polarization dependent loss

As reported in Chapter 2, in a dual polarization scenario the optical signal can be expressed as a 2×1 vector $A \triangleq \begin{bmatrix} A_x, & A_y \end{bmatrix}^T$. Each PDL element is described by a 2×2 matrix as in [71, 75]:

$$\boldsymbol{M}_k \triangleq \boldsymbol{U}_k \begin{bmatrix} \sqrt{1+\gamma} & 0 \\ 0 & \sqrt{1-\gamma} \end{bmatrix} \boldsymbol{U}_k^{\dagger} \tag{4.1}$$

where U_k is the Jones representation of a random matrix describing a uniform-distributed rotation over the Stokes space [65], while $\dagger$ means transpose conjugate, and γ is the power attenuation/gain of the single polarization. It is worth noting (4.1) preserves the total power, hence if a polarization is attenuated, the other is inflated. Such a representation for the PDL element consists in considering the EDFAs generating the PDL as operating in constant output power mode, such that the total signal power after the PDL element is equal to the launch power [71]. The PDL is typically characterized by the parameter:

$$\rho = (1+\gamma) \,/\, (1-\gamma) \ ,$$

which defines the strength of the PDL element. Although such a parameter can be random as well, it is typically considered equal for all PDL elements for the sake of simplicity. Since each PDL element is randomly-oriented, PDL accumulation is not linear along distance [65,66]. The distribution of ρ assumes

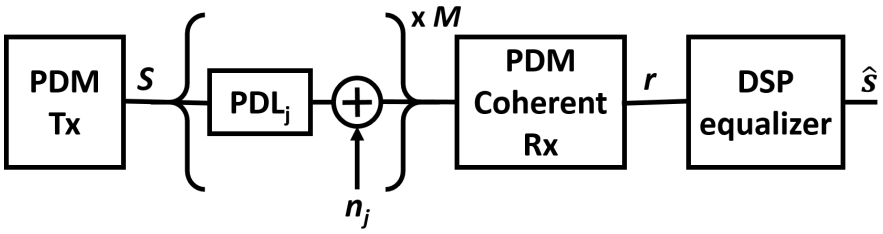

Figure 4.1: Scheme of the RCM for PDM optical transmissions. Each span is modeled as a concatenation of a randomly-oriented PDL element and an additive ASE noise source.

Maxwellian statistics when is expressed in dB and many PDL elements in the line are considered [65].

By passing among many PDL element during the propagation, the received signal can be expressed as:

$$A(z,t) = \boldsymbol{T}_{1k} A(0,t)$$

where:

$$\boldsymbol{T}_{1k} \triangleq \boldsymbol{M}_k \boldsymbol{M}_{k-1} \dots \boldsymbol{M}_1 \tag{4.2}$$

is the matrix describing the PDL accumulation up to coordinate z, with k the number of PDL elements crossed by the signal.

4.2.1 The reverse channel method

The reverse channel method (RCM) is a model proposed by Tao *et al.* to predict the SNR variability due to PDL when ASE is the only source of noise in the transmission [75]. The RCM works with a simplified optical link described by a sequence of PDL elements and additive noise sources. Each pair describes the ASE noise and the PDL introduced by a single EDFA over the link. The scheme of the link is reported in Fig. 4.1. The transmitted signal s is affected by all PDL elements of the line, while the generic ASE noise process $n_j = [n_{j,x}\,,\, n_{j,y}]^T$,

generated after the j^{th} PDL element, experience only the PDL from the span $j+1$ to the end of the link. Following the notation of Fig. 4.1, the signal after coherent detection is thus:

$$r = \boldsymbol{T}_{1,N_s} \cdot s + \boldsymbol{T}_{2,N_s} \cdot n_1 + \cdots + \boldsymbol{T}_{N_s,N_s} \cdot n_{N_s-1} + n_{N_s}$$

where $\boldsymbol{T}_{ij}$ is the PDL accumulation from element i to element j and N_s is the number of PDL elements in the link, corresponding by assumption to the number of spans. After detection, a perfect PDL equalizer is considered, with a zero-forcing channel inversion, which perfectly recovers the total PDL affecting the signal. Such a block thus applies the matrix T_{1,N_s}^{-1}, yielding:

$$\hat{s} = s + \boldsymbol{T}_{11}^{-1} \cdot n_1 + \cdots + \boldsymbol{T}_{1,N_s-1}^{-1} \cdot n_{N_s-1} + \boldsymbol{T}_{1,N_s}^{-1} \cdot n_{N_s} \,. \tag{4.3}$$

From (4.3) is clear that the PDL after equalization affects only the noise. Moreover, the PDL contribution to each noise process can be calculated iteratively starting from $\boldsymbol{T}_{11}^{-1} \equiv \boldsymbol{M}_1^{-1}$ and progressively adding PDL contributions to the inverse PDL matrix. At each step the resulting matrix is exactly $\boldsymbol{T}_{1j}^{-1}$, i.e., the PDL contribution affecting the ASE process n_j. The covariance matrix of the received noise $n \triangleq \hat{s} - s$:

$$\mathrm{Cov}\left[n\right] \triangleq \sigma_{\mathrm{ASE},x}^2 \cdot \sum_{j=1}^{N_s} \boldsymbol{T}_{1j}^{-1} \left(\boldsymbol{T}_{1j}^{-1}\right)^{\dagger}$$

where $\sigma_{\mathrm{ASE},x}^2$ is the ASE variance of a single polarization tributary in absence of PDL. It is worth noting in such a model the PDL affects the SNR of each polarization tributary by a factor of the ASE noise variance, thus it can be defined as:

$$\mathrm{SNR}_{x,y} \triangleq \frac{P}{\sigma_{\mathrm{ASE}}^2 \cdot \mathrm{PDL}_{x,y}} \,, \tag{4.4}$$

where $\frac{P}{\sigma_{\mathrm{ASE}}^2}$ is the SNR without PDL and $\mathrm{PDL}_{x,y}$ are the terms on the diagonal of the global PDL matrix $\sum_{j=1}^{N} \boldsymbol{T}_{1j}^{-1} \left(\boldsymbol{T}_{1j}^{-1}\right)^{\dagger}$. When the SNR is expressed in dB, the PDL results thus in an additive term.

The SNR randomness induced by PDL in such a model can be estimated by Monte Carlo simulations. Such a model will be compared to our numerical investigation, based on SSFM simulations, to check the impact of PDL in nonlinear regime.

4.3 Numerical setup

In this section we describe the numerical setup used for the numerical investigation.

The signal used for the propagation was a WDM comb composed of 13 channels. The number of channels has been chosen as a good compromise to keep a small simulation time for each propagation while having a significant amount of cross-channel nonlinearity on the signal. However, since the PDL is modeled as a frequency independent matrix, the bandwidth of the signal should not impact its interaction with the fiber nonlinearity, thus motivating our choice on a relatively small number of channels. All channels were considered with a symbol rate of $R = 32$ Gbaud, evenly spaced by $\Delta f = 37.5$ GHz and modulated using PDM-QPSK modulation format with root raised cosine pulses with roll off 0.01. The alignment between the signal SOP and the principal axis of the PDL elements has been showed to impact the SNR [70, 76]. To account for such a dependence we considered a randomly-oriented SOP for each WDM channel, uniformly distributed on the Poincaré sphere.

The transmitted signal was propagated on a link composed by 25 spans of standard SMF (dispersion $D = 16.7$ ps/nm/km, length $L = 100$ km, attenuation $\alpha = 0.22$ dB/km, nonlinear coefficient $\gamma = 1.3$ 1/W/km). We considered for the propagation either a DU link and a DM one. The difference between the two is that in the DU link the accumulated fiber dispersion is recovered only at the receiver, while in the DM case part of the accumulated dispersion is recovered at the end of each span, typically by using a dispersion compensating fiber. To clarify the difference, in Fig. 4.2 the accumulated dispersion along distance is sketched for the two cases. Please note DU links are characterized

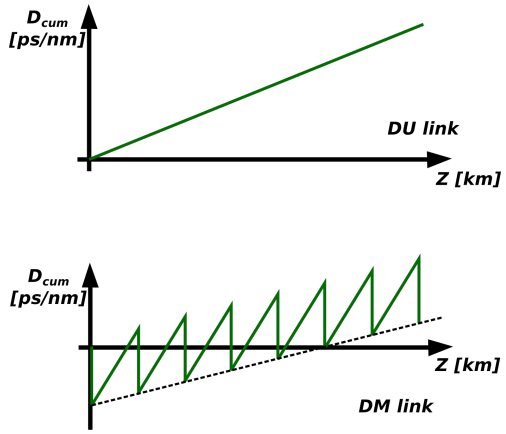

Figure 4.2: Sketch of the accumulated dispersion D_{cum} along distance for a DU link (top) and a DM link (bottom).

by a smaller nonlinearity than the DM, thus they are well-suitable for modern transmissions, where the large amount of accumulated GVD can be digitally recovered at the receiver. The DM configuration, however, is still present in already placed optical links, thus it is important to study also such a case. For our investigation, we considered a DM30 link, i.e., a DM link where after each span the amount of residual dispersion left on the signal is $D_{\text{RDPS}} = 30$ ps/nm. Typically, in DM links the input signal is pre-dispersed to reduce the nonlinearity on the link. We consider in the DM30 case a pre-dispersion of $D_{\text{PRE}} = -550$ ps/nm. In any case of study the residual dispersion on the signal at the receiver has been ideally compensated.

It has been shown in [82,83] that PMD has a negligible impact on the PDL for typical values of PMD and PDL in the fiber. This allows studying the two effects separately. The PMD on the fiber was thus set to zero in order to focus only on the variability generated by the PDL effect. Fiber propagation was modeled using the SSFM applied to the ME. The SSFM has been implemented by using the FWM-CLE rule (see Section 3.5) with symmetric step SSFM and a value of $\Phi_{\text{FWM}} = 20$ rad to set the first step length. This setup grants a

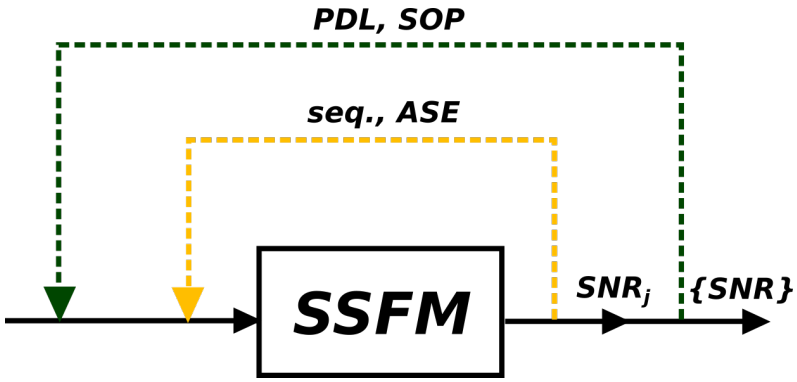

Figure 4.3: Nested loop structure to correctly estimate the SNR realizations due to PDL. SNR$_j$ is the correct SNR of the single PDL realization, cleaned by the randomness due to information sequences and ASE noise. {SNR} is the set of all PDL-affected SNR realizations. The block SSFM indicates the signal propagation.

numerical error on the SNR $\sim$ 0.01 dB, a reliable accuracy to study SNR fluctuations of fractions of dB. Fiber loss was recovered at the end of each span using EDFA characterized by a noise figure of 6 dB.

The receiver was composed by a matched filter followed by a data-aided, 5-tap least square butterfly filter to recover the signal polarization. The SNR of each polarization has been estimated evaluating the variance of the received clouds. A single PDL element was placed before each EDFA, randomly-oriented with respect to polarization tributaries. Each PDL element has been considered with $\rho = 0.5$ dB.

It is worth noting that, especially in DM links, the stochastic nature of the nonlinear impairments can be caused by the specific realization of the transmitted sequence [84]. The variability of the SNR due to the transmitted sequence can impair the SNR measurements if not correctly taken into account. Such a stochastic phenomenon, however, changes faster in time than the PDL, which shows a variability in the order of hours [67]. To isolate the SNR fluctuations due to PDL we thus used a two-nested Monte Carlo loop like the one used

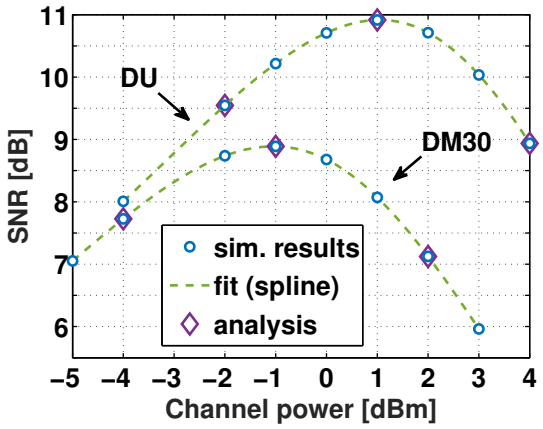

Figure 4.4: SNR vs channel power for both the DM30 and DU link configurations considered. The three powers used for the investigation are marked as "analysis". Each point of the curve has been obtained by averaging the SNR over 20 different PDL realizations.

in [84] and depicted in Fig. 4.3. The outer loop accounts for slow time-varying phenomena such as the PDL and the SOP variation of the signal, while the inner loop accounts for the fast phenomena such as the transmitted sequence realization and the ASE noise. The SNR variation with the fast phenomena has been averaged to mitigate the Monte Carlo uncertainty. The average SNR is thus the result of a double expectation:

$$\mathbb{E}[\mathrm{SNR}] = \mathbb{E}_{\psi}[\mathbb{E}_{\rho}[\mathrm{SNR}|\psi]]$$

where ψ is the set of the fast-varying random processes and ρ the set of the slowly-varying random processes.

4.4 Numerical results

We first simulated the system average performance in terms of SNR vs power to identify the regions where nonlinearity is the dominant effect. Such curves

	P_{LIN}	P_{NLT}	P_{NL}
DM30	-4	-1	2
DU	-2	1	4

Table 4.1: Channel powers [dBm] used for the analysis of the PDL for the DM30 and DU link configurations.

are depicted in Fig. 4.4 for both the DU and DM30 link. Each curve has been obtained by averaging the SNR over 20 random realizations of PDL and signal SOP. The SNR in Fig. 4.4 has been obtained from the single polarization BER, under the hypothesis of additive circular Gaussian noise and for QPSK modulation, as:

$$\mathrm{SNR} \triangleq 2\left\{\mathrm{erfc}^{-1}\left[2\left(\frac{1}{2}\mathrm{BER}_x + \frac{1}{2}\mathrm{BER}_y\right)\right]\right\}^2 \tag{4.5}$$

where:

$$\mathrm{BER}_i = \frac{1}{2}\mathrm{erfc}\left(\sqrt{\mathrm{SNR}_i/2}\right) \quad i \in [x, y] \tag{4.6}$$

is the bit error rate BER of the single x-y polarization and erfc is the complementary error function.

We considered for the analysis three different powers for each link configuration: the nonlinear threshold (NLT) P_{NLT}, i.e., the power giving the maximum SNR of the transmission [52], and two other powers, P_{LIN} and P_{NL}, for a transmission in linear and nonlinear regime, respectively. Such powers have been chosen as $P_{\mathrm{LIN}} = P_{\mathrm{NLT}} - 3$ dBm and $P_{\mathrm{NL}} = P_{\mathrm{NLT}} + 3$ dBm, i.e., way apart from the NLT to claim the nonlinear noise and the ASE noise negligible,respectively. Such powers are highlighted in Fig. 4.4 by the label "analysis", and are reported in Table 4.1.

For all powers we first check the minimum number of fiber propagations at a given PDL realization, i.e., the number of inner loop propagations of Fig. 4.3, to stabilize the SNR from the Monte Carlo uncertainty induced by the fast-varying stochastic phenomena. As expected by [84], the longest number of propagations needed to obtain a reliable SNR realizations was obtained for

the DM30 link by averaging the performances of 16 sequences of 4096 symbols each, for a total of 2^{16} symbols, and this value has been employed in all the following results for any power value and link configuration.

4.4.1 Probability density function of the SNR

In order to fairly compare the probability density functions (PDF) of all the three powers we analyzed the overall equivalent SNR deviation in dB from its mean value, i.e.:

$$\Delta\mathrm{SNR} \triangleq \mathrm{SNR}_{\mathrm{dB}} - \mathbb{E}\left[\mathrm{SNR}_{\mathrm{dB}}\right]$$

where SNR is the actual estimated SNR for each PDL realization and $\mathbb{E}\left[\mathrm{SNR}_{\mathrm{dB}}\right]$ is the average of the SNR expressed in dB overall random PDL realizations. Figure 4.5(top) shows the PDF of ΔSNR for the three powers considered and the DU link configuration. Each PDF has been obtained by analyzing 800 different PDL and signal SOP realizations. Each PDF shows an unbalanced tail toward the low SNR values, in agreement with literature [71, 72, 78]. For increasing power, it can be seen the PDF of ΔSNR spreads because of the PDL interplay with the fiber nonlinearity. Such a spreading is already present at P_{NLT}, thus it cannot be considered negligible in the transmission.

In Fig. 4.5(top) we also plotted the same PDF obtained by the reverse channel method, labeled in the plot as RCM, for both P_{NLT} and P_{NL}. Since such a model does not consider the nonlinearity of the fiber, one should expect the PDF obtained by the RCM to always match the PDF of the simulation at P_{LIN}, i.e., in the linear regime. This does not happen because of the high values of SNR on which the RCM works for not considering the nonlinearity at high power. At this high SNR, the intrinsic nonlinearity of equation (4.5) spreads the PDF of the overall transmission SNR, resulting in the RCM to overestimate its real statistics. Similar results have been found for the DM30 link under analysis, as reported in Fig. 4.5(bottom). Even here the spread of the PDF for increasing power is present, highlighting that the interaction between PDL and fiber nonlinearity is of concern even in these link configurations. The

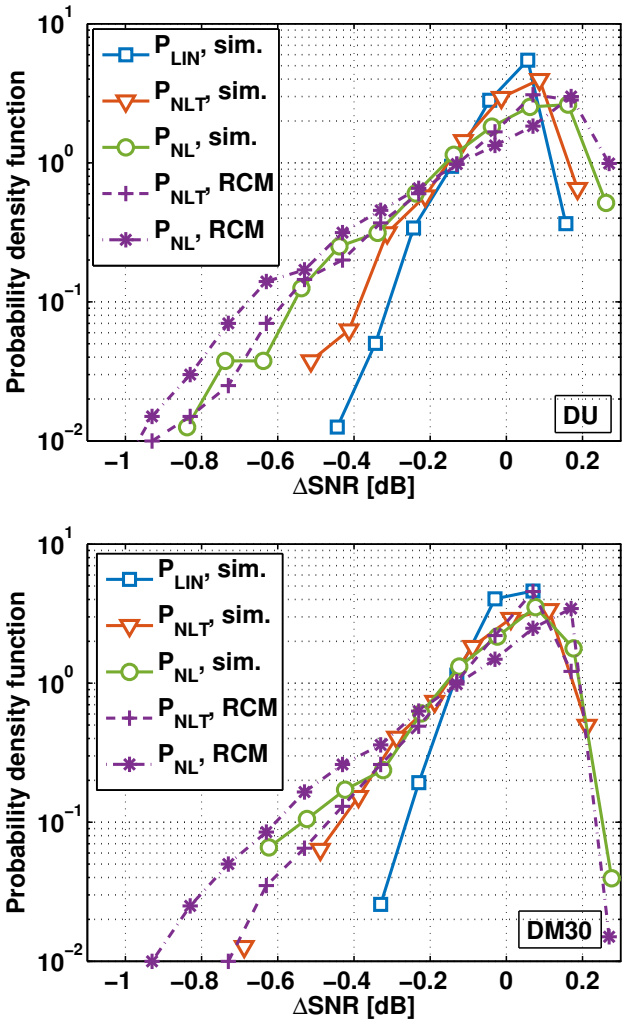

Figure 4.5: (solid marked curves) $\Delta\text{SNR} \triangleq \text{SNR} - \mathbb{E}[\text{SNR}]$ distribution for P_{LIN}, P_{NLT} and P_{NL}, estimated from simulations. (dashed and dash-dotted curves) ΔSNR distribution from RCM [75] predictions at P_{NLT} and and P_{NL}, respectively. 25×100 km SMF link. DU link (top) and DM30 link (bottom). PDFs computed over 800 SNR samples each.

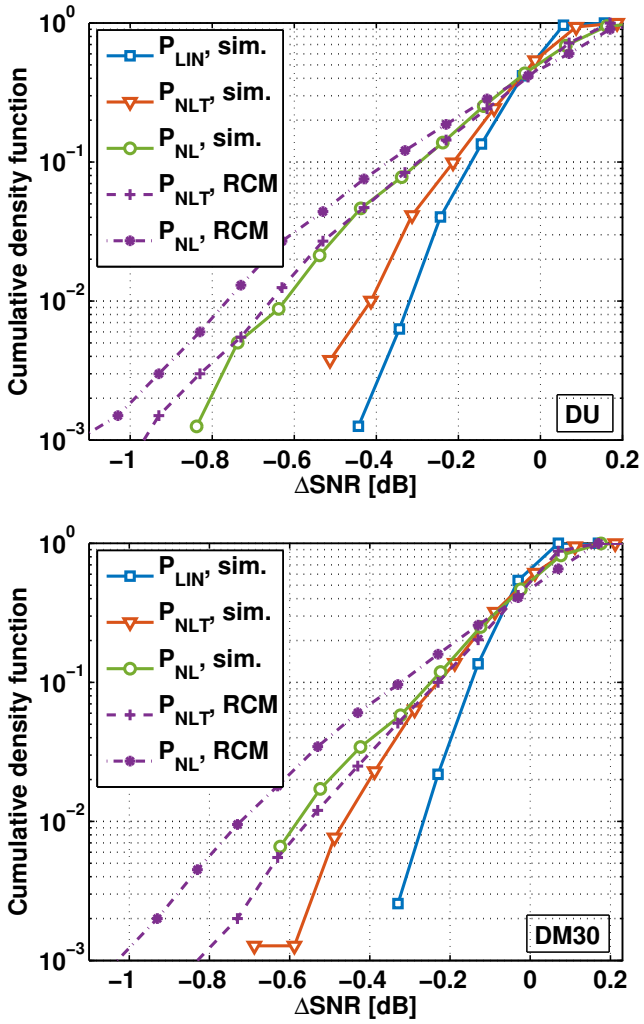

Figure 4.6: (solid marked curves) $\Delta\text{SNR} \triangleq \text{SNR} - E\,[\text{SNR}]$ cumulative distribution for P_{LIN}, P_{NLT} and P_{NL}, estimated from simulations. (dashed and dash-dotted curves) ΔSNR cumulative distribution from RCM [75] predictions at P_{NLT} and P_{NL}, respectively. 25×100 km SMF link. DU link (top) and DM30 link (bottom). PDFs computed over 800 SNR samples each.

broadening of the SNR PDF for increasing power affects the setting of the SNR margins to avoid system outage. To address this problem we plotted in Fig. 4.6 the cumulative distribution function (CDF) of the same ΔSNR realization as Fig. 4.5. ΔSNR can be ascribed to the SNR penalty, i.e., the SNR margin, that should be allocated in the transmission to obtain a system outage with a probability described by the curves.

For both the DU and DM30 link the simulations show the margin needed for P_{NLT} is higher than the one for P_{LIN} with a spread of ~ 0.1 dB for the DU case and ~ 0.2 dB, at outage probability $P\,[\mathrm{outage}] = 10^{-2}$. Even if such differences seem small at this outage probability, they can become significant at operative values of outage probability, which are around 10^{-5}. Such values have been impossible to reach in our numerical investigation, due to the high computational time required for the simulations. The prediction of the RCM, as expected, overestimates the system margins for increasing power due to an overestimation of the PDF spread reported in Fig. 4.5.

4.4.2 Single polarization SNR

To explain the broadening of the PDF for increasing power, we analyzed the PDF of the single polarization SNR, i.e., $\mathrm{SNR}_{x,y}$, for the DU link case. Since PDL involves a uniform distributed rotation in the Stokes space,the SNR of each x, y polarization has approximately the same statistics. In this work instead we found more interesting to classify the SNR into the best- or worst-performing between the two polarization tributaries, i.e.:

$$\begin{aligned} \mathrm{SNR}_{\mathrm{best}} &\triangleq \max\left[\mathrm{SNR}_x, \mathrm{SNR}_y\right] \\ \mathrm{SNR}_{\mathrm{worst}} &\triangleq \min\left[\mathrm{SNR}_x, \mathrm{SNR}_y\right] . \end{aligned}$$

Such a classification highlights some properties of the PDL-affected SNR that would have been missed by plotting only the PDF of $\mathrm{SNR}_{x,y}$, as depicted in Fig. 4.7(top) and Fig. 4.8(top) for the two extreme cases of P_{LIN} and P_{NL}, respectively. In linear regime, i.e., P_{LIN} (Fig. 4.7(top)), the two PDF of $\mathrm{SNR}_{\mathrm{best}}$ and $\mathrm{SNR}_{\mathrm{worst}}$ show a strong symmetry around the average SNR. In fact, each

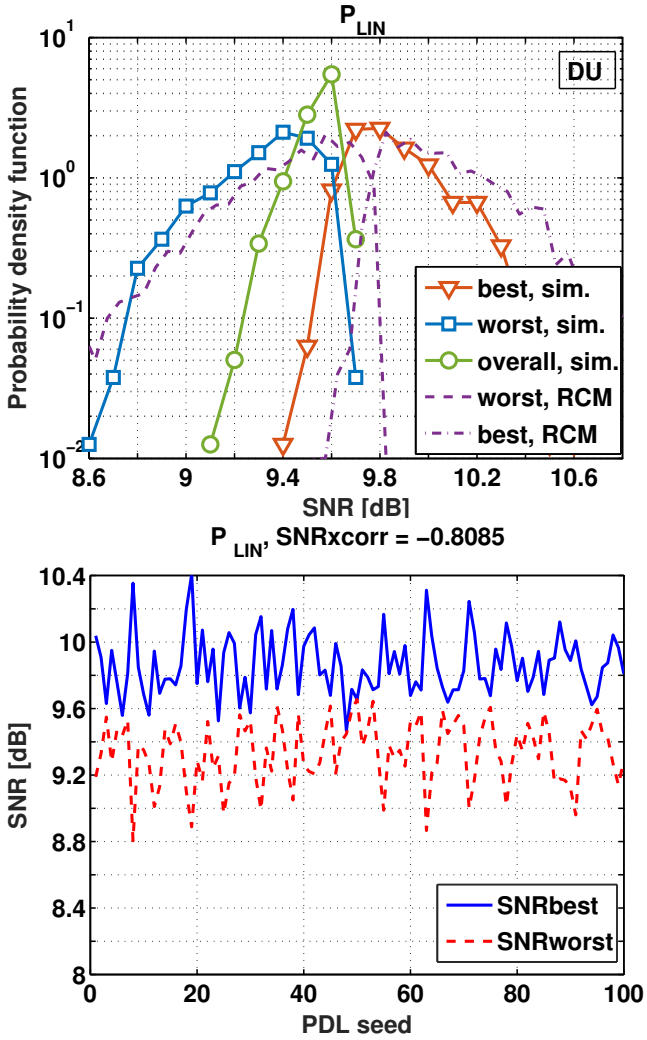

Figure 4.7: (top) Single polarization SNR distributions for best/worst-performing tributary (triangles/square), as well as the overall SNR (circles), estimated from simulations at $P_{\mathrm{LIN}} = -2$ dBm. 25×100 km DU-SMF link. 800 SNR samples. (dashed and dotted curves) SNR distribution from RCM [75] predictions. (bottom) First 100 realizations of $\mathrm{SNR}_{\mathrm{best}}$ and $\mathrm{SNR}_{\mathrm{worst}}$ obtained by simulations.

polarization PDF shows the same unbalanced tail, but toward opposite directions. Such a symmetry is due to the total energy conservation of each random PDL block, see (4.1). In fact, after recovering the random polarization rotation at the receiver, PDL is nothing more than a random symmetric redistribution of the total signal power between the two polarizations. For low powers, i.e., in linear transmission regime, the dominant noise is the signal-independent ASE noise, thus such a symmetry is maintained in the SNR of the best/worst polarization. Such an observation is confirmed by Fig. 4.7(bottom), which shows the first 100 realizations of $\mathrm{SNR}_{\mathrm{best}}$ and $\mathrm{SNR}_{\mathrm{worst}}$ for P_{LIN}. An antithetic correlation is clearly visible. Such a negative correlation helps to reduce the impact of the PDL on the overall SNR calculated as in (4.5) from the average BER of the two polarizations. This is visible in Fig. 4.7(top) by the circle markers, which highlights that the PDF of the overall SNR is narrower compared to the best/worst polarization.

In linear transmission regime the dominant noise is the ASE noise, thus not surprisingly the same classification of best/worst-performing SNR for the RCM model shows almost identical results, as depicted in Fig. 4.7(top), by the dashed curves. A minor shift between simulations and RCM PDF is present. This is due to the small residual contribution of nonlinearity to the SNR.

By increasing the transmission power up to P_{NL}, the symmetry between the PDF obtained in the linear regime breaks down. In this case, reported in Fig. 4.8(top), the two PDF of $\mathrm{SNR}_{\mathrm{best}}$ and $\mathrm{SNR}_{\mathrm{worst}}$ tilt in the same direction, with a more pronounced tail toward the low SNR values. Such an asymmetry is indicative of a different interaction between the PDL and the fiber nonlinearity, which is the dominant noise at high power, with respect to the interaction with ASE noise. This is confirmed in Fig. 4.8(bottom), where we collected again 100 realizations of $\mathrm{SNR}_{\mathrm{best}}$ and $\mathrm{SNR}_{\mathrm{worst}}$ like what done in Fig. 4.7(bottom). Contrary to the ASE case, now a strong correlation exists. Such a change in the dynamics of $\mathrm{SNR}_{\mathrm{best}}$ and $\mathrm{SNR}_{\mathrm{worst}}$ is thus the cause of the spreading of the overall SNR found in Fig. 4.5.

The different interaction of the PDL with nonlinearity is not surprisingly

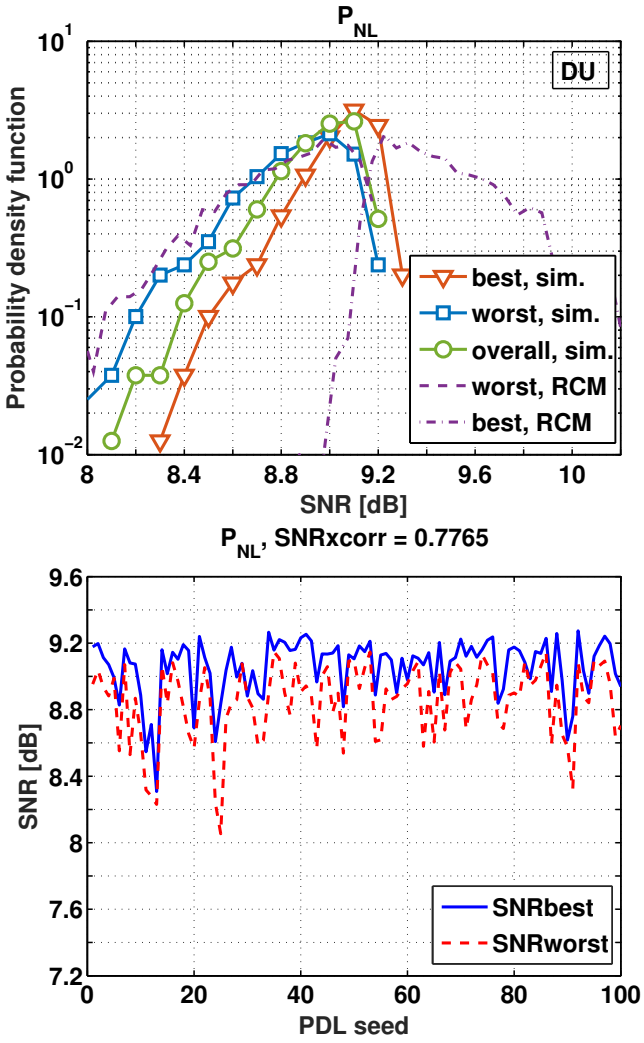

Figure 4.8: (top) Single polarization SNR distributions for best/worst-performing tributary (triangles/square), as well as the overall SNR (circles), estimated from simulations at $P_{\mathrm{NL}} = 4$ dBm. 25×100 km DU-SMF link. 800 SNR samples. (dashed and dotted curves) SNR distribution from RCM [75] predictions, shifted down to match the simulations average SNR to ease the comparison. (bottom) First 100 realizations of $\mathrm{SNR}_{\mathrm{best}}$ and $\mathrm{SNR}_{\mathrm{worst}}$ obtained by simulations.

missed by the RCM model, since it completely neglects the nonlinearity of the fiber. In fact, even in nonlinear transmission regime the RCM shows the same symmetry between best- and worst-performing SNR as in linear regime, as depicted in Fig. 4.8(top). It is worth noting that by neglecting the nonlinearity, the RCM shows for P_{NL} an average SNR way higher than the real one. For the sake of comparison, the SNR PDF relative to the RCM in this case has been manually shifted to the average SNR value of the simulations.

A similar trend has been observed by analyzing the best- and worst-performing polarizations in the DM30 link, as depicted in Fig. 4.9 for P_{LIN} and P_{NL}.

4.4.3 Individual contributions to nonlinearity

In Section 4.4.2 we demonstrated that the PDL interacts with fiber nonlinearity with a different mechanism with respect to ASE noise. To further investigate this aspect, we decoupled the nonlinearity into its main contributions. A similar investigation has been made in [36]. However, in [36] the decoupled nonlinearity has been investigated only in terms of the average penalty, i.e., the average contribution of all the PDL realizations. In this Section, we enlarge the topic by analyzing the distribution of the single polarizations of such nonlinearity contributions.

Referring to (2.25), the contributions to the nonlinearity can be classified in the SPM, XPM, and FWM. In dual polarization, the two degenerate cases of XPM can be further separated into a scalar XPM, i.e., the nonlinear phase rotation due to interfering channels and common to both channel polarizations, and an XPolM [47], i.e., the nonlinear SOP rotation on the Poincaré sphere inducing nonlinear polarization crosstalk.

Once separated the nonlinear contributions as (2.25), we simulated the propagation of the WDM signal selectively activating in the SSFM the nonlinear as in [36]. The link was composed of 25×100 km of SMF in DU configuration. The channel power was $P_{\mathrm{NLT}} = 1$ dBm. This time we focused only on the interaction between PDL and fiber nonlinearity, thus we forced all the EDFA noiseless. The same 800 PDL realizations of the previous section have

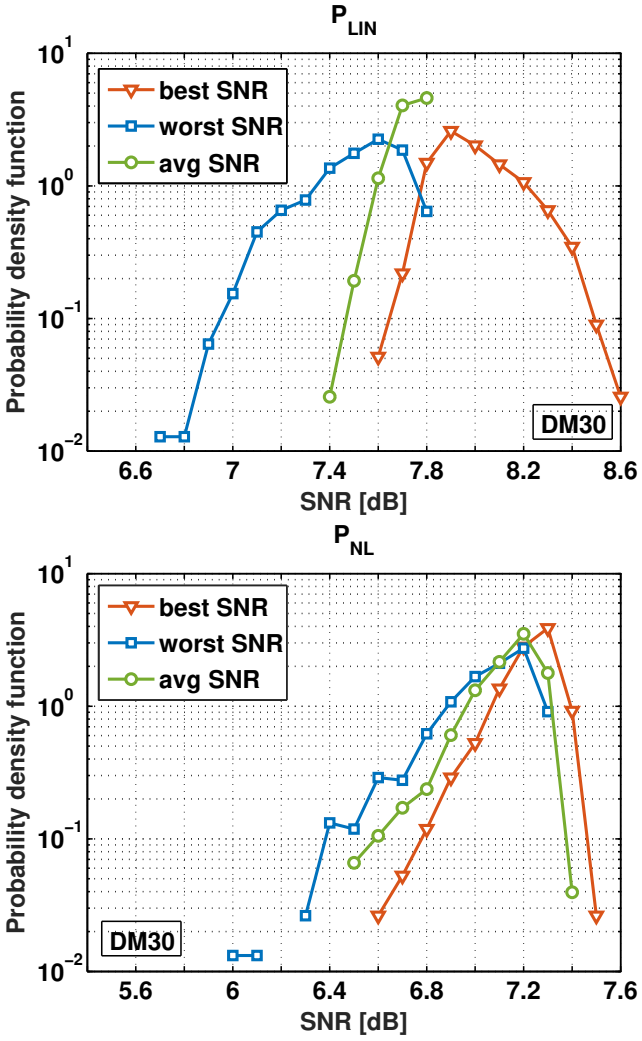

Figure 4.9: Single polarization SNR distributions for best-/worst-performing tributary (triangles/square), as well as the overall SNR (circles), estimated from simulations at (bottom) $P_{\mathrm{NL}} = -4$ dBm and (bottom) $P_{\mathrm{NL}} = 2$ dBm for the DM30 link under investigation. 25×100 km DM30-SMF link. 800 SNR samples.

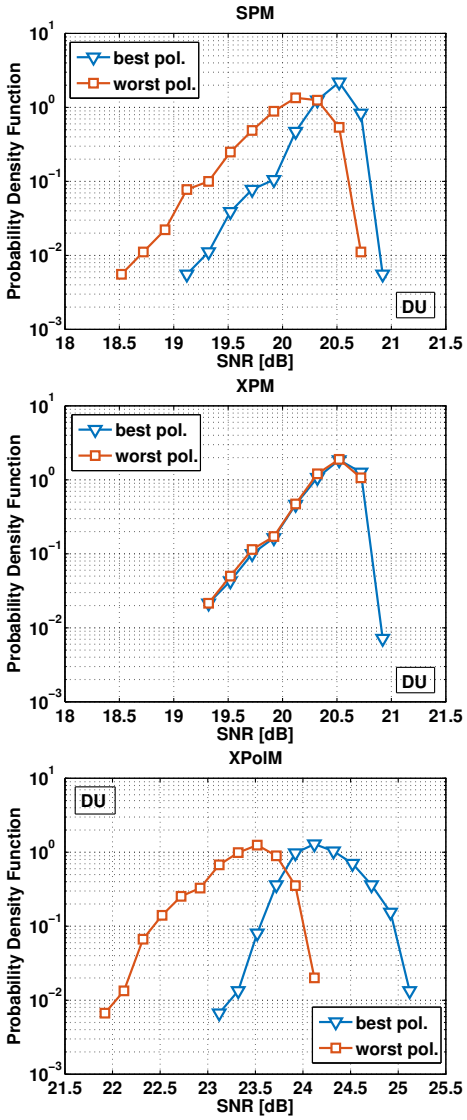

Figure 4.10: PDF of $\mathrm{SNR}_{\mathrm{best}}$ and $\mathrm{SNR}_{\mathrm{worst}}$ obtained by simulating the propagation of the WDM signal described in Sec. 4.3 in a 25×100 km SMF link, in presence of only one nonlinear contribution at a time. Channel power is $P_{\mathrm{NLT}} = 1$ dBm in absence of ASE noise. 800 PDL seeds. a) SPM only, b) XPM only and, c) XPolM only.

been investigated. The single polarization SNR has been analyzed, classified as best- and worst- performing as in Section 4.4.3. The resulting PDF of the SNR are reported in Fig. 4.10. The SNR in the three different plots refers just to the SNR of the single nonlinear contribution, i.e.:

$$\mathrm{SNR} \triangleq \frac{P}{\sigma_i^2} . \qquad i \in [\mathrm{SPM}, \mathrm{XPM}, \mathrm{XPolM}] .$$

For both the SPM and XPM case, the two PDF of $\mathrm{SNR}_{\mathrm{best}}$ and $\mathrm{SNR}_{\mathrm{worst}}$ show a more pronounced tail in the same direction, i.e., toward low SNR values. On the contrary, the PDF with only XPolM shows a similar symmetry already observed in presence of ASE only in linear regime, see Fig. 4.7.

The explanation of the different interaction between PDL and the nonlinear contributions can be ascribed to the different nature of the contributions themselves. In fact, both SPM and XPM are scalar phase rotations, depending on the total instantaneous power of the signal. They thus affect the two polarizations identically. Such a power fluctuates randomly along distance due to PDL, and such an oscillation is reflected identically on the single polarization SNR. XPolM instead is a polarization-dependent matrix term, thus the SNR of the two different polarizations is affected in a different way.

Such a behavior is behind the symmetry inversion of $\mathrm{SNR}_{\mathrm{best}}$ and $\mathrm{SNR}_{\mathrm{worst}}$ found in Section 4.4.2 for increasing power. In a typical DU link the dominant nonlinearity contributions are SPM and XPM [85], as visible by the x-axis of the plots in Fig. 4.10. By increasing the signal power, the nonlinearity becomes more significant in the SNR, thus the variability of SPM and XPM dominates the SNR fluctuations due to PDL.

4.5 Conclusions

We have numerically studied how the PDL affects the SNR of the transmission when fiber nonlinearity is significant. We showed that the interaction between PDL and nonlinearity contributes to the SNR variability by spreading its distribution. Such a spread is already of concern at the nonlinear threshold, and

cannot be neglected if a precise setting of the system margins is needed. Moreover, we investigated the reasons of the spreading of SNR distribution, showing that it is due to a symmetry break in the distributions of the single polarizations SNR for increasing power when such polarizations are classified as best- and worst- performing. Finally, we investigated the variability of nonlinearity by decoupling it in its three main contributions, namely the SPM, XPM, and XPolM, showing that the above-mentioned symmetry break of the SNR distribution can be ascribed to SPM and XPM.

Chapter 5

The extension of the Gaussian noise model to PDL

In this chapter we extend the GN model [14, 15, 17] to account the impact of the PDL in nonlinear regime. In the first section we will introduce the reader to the topic of the nonlinear modeling for a fast and reliable estimation of the system performance. Next, we will present the model theoretically and we will extend it by including the PDL in the theory. Finally, we will validate the extended model both numerically and experimentally.

5.1 Introduction

There is recently a surge of operators' interest in electrical SNR monitoring in multi-vendor optical networks [86]. The SNR of the transmission over each network node can in fact be used as a quality-of-transmission parameter to optimize the network performance [22]. To this goal, such a monitoring should be accurate and fast. In modern optical communications the SNR is strongly impacted by fiber nonlinearity, thus a tool to estimate it is mandatory.

Other important impairments to the SNR are polarization related effects, such as PMD and PDL. They are characterized by a stochastic nature, thus

their impact on the transmission is variable in time. Moreover, they can interact with fiber nonlinearity, changing its statistics randomly during the propagation.

Assessing the impact Kerr nonlinearity on the transmission has been in the recent years an important field of the research in optical communications. To this goal, different models has been found to predict the statistics of nonlinearity [14–19, 21]. The common assumption of all these models is a first-order regular perturbation approximation of the NLI, that we briefly analyzed in Section 2.1.4. Since the NLI is a small additive perturbation, a SNR can be evaluated and possibly converted in BER under the assumption of a received Gaussian-distributed NLI. If such assumption is taken from the beginning of the link, the GN model is derived. The GN model received significant attention due to its simplicity and the good compromise between accuracy and computational effort [87].

Even in the common framework of the first-order perturbation, different approaches has been used to characterize the NLI. For example, [16, 21] use a time domain approach to characterize the nonlinear interference noise by analyzing the collision between different pulses of the signal, whose main feature is to highlight the phase noise nature of the nonlinearity. Other approaches has been studied to characterize the NLI both in the frequency [14, 15, 18] and the spatial [17, 19] domains, with converging results between the two points of view.

The analysis of the GN model started by considering both the NLI and the transmitted signal as a Gaussian process [14, 15, 17], and it has been later improved to include non-Gaussian modulation formats [18, 19], Raman amplification [88] and nonlinear signal-noise interactions [89]. More recently, attempts to also include high-bandwidth stimulated Raman scattering have been made in [62].

So far, the GN model still misses some high-level features as the influence on the NLI of polarization related impairments such as PMD and PDL, as well as the impact of digital signal processing at the receiver. Including such impairments directly in the model will both improve the accuracy of it as well

as allowing fast and reliable investigations of stochastic effects such as PMD and PDL in nonlinear regime.

Recent investigations of the ME in PMD-influenced transmissions over large bandwidths showed a minor impact of the PMD in nonlinear transmissions [46]. However, we showed in Chapter 4 that the interaction between PDL and fiber nonlinearity may have a significant impact on the SNR distribution. Usually the impact of the PDL is modeled by considering only its interaction with ASE [75], thus not considering its interaction with nonlinearity. However, we saw in Charter 4 that nonlinearity reduces the accuracy of such existing models in predicting the SNR margins of the system in presence of PDL. A new model predicting the PDL-nonlinearity interaction is thus extremely useful.

The following work attempts to fill such a gap by extending the GN model for the estimation of the NLI variance to include PDL, and by validating it both numerically and experimentally. The experiments have been performed in Nokia Bell Labs, Paris-Saclay, France in Nov.-Dec. 2017.

5.2 Theoretical model

In this Section we theoretically present the model and its extension to include the PDL and nonlinearity interaction.

5.2.1 GN model theory

In this section we recap some of the most important features of the GN model, as well as highlighting some interesting parallelisms between the frequency domain approach [14, 15] and the spatial approach [17].

We start by reporting the NLI, i.e., the perturbed solution, of an optical PDM signal $U = [U_x, U_y]^T$ in the framework of the ME at the end of a link of length L:

$$U_1(L,t) = -j\gamma\frac{8}{9}\int_0^L e^{-\mathcal{L}\xi}\mathcal{N}\left(e^{\mathcal{L}\xi}U(0,t)\right)\mathrm{d}\xi \tag{5.1}$$

where $e^{\mathcal{L}\xi} \triangleq \mathcal{F}\left\{e^{-j\frac{\beta_2}{2}\omega^2\xi}\right\}$ and $\mathcal{N}(U) \triangleq e^{-g(\xi)}\left(U^\dagger U\right)U$, U the unperturbed solution of the field. $g(\xi)$ is the link gain profile defined in (2.32). Please note that (5.1) refers to a reference system that tracks the GVD of the signal. Such a point of view is equivalent to consider a perfect GVD equalization at the receiver. We are considering a signal with uncorrelated polarizations, i.e.,

$$\boldsymbol{R}_0(\tau) = \mathbb{E}\left[U(0,t)\,U^\dagger(0,t)\right] = R_U\boldsymbol{I} \tag{5.2}$$

with R_U the autocorrelation function of the single polarizations, here considered identical.

If we remove GVD, i.e., $e^{\mathcal{L}\xi} = 1$, the NLI takes the following expression:

$$U_1(L,t) = -j\Phi\,|U(0,t)|^2\,U(0,t) \tag{5.3}$$

where the factor:

$$\Phi \triangleq \frac{8}{9}\gamma\int_0^L e^{\int_0^\xi -\alpha+\sum_k \log(G_k)\delta(x-z_k)\mathrm{d}x} \tag{5.4}$$

is the nonlinear phase accumulated in the propagation.

The autocorrelation of such a nonlinearity is:

$$\begin{aligned}\boldsymbol{R}_{\mathrm{NLI}}(\tau) &= \mathbb{E}\left[U_1(L,t+\tau)\,U_1^\dagger(L,t)\right]\\ &= \Phi^2\mathbb{E}\left[|U(0,t+\tau)|^2\,U(0,t+\tau)\,U^\dagger(0,t)\,|U(0,t)|^2\right]\end{aligned}$$

which involves six-order moments of U. With the assumption of a Gaussian input signal and in absence of linear effects such as PMD, such a correlation can be evaluated using the complex Gaussian moment theorem [90], obtaining [17]:

$$\boldsymbol{R}_{\mathrm{NLI}}(\tau) = 3\Phi^2\,|\boldsymbol{R}_0|^2\,\boldsymbol{R}_0\,. \tag{5.5}$$

It is interesting to note that such an autocorrelation is, in the form, identical to the NLI in (5.3). The nonlinear coefficient γ is present squared in (5.5) through Φ^2, which is not surprising since R_{NLI} has the unit measure of power.

The factor 3 comes from the degeneracy factor of both intra- and inter-channel cross phase modulation from one polarization, plus a single contribution from the other one [17].

We include now the GVD. In this case, the nonlinearity is better described in the frequency domain by a FWM process [14, 15]. Equation (5.1) in the frequency domain is:

$$\tilde{U}_1(L,f) = -j\gamma \iint_{-\infty}^{\infty} \eta(L,f,f_1,f_2)\,\tilde{U}(0,f_1) \times \left[\tilde{U}^{\dagger}(0,f_1+f_2-f)\,\tilde{U}(0,f_2)\right] \mathrm{d}f_1 \mathrm{d}f_2 \mathrm{d}\xi \tag{5.6}$$

with $\eta(L,f,f_1,f_2)$ the kernel of the nonlinearity, which contains all the properties of the link. Equation (2.31) accounts for the nonlinearity from input to output.

The Fourier transform of the correlation function in the frequency domain is the power spectral density (PSD). It is worth noting in absence of PMD and PDL the invariance of the form between the first order perturbation (5.6) and its PSD still holds. In fact, the PSD of the NLI $\tilde{G}_{\mathrm{NLI}}$ is, by Gaussian assumptions of the input signal [15]:

$$\tilde{G}_{\mathrm{NLI}} = \gamma^2 \frac{16}{27} \iint_{-\infty}^{\infty} |\eta(L,f,f_1,f_2)|^2\,\tilde{G}(f_1) \times \left[\tilde{G}(f_1+f_2-f)\,\tilde{G}(f_2)\right] \mathrm{d}f_1 \mathrm{d}f_2\,. \tag{5.7}$$

In this case, the kernel is squared in modulus, thus generalizing the square of gamma mentioned before, and we have $\tilde{G}(f) = \left|\tilde{U}_x(0,f)\right|^2$ the PSD of the single polarization input signal. The factor $\frac{16}{27}$ comes from the PDM assumption, and its origin is explained in [17, App.C]. We want to stress out that the invariance in form of (5.3), (5.5) and (5.6), (5.7) come by the assumption of uncorrelated and identical distributed polarizations, as well as the absence of any kind of effect along the line that can induce correlation between them.

The accumulation of the PDL is distributed along the line. However, the GN model as in (5.6) is a global description of the nonlinearity from input to

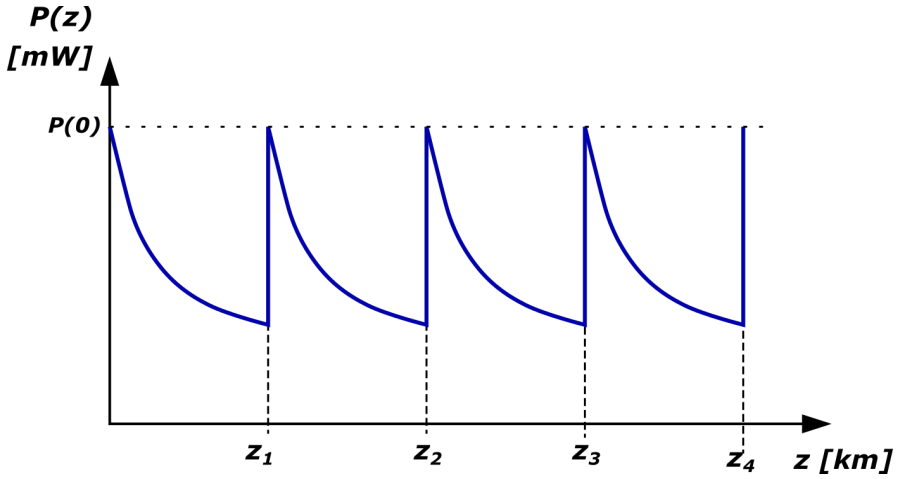

Figure 5.1: Scheme of the signal power profile in a transparent link. $P(0)$ is the input signal power, while z_k are the coordinates of the amplifiers.

output, thus it lacks the local evolution of the signal along the line. This is not a problem in reality since the kernel is a linear function with distance, thus it can be broken into its single span contributions. To this goal, we consider a transparent periodic link as the one depicted in Fig. 5.1, i.e., with equal spans of length L_s where each amplifier completely recover the span loss. In this case the signal input power $P(0) = P(iL_s)\,,\ \ i = 1, ..., N_s$, with N_s the number of spans, while the gain profile (2.32) becomes:

$$g(\xi) = e^{-\alpha(\xi - \lfloor \xi / L_s \rfloor)}$$

with $\lfloor . \rfloor$ the floor function. This way, (2.31) can be separated in the sum of each span contribution as:

$$\eta(L, f, f_1, f_2) = \sum_{n=1}^{N_s} \eta_n(L_s, f, f_1, f_2) \tag{5.8}$$

with the single span kernel of the n^{th} span as:

$$\eta_n(L_s, f, f_1, f_2) \triangleq \eta_1(L_s, f, f_1, f_2)\, e^{-j(2\pi)^2 \beta_2 (f - f_1)(f - f_2)(n-1)L_s} \tag{5.9}$$

The complex exponential in (5.9) accounts for the GVD accumulated by the signal up to span $n-1$. The perturbed solution $\tilde{U}_1(z,f)$ can thus be expressed as:

$$\tilde{U}_1(L,f) = \sum_{n=1}^{N_s} \tilde{U}_{1,n}(L,f) \tag{5.10}$$

with:

$$\begin{aligned}\tilde{U}_{1,n}(L,f) \triangleq -j\gamma \iint_{-\infty}^{\infty} \eta_n(L_s,f,f_1,f_2)\,\tilde{U}(0,f_1) \\ \times \left[\tilde{U}^{\dagger}(0,f_1+f_2-f)\,\tilde{U}(0,f_2)\right] \mathrm{d}f_1 \mathrm{d}f_2\end{aligned}$$

the single span contribution of the NLI. The PSD in this case becomes a double summation, i.e.:

$$\begin{aligned}\tilde{G}_{\mathrm{NLI}} = \gamma^2 \frac{16}{27} \sum_{n,k=1}^{N_s} \iint_{-\infty}^{\infty} \eta_k^*(L_s,f,f_1,f_2)\,\eta_n(L_s,f,f_1,f_2) \\ \times \tilde{G}(f_1)\left[\tilde{G}(f_1+f_2-f)\,\tilde{G}(f_2)\right] \mathrm{d}f_1 \mathrm{d}f_2\,.\end{aligned} \tag{5.11}$$

It is worth noting that for $k=n$ we get the PSD of the nonlinearity generated on the n^{th} span of the link. All the other cases with $k \neq n$ are instead the cross-PSDs describing the correlation of the NLI generated on different span.

We now proceed to include the PDL in the model.

5.2.2 Inclusion of the PDL in the model

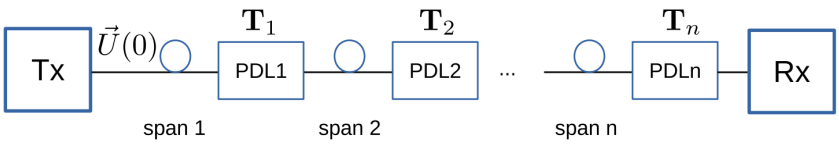

Figure 5.2: Scheme of the transmission link with PDL.

We focus, as in Chapter 4, over a single PDL elements M_i as defined in (4.1) added in the link at the end of each span before amplification. A simple scheme of the resulting link is depicted in Fig. 5.2. This way the perturbed solution U_1 at a generic coordinate s becomes:

$$U_1(L,t) \triangleq -j\gamma \int_0^L \boldsymbol{T}_{n+1,N_s} e^{-\mathcal{L}\xi} \mathcal{N}\left(\boldsymbol{T}_{1n} e^{\mathcal{L}\xi} U(0,t)\right) \mathrm{d}x\,.$$

where T_{1n} is the PDL accumulation matrix defined in (4.2), with $n = \lfloor s/L_s \rfloor$. The special case $\boldsymbol{T}_{10} \equiv \boldsymbol{I}$ is the identity matrix. If the total PDL of the link can be equalized by the receiver, we can change the reference system as $u_1(L,t) \triangleq \boldsymbol{T}_{1N_s}^{-1} U_1(L,t)$ obtaining:

$$u_1(L,t) \triangleq -j\gamma \int_0^L \boldsymbol{T}_{1n}^{-1} e^{-\mathcal{L}\xi} \mathcal{N}\left(\boldsymbol{T}_{1n} e^{\mathcal{L}\xi} U(0,t)\right) \mathrm{d}x \tag{5.12}$$

where $\boldsymbol{T}_{1,n-1}^{-1} = \boldsymbol{T}_{1N_s}^{-1} \boldsymbol{T}_{jN_s}$ comes from the reciprocity of the PDL matrix.

As in Section 5.2.1, we start without considering the GVD. By including the PDL in (5.3), considering the same transparent link of Fig. 5.1 and exploiting the nonlinear operator, the perturbed solution become:

$$u(L,t) = -j\Phi_s \sum_{n=0}^{N_s-1} \boldsymbol{T}_{1n}^{-1} \left(U^\dagger \boldsymbol{P}_{1n} U\right) \boldsymbol{T}_{1n} U \tag{5.13}$$

where:

$$\Phi_s \triangleq \frac{8}{9}\gamma \int_0^{L_s} e^{-\alpha\xi} \mathrm{d}\xi = \frac{8}{9}\gamma L_{\mathrm{eff}}(L_s)$$

is the single nonlinear phase of each identical spans of the link. The matrix:

$$\boldsymbol{P}_{1n} \triangleq \boldsymbol{T}_{1,n}^\dagger \boldsymbol{T}_{1,n}$$

is a positive definite matrix which accounts for the power imbalance due to PDL within the first n spans. Since $\left(U^\dagger \boldsymbol{P}_{1n} U\right)$ in (5.13) is a scalar term, (5.13)

can be further simplified as:

$$u\left(L,t\right)=-j\Phi_{s}\sum_{n=0}^{N_{s}-1}\left(U^{\dagger}\boldsymbol{P}_{1n}U\right)U\,.$$

PDL induces correlation between polarizations. The general case of $\boldsymbol{R}_{\mathrm{NLI}}$ with correlated polarizations is depicted in [17]. By following similar steps as in [17, App.C], the autocorrelation $\boldsymbol{R}_{u}\left(0\right)\triangleq\mathbb{E}\left[u\left(L,t\right)u^{\dagger}\left(L,t\right)\right]$ is:

$$\boldsymbol{R}_{u}=\Phi^{2}\left|R_{U}\right|^{2}R_{U}\sum_{k,n=0}^{N_{s}-1}\left(\mathrm{Tr}\left[\boldsymbol{P}_{in}\boldsymbol{P}_{ik}\right]\boldsymbol{I}+\boldsymbol{P}_{in}\boldsymbol{P}_{ik}\right)\,. \tag{5.14}$$

with Tr matrix trace. We note in (5.14) that the correlation between polarizations given by the PDL breaks the invariance of the form between the electric field and the autocorrelation mentioned previously. However, such a invariance is maintained in the scalar part of (5.14) by $\left|R_{U}\right|^{2}R_{U}$. This should not surprise since any scalar can be factored out by a matrix operation as the PDL. The matrix term:

$$\sum_{k,n=0}^{N_{s}-1}\left(\mathrm{Tr}\left[\boldsymbol{P}_{in}\boldsymbol{P}_{ik}\right]\boldsymbol{I}+\boldsymbol{P}_{in}\boldsymbol{P}_{ik}\right) \tag{5.15}$$

accounts for the correlation between polarizations due to the accumulation of the PDL.

Including the GVD even the scalar part of (5.14) becomes span dependent. However, the GVD is a scalar term as well as R_{U}, thus it can be factored out from (5.15). By similar arguments as the ones used in Section 5.2.1, the PSD of the NLI affected by PDL can be expressed by (5.11) as:

$$\tilde{G}_{u}\left(f\right)\triangleq\sum_{k,n=0}^{N_{s}-1}\left(\mathrm{Tr}\left[\boldsymbol{P}_{in}\boldsymbol{P}_{ik}\right]\boldsymbol{I}+\boldsymbol{P}_{in}\boldsymbol{P}_{ik}\right)S_{kn}\left(f\right) \tag{5.16}$$

with:

$$\begin{aligned}S_{kn}\left(f\right)=&\iint_{-\infty}^{\infty}\eta_{k}^{*}\left(L_{s},f,f_{1},f_{2}\right)\eta_{n}\left(L_{s},f,f_{1},f_{2}\right)\\&\times\tilde{G}\left(f_{1}\right)\left[\tilde{G}\left(f_{1}+f_{2}-f\right)\tilde{G}\left(f_{2}\right)\right]\mathrm{d}f_{1}\mathrm{d}f_{2}\,.\end{aligned}$$

The single realization PDL can thus be modeled as a matrix coefficient (5.15) that weights each cross-PSD among spans, which can be obtained directly by the original GN model. This allows to a fast statistical investigation of the randomness of the SNR by iterating only on the realizations of the PDL, in the order of seconds.

5.3 Model validation

The proposed model has been double-checked with both SSFM simulations and experiments. The results are reported in the next sections.

5.3.1 Experimental validation

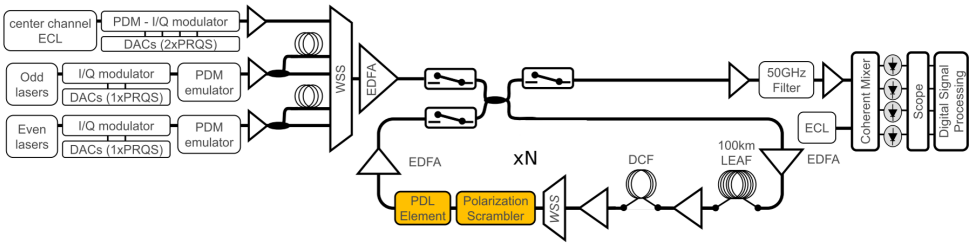

Figure 5.3: Experimental setup. The highlighted blocks are the PDL element and the polarization scrambler used to emulate the random PDL after each span.

The setup for the experiments is reported in Fig 5.3. The optical signal was composed of 13 PDM channels each modulated with PDM-QPSK at symbol rate 32.5 GBaud and evenly spaced of 50 GHz. To emulate the interfering channel we used the common technique to modulate all the even and odd channels with a single modulator, respectively, and then we decorrelated them by using different temporal delays. The channel under test was set to the central one of the comb. The link was composed by 10×100 km spans of LEAF fiber, characterized by a dispersion coefficient $D = 4.3$ ps/nm/km, with

a residual dispersion of 40 ps/nm after each span (DM40), implemented by a recirculating loop. The random PDL was emulated by a fixed PDL element of $\rho = 1$ dB, preceded by a polarization scrambler which randomly rotate the SOP of the signal with respect to the PDL element in each loop [78,81]. The frequency of the polarization scrambler was synchronized with the recirculating loop to ensure the same PDL realization on all the symbols of the transmitted sequence. The acquisition windows of the oscilloscope at the receiver was tuned with it as well. At the receiver side, after recovering the linear impairments and the carrier phase, the SNR of each polarization tributary has been estimated by the variance of the symbols clouds.

To obtain the variance of the NLI, we integrated $\tilde{G}_{\mathrm{NLI}}$ of (5.16) as:

$$\sigma^2_{\mathrm{NLI}} = \int_{-\infty}^{+\infty} |H(f)|^2 \tilde{G}_{\mathrm{NLI}}(f)\,\mathrm{d}f$$

with $H(f)$ the transfer function of the detection filter at frequency f.

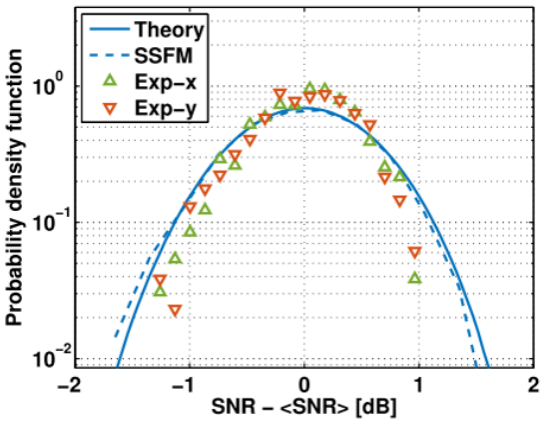

Figure 5.4: Probability density function of the SNR affected by PDL in linear regime. (solid curve) RCM linear model of [75]. (dashed curve) SSFM simulations. (markers) experimental results.

We first check the setup by setting the signal power 4 dB below the non-

linear threshold (estimated in absence of PDL), thus in linear regime, in order to ensure the ASE as the dominant noise. The corresponding PDF of 1500 realizations of the single polarizations SNR is plotted in Fig. 5.4 with markers. We checked it both with accurate SSFM simulations and with the RCM model of [75], described in Section 4.2. Figure 5.4 shows a good match of the experiments with both the model and simulations, thus confirming the correctness of our emulation of the PDL in the experimental setup.

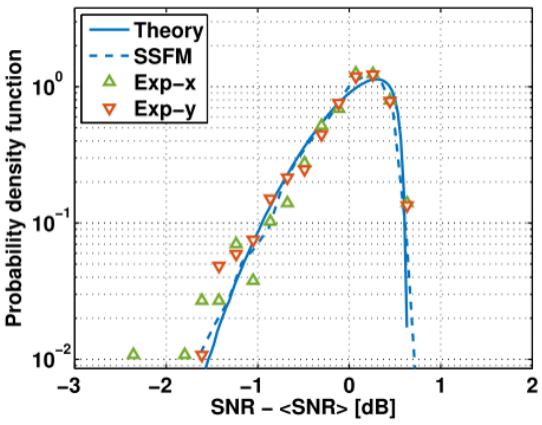

Figure 5.5: Probability density function of the SNR affected by PDL in nonlinear regime. (solid curve) PDL-extended GN model. (dashed curve) SSFM simulations. (markers) experimental results.

We then checked the GN model extension by switching the signal power 4.5 dB above the nonlinear threshold, i.e., with σ^2_{NLI} as dominant noise, by comparing the PDF of the single polarization tributaries of the experiments with our proposed GN model extended to PDL in Fig. 5.5. In this case we applied the standard GN model (5.11) to obtain the NLI power spectral density estimation without the PDL, and then we iterated over the PDL realizations by using (5.15). We observe an excellent match of the new model with both the experimental PDF and SSFM simulations, confirming the good accuracy of our proposal.

5.3.2 Numerical investigation

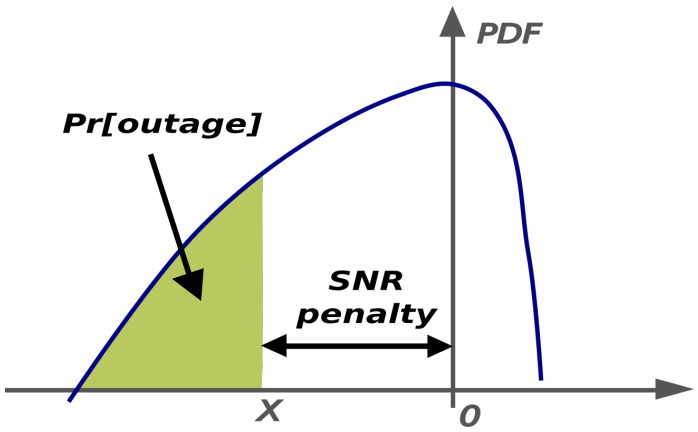

Figure 5.6: Sketch of the SNR penalty of the transmission to obtain a given outage probability.

We investigated the accuracy of our proposed model by numerical SSFM simulations over a DU link. This time the accuracy of the model has been evaluated in terms of SNR penalty, which corresponds to the SNR margin that should be left in the transmission to obtain a given outage probability. Such a probability can be obtained by integrating the PDF of the SNR as:

$$\Pr\left[\text{outage}\right] \triangleq \int_{-\infty}^{X} p_{\text{SNR}}\left(x\right) \mathrm{d}x$$

with $p_{\text{SNR}}\left(x\right)$ the PDF of the SNR obtained by the investigation. The SNR penalty is the distance between the value X and the average SNR (0 in our figures) that gives the corresponding $\Pr\left[\text{outage}\right]$. A sketch of the idea is depicted in Fig. 5.6.

The signal of this investigation was a 21-channel WDM comb with PDM-16QAM modulation, with channel spacing 37.5 GHz. The link was composed of $N_s \times 100$ km of SMF fibers, with N_s going from 1 to 20, with a single

randomly-oriented PDL element at the end of each span. This time the PDL was set to a more realistic value of $\rho = 0.5$ dB.

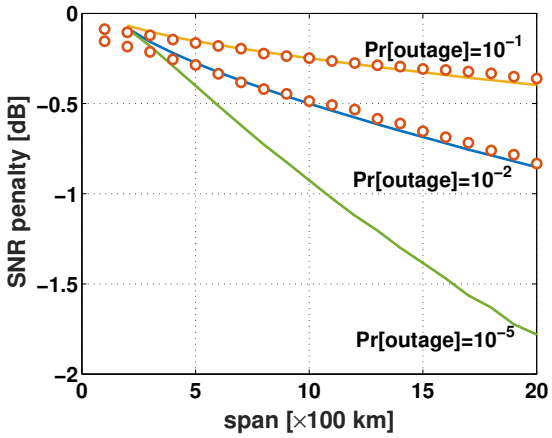

Figure 5.7: SNR penalty [dB] at PDL outage probability of 10^{-1}, 10^{-2} or 10^{-5}. PDL per span: 0.5 dB. Solid lines: GN-model extended to PDL. Symbols: simulations.

The simulations were performed by propagating a different sequence of 28800 symbols for each PDL realization, in order to ensure accurate Monte Carlo estimations. We collected the received sequence after each span and then estimated the SNR by the variance of the received clouds. The PDF of the SNR for both the simulation and the proposed model have been estimated with the histogram method by collecting at least 2000 different PDL realizations. We then evaluated the surrounding area of the left tail of the PDF to obtain the outage probability. The SNR penalty has been the evaluated by interpolation to fixed values of outage probability of 10^{-1}, 10^{-2} and 10^{-5}.

The SNR penalty for the three different value of outage probability is reported in Fig. 5.7 as a function of the number of spans. Solid curves refer to the SNR penalty estimated with the model, while the markers with the simulations. The model shows even in this setup an excellent accuracy with the simulations for outage probability of 10^{-1} and 10^{-2}, confirming the validity of

the proposed theory. The outage probability of 10^{-5} was too low to be estimated by the SSFM in a feasible time due to the insufficient amount of PDL realizations needed to achieve such a probability. The model instead can be used to investigate the system margins even for such a low probability.

It is worth noting the SSFM shows a residual randomness of the SNR even after 1 span, whereas it should be absent since in our setup the PDL is equalized even on the noise in the first span. Such a fluctuation of the SNR is due to the residual Monte Carlo randomness, which cannot be completely avoided in a numerical simulation.

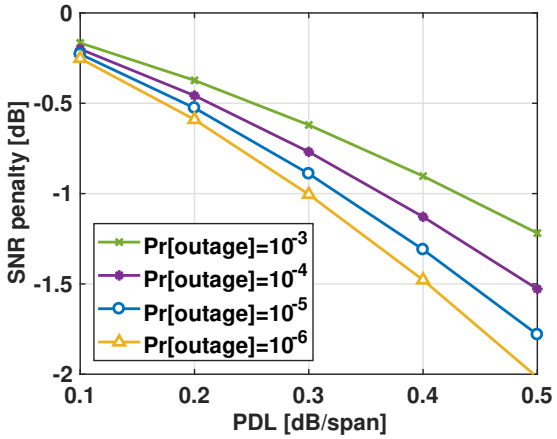

Figure 5.8: SNR penalty at various values of outage probability (see Fig. 5.7) and $N = 20$ spans versus PDL.

Finally, we used our new proposal to investigate the SNR penalty at different values of outage probability ranging from 10^{-3} to 10^{-6} after 20 spans for variable PDL. Such a penalty is reported in Fig. 5.8. The model shows that after 20 spans a small penalty on the SNR around 0.3 dB exists even at the small value of PDL of 0.1 dB.

5.4 Conclusions

We proposed an extension of the GN model already existent in literature to account for the randomness of the SNR due to the PDL-NLI interaction, which is missed by the existing model to estimate the PDL and can be of concern in nowadays transmissions where the Kerr nonlinearity is significant. The proposed model showed an excellent match with both SSFM simulations and experiment. Due to its light computational effort, the proposal can be used for fast statistical estimations of the impact of the PDL on the SNR of the transmission, thus it can be an ideal tool for the SNR monitoring in modern elastic optical networks.

Chapter 6

Conclusions

In this work we addressed two open problems regarding the estimation of the nonlinear interference arising in modern fiber-optic communications, namely the accuracy of the split-step Fourier method (SSFM) and the modeling of the interaction of the polarization dependent loss (PDL) and fiber-optic Kerr nonlinearity. Both problems arise from the need in optical communication to precisely estimate the impact of the Kerr nonlinearity on the signal-to-noise ratio (SNR) of optical transmissions, which is nowadays one of the main limiting factors of the capacity of an optical channel.

The first part of the thesis was about the SSFM algorithm, focusing on the scaling of its accuracy to simulate the propagation of optical signals in transmissions affected by Kerr nonlinearity. We have analyzed the accuracy of this technique by treating the numerical error of the simulation as a perturbation to the signal, thus worsening the SNR of the transmission. Such a study was motivated by the urge of modern optical communications to increase the transmission bandwidth beyond the current 5 THz of the Erbium-doped fiber amplifiers available band, where the accuracy of the SSFM has never been investigated.

By exploiting a parallelism between the SSFM and some known numerical quadrature rules we have detailed the reasons of the numerical error, showing

that it is set by an over-estimation of the four-wave mixing (FWM) along each step of the simulation. Following this analysis, we proposed a new parameter to set the first step of the simulation based on bounding the FWM oscillations on it to grant a constant error on the SNR of the transmission for variations of signal power, bandwidth, and fiber dispersion. In particular, we showed that the first step of the simulation should be power-independent, and it should shrink quadratically with the signal bandwidth. Such a new rule has been compared with two popular rules to set the step of the simulation, namely the nonlinear phase criterion [8] and the constant local error rule [9,10]. We showed that such rules, which are both power-dependent, do not yield a constant SNR error in decibel. In particular, the nonlinear phase criterion showed a decrement of the simulation accuracy for increasing bandwidth, thus it is not reliable to simulate the propagation of wideband signals, while the constant local error rule, due to its power-dependence, shrinks the step of the simulation excessively, resulting in over-accurate simulations at the expense of computational effort. Contrarily, our proposal showed an almost-constant accuracy for signal bandwidths up to 5 THz, thus it can be an extremely useful plug-and-play universal parameter to set up SSFM simulations for ultra-wide bandwidth transmissions.

Moreover, we compared in terms of computational effort the two different step-updating rules, that can be inferred by the nonlinear phase criterion and the constant error rule, both triggered with a first step chosen by our proposal. We showed that, for practical accuracies around 0.01 dB of SNR error, the step-updating rule inferred by the constant local error rule shows a lighter computational effort with respect to the one inferred by the nonlinear phase criterion, while the latter results more convenient when the propagation is simulated with a small number of steps, thus it can be preferable in applications where keeping a light computational effort is the main goal, such as in the framework of digital back-propagation, even at the expense of the accuracy.

In the second part of this work, we analyzed the impact of random PDL in nonlinear optical transmissions, showing that its interaction with nonlinearity changes the statistics of the SNR in a different way than the interaction be-

tween PDL and the ASE noise. We have shown that this different interaction, which is not accounted for by any model to predict the PDL impairments, can significantly vary the statistics of the SNR at the receiver in nowadays transmissions where the variance of the nonlinear interference is comparable to the ASE noise one. Such a result should be accounted for in the set up of the SNR margin to limit the outage of the optical system.

Finally, we extended the Gaussian noise (GN) model for the prediction of the NLI variance to include the nonlinear interaction between signal and PDL. Our extension showed an excellent match both with accurate numerical simulation and experiments and resulted in extremely fast computational time. It is thus an ideal tool to perform fast statistical investigations on the randomness of the SNR affected by the PDL.

Appendix A

Nonlinear phase noise investigation

In this thesis we investigated the SNR of optical links by proposing useful rules to i) setup a numerical SSFM simulation (Chapter 3), ii) understand the impact of XPM and XPolM in presence of PDL (Chapter 4), and iii) estimate the outage probability due to PDL in nonlinear regime (.Chapter 5).

Although the SNR is widely adopted as a key paramenter for system design, the designer may be more interested in other metrics, such as the BER. Under the assumption of circular noise, the conversion between BER and SNR is straightforward. For instance, we have:

$$\text{BER} = \frac{1}{2}\text{erfc}\left(\sqrt{\text{SNR}/2}\right) \qquad \text{QPSK}$$

$$\text{BER} \simeq \frac{2}{\log_2 M}\left(1-\frac{1}{\sqrt{M}}\right)\text{erfc}\left(\sqrt{\frac{3}{2\left(M-1\right)}\text{SNR}}\right) \qquad \text{M-QAM}$$

However, expetially in short links, the NLI may not be circular due to a non-negligible amount of phase noise. Aim of this section is to clarify this aspect and show that after a carrier phase estimation with practical number of taps the circualrity is still a reasonable approximation.

A.1 The problem of nonlinear phase noise

Most of the models of NLI characterization in optical fibers share the common assumption that the NLI generated on the link can be approximated by an additive signal with Gaussian statistics [14, 15, 17–19], and are thus known by the name of Gaussian noise (GN) model. The validity of such an assumption has been confirmed by several simulations and experiments in dispersion uncompensated (DU) links [15, 19, 54–56]. However, These models result inaccurate in particular scenarios, such as links over short distances or dispersion managed (DM) transmissions. One of the reason of such an inaccuracy can be attributed to the strong nonlinear phase noise in these links [94–96], thus the circularity assumption of the GN model can be argued. In particular, in [94] the authors showed that the amount of phase noise on the NLI in WDM transmission strongly depends on channel spacing. As a consequence, while the NLI induced by neighboring channels is a balanced mix of phase and circular noise, the NLI induced by far away channels is mostly phase noise. Such a phase noise could be recovered in the DSP at the receiver, thus not impacting the transmission performance.

The aim of this Appendix is to study more in depth the phase noise nature of the NLI for variable channel spacing by numerically analyze the role of phase noise in the XPM and XPolM contribution to the NLI, and its possible mitigation by using a carrier phase estimation (CPE) at the receiver.

A.2 Phase noise estimation

The received symbol y_k over a channel characterized by a random phase shift and additive circular noise can be generally modeled as :

$$y_k = x_k e^{j\phi_k} + n_k^{\text{circ}} \tag{A.1}$$

where x_k is the transmitted symbol, ϕ_k the phase noise and n_k^{circ} is an additive noise contribution which we assumed circular. It is worth noting such a model

for the received signal is coherent with the nature of the nonlinearity since XPM alone is a pure phase rotation, as depicted in Section 2.1.2, while the XPolM induces a distributed interference characterized by random paths over the Poincaré sphere. Hence, XPolM is expected to induce equally distributed distortions between the real and imaginary part components.

Under the perturbative assumption of the nonlinearity described in Section 2.1.4, the exponential in (A.1) can be approximated by its Taylor series truncated to the first order, and the received signal can be expressed as [93]:

$$y_k = x_k + jx_k\phi_k + n_k^{\text{circ}},$$

i.e., the phase shift can be modeled as an additive noise, generally complex due to the dependence on the transmitted symbol. To isolate the phase noise ϕ_k from n_k^{circ} we first extract the In-phase and Quadrature components (I/Q) of the received signal by using the decomposition described in [95,96]. We thus considered:

$$\hat{r}_k \triangleq y_k e^{-\angle x_k} - |x_k|$$

where the received symbol is first de-rotated of the transmitted symbol phase $\angle x_k$ and then we subtracted $|x_k|$ to obtain the zero-mean process $\hat{r}$. The real and imaginary parts of $\hat{r}_k$ are the I/Q components of the noise, with variances σ_I^2 and σ_Q^2. It is worth noting that such a point of view is different from the GN model [15,19], whose reference system is the same as the one of the transmitted signal constellation.

Since we assumed n_k^{circ} as circular, its contribution to σ_I^2 and σ_Q^2 is identical since the phase rotation $e^{-\angle x_k}$ does not change its statistics. By these considerations and by assuming independence between ϕ_k and x_k we can express:

$$\sigma_I^2 = \frac{\sigma_C^2}{2} \qquad \sigma_Q^2 = \mathbb{E}\left[|x_k|^2\right]\sigma_\phi^2 + \frac{\sigma_C^2}{2}$$

where σ_C^2 and σ_ϕ^2 are the variances of the circular and phase noise, respectively, which can be thus easily estimated from σ_I^2 and σ_Q^2.

A.3 Simulation setup

In this work we focus on the nonlinearity generated by cross-channel effects, i.e., XPM and XPolM (see (2.25)). We thus considered a two-channel system and studied the nonlinear impairments on a reference channel, kept fixed in frequency, due to an interfering one at variable channel spacing. Each channel was modulated by PDM-16QAM signals with root raised cosine supporting pulses of roll-off $r = 0.1$ at symbol rate of $R = 32$ GBaud. The power per-channel was $P = 0$ dBm. The channel spacing Δf between the two channels has been varied from 37.5 to 975 GHz by steps of 37.5 GHz. We overlooked the laser phase noise to focus only on the phase noise produced by nonlinearity. The link was composed of 5×100 km SMF spans ($\alpha = 0.2$ dB/km, $D = 17$ ps/nm/km, $\gamma = 1.3$ mW^{-1}·km^{-1}) dispersion uncompensated, with a perfect equalization of the GVD at the receiver. PMD was not present. In order to focus on the nonlinearity we considered at the end of each span a noiseless EDFA that perfectly compensated the span loss.

The receiver was composed of a matched filter followed by a fractionally spaced data-aided least squared equalizer with 5 taps to compensate any residual linear distortion as well as recovering the average nonlinear phase rotation of the signal. When CPE is present, we considered a blind phase search algorithm [97] (BPS) with a number of taps optimized case by case. It is worth noting that, differently from [97], we used a total of number of test phases equal to 3000 to keep the quantization error negligible compared to the NLI.

SSFM simulations were conducted by solving the ME by separation of the nonlinear effects as in (2.25) and switching "on" only the desired contributions. In order to obtain reliable statistics of the NLI we averaged the results of a total of 10 different propagations case by case, each one corresponding to a transmission of 4096 symbols discretized with 64 samples per symbol. Note that such values respect the suggestions on the signal discretization reported in Appendix 3.6.5.

After signal detection we extracted the NLI I/Q components using the

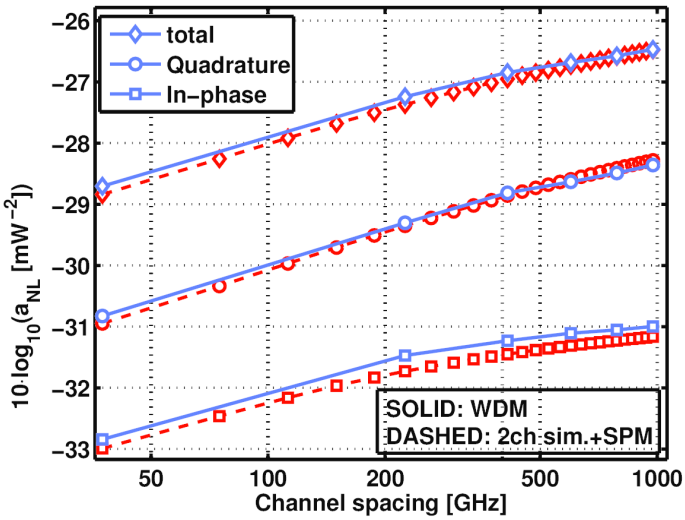

Figure A.1: I/Q contributions to the NLI variance (a_{NL}) of a WDM simulation (solid) compared with the cumulative sum of the two-channel simulations + SPM (dashed).

procedure described in Section A.2 and we estimated the normalized NLI variance [53] a_{NL} as defined in (3.5) for each component.

A.4 Impact of the phase noise on the NLI

We first check the importance of FWM in the considered case of study by simulating WDM transmissions for increasing number of channels and using it as a benchmark for our two-channel simulations. We estimated the a_{NL} of the WDM transmissions and compared it with the cumulative sum of the a_{NL} of the two-channel simulations up to the same bandwidth of the WDM. To include the SPM in the resulting variance we added the variance of a single channel propagation. Figure A.1 shows the above-mentioned comparison for each I/Q component and the total a_{NL}, defined as the sum of the I and Q

variances. The good match of the curves in Fig. A.1 indicates that the FWM in this case of study is negligible, thus validating the approach based on the separation of nonlinear contributions.

We now focus on just the two-channel cases. Figure A.2 shows the a_{NL} of each NLI contribution as a function of the channel spacing for the XPM (Fig. A.2 left) and XPolM (Fig. A.2 right) alone. We can see in both figures a gap between the In-phase and Quadrature components (solid squares and circles) increasing for increasing channel spacing. Such a gap indicates that phase noise is present, inflating the quadrature component of the NLI. The gap is maximized with only the XPM, reaching a difference of almost 10 dB for a channel spacing of 975 GHz, coherently with the local phase nature of the XPM for big walk-off between channels, as reported in (2.20). Such a gap between the I/Q components is smaller in the XPolM case, reaching a maximum gap of almost 5 dB at 975 GHz. The reason for this smaller gap is that, contrarily to XPM which is a unitary scalar operation, XPolM is a unitary matrix operation affecting both the phase and the amplitude of the signal, thus inflating the In-phase component of the NLI.

For the sake of comparison we also plotted the normalized variances of the real and imaginary parts of the NLI, labeled "Real" and "Imaginary" in the reference system of the signal constellation, which is the point of view shared by the GN model. Not surprisingly, the two variances in this reference system essentially coincide, confirming that the GN model is unaware of the phase noise nature of the NLI.

Figure A.3 shows the same curves of the NLI components but with both XPM and XPolM acting together, which is the more realistic case of a real transmission. We note that since the two components are comparable, a phase noise still exist even in this case.

It is worth noting that in all the three considered cases (XPM, XPolM, XPM+XPolM) the variance of the real and imaginary part of the NLI, i.e., the one referring to the GN model point of view, are exactly 3 dB lower than the total I+Q variance, i.e., the total energy of the NLI is identical in both points

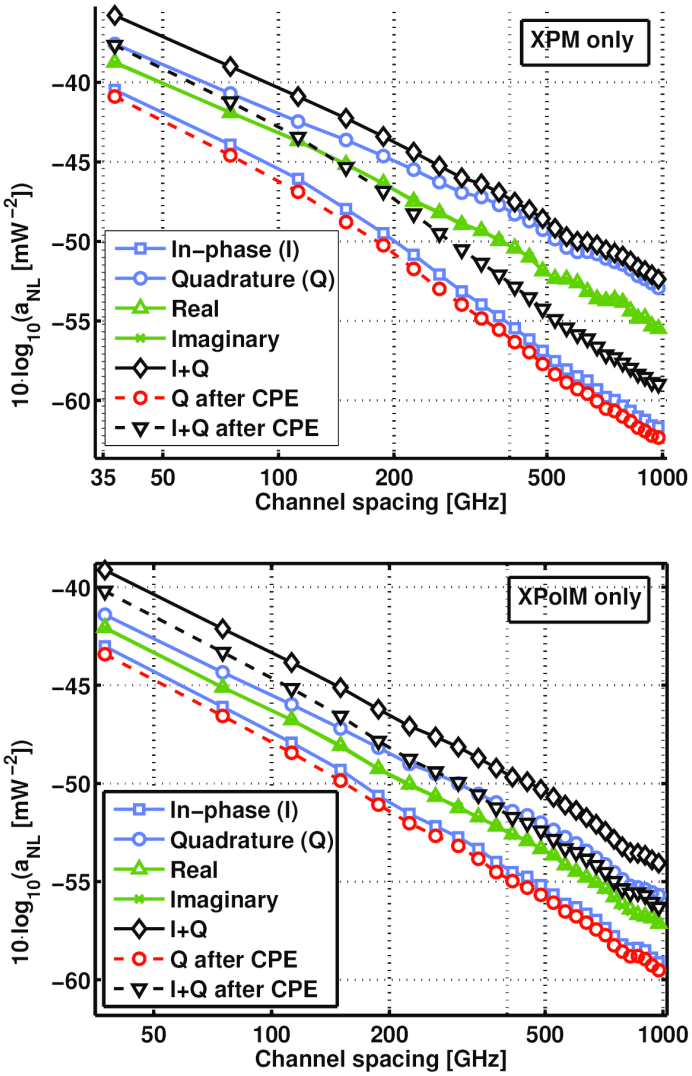

Figure A.2: a_{NL} contributions versus channel spacing in different reference systems (see Section 2). "I" and "Q": In-phase /Quadrature components; "Real", "Imaginary": components of the GN model reference system; "I+Q": overall NLI normalized variance. Dashed curves: variances after phase estimation. a) XPM only and b) XPolM only. Two-channel simulations.

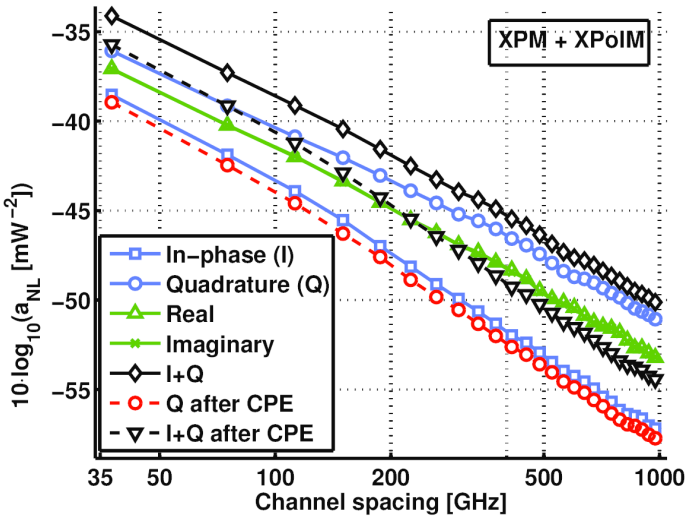

Figure A.3: a_{NL} contributions versus channel spacing in different reference systems (see Section 2). "I" and "Q": In-phase /Quadrature components; "Real", "Imaginary": components of the GN model reference system; "I+Q": overall NLI normalized variance. Dashed curves: variances after phase estimation. a) XPM only and b) XPolM only. Two-channel simulations.

of view.

The phase noise we observed in Fig. A.2 and Fig. A.3 could be mitigated at the receiver by using a CPE. To study the impact of such a mitigation we considered a BPS algorithm [97] placed after the receiver. The beneficial gains of the CPE are highlighted in Fig. A.2 and A.3 by the dashed curves. We can see that in all cases the optimized CPE reduces the variance of the quadrature component to a level slightly lower than the in-phase one for all the three cases of XPM, XPolM and XPM+XPolM (dashed circles). This observation suggests that the phase noise has been completely removed by the CPE, with a maximum gain in the total NLI variance in the XPM+XPolM case of almost 5 dB for a channel spacing of 975 GHz. Please note that the total variance after the CPE is different from the one shared by the GN model, which is unaffected by the CPE due to the circularity assumption of the noise. Such results are coherent with the ones found in [20, Fig. 20].

A.5 Spectral analysis of the phase noise

To better understand the nature of the NLI we performed a spectral analysis of the in-phase and quadrature components. To this goal, we tested 50 random seeds of transmitted sequences and estimated the power spectral density (PSD) by using the Welch method. We focused on the two extreme scenarios of Fig. A.3, i.e., a channel spacing of 37.5 GHz and 975 GHz for the case of XPM+XPolM. Such PSDs are reported in Fig. A.4 in terms of normalized frequency to the symbol rate. From the figure we note that the GN model point of view presents PSDs of the real and imaginary components almost identical, confirming the circularity of the noise in such a point of view. Moreover, the flatness of the PSDs suggest that such a noise is essentially white since the NLI term $jx_k\phi_k$ is independent symbol by symbol. The I and Q components, instead, show a non-flat spectrum, with a depression and an inflation on the lower frequencies of the spectrum. Such a behavior can be ascribed to the phase noise, which shows a correlation on the quadrature component due to

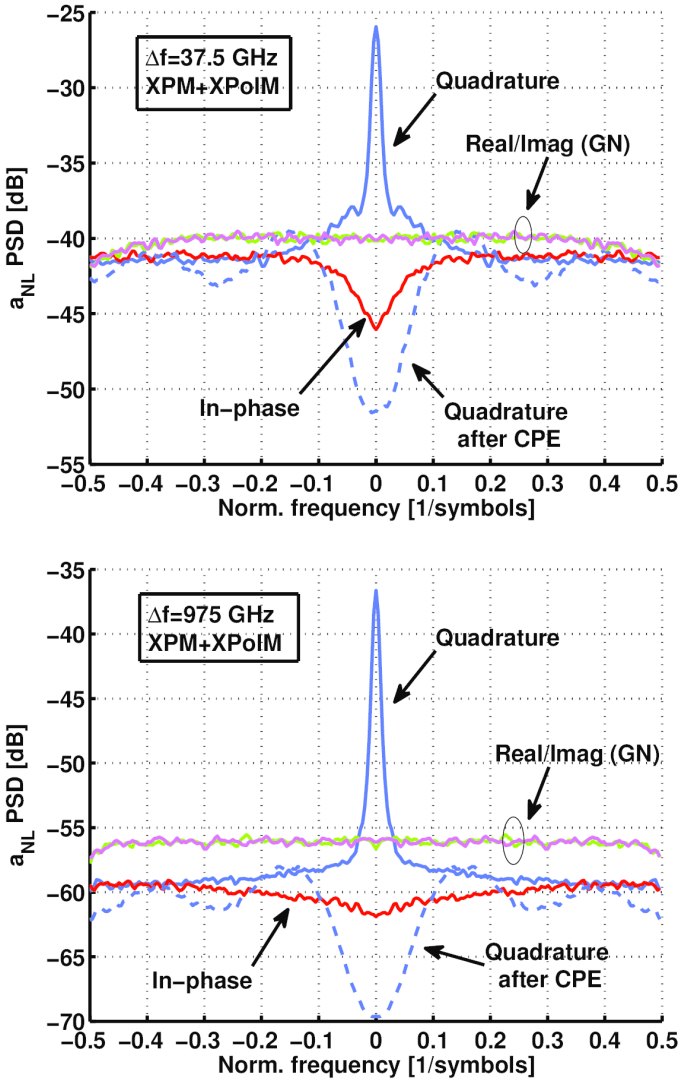

Figure A.4: PSD of normalized NLI components for a) $\Delta f = 37.5$ GHz and b) $\Delta f = 975$ GHz with both XPM and XPolM, in a frequency normalized to symbol rate. Solid lines: without CPE; dashed line: with CPE.

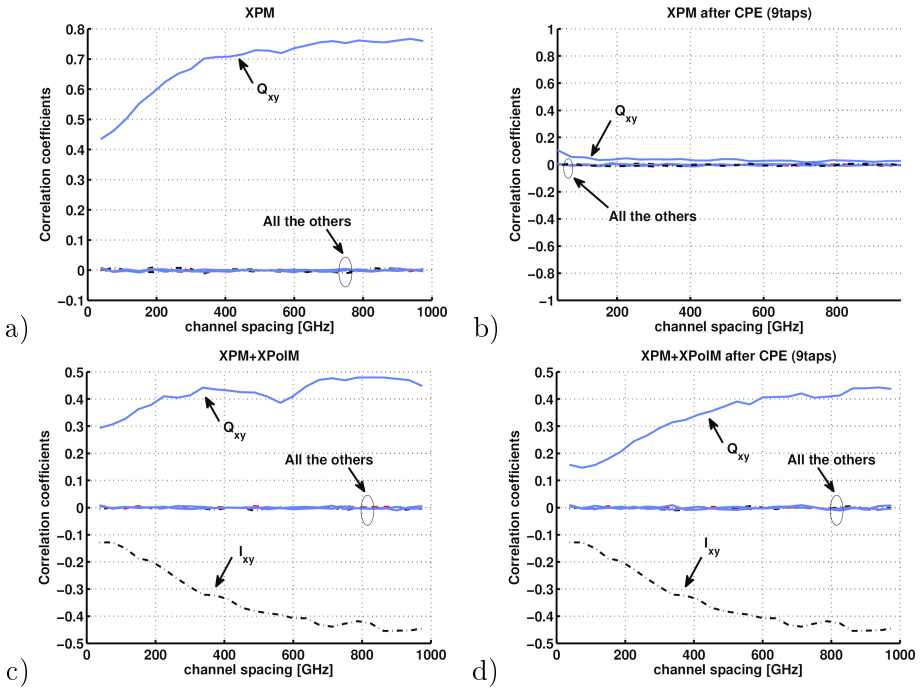

Figure A.5: Cross-correlation coefficients between I and Q components of X and Y polarizations. First row: XPM; Last row: XPM+XPolM. First column: before CPE; last column: after CPE. I_{xy} (Q_{xy}): correlation between in-phase (quadrature) components of X and Y polarizations. Two-channel simulations.

the walk-off between the two channels.

The dashed curve refers to the PSD of the quadrature component after the CPE with an optimized number of taps. In this case the CPE removed the phase noise from the quadrature PSD, which is now similar to the in-phase PSD. This suggests the optimal number of taps for the CPE is the one that makes its bandwidth closest to the bandwidth of the phase noise. In fact, after removing the slow frequency components due to the correlated phase noise, the CPE cannot cope with the flat high frequencies of the PSD, which act as a white noise leading to unrecoverable symbol errors and cycle slips.

Finally, we studied the correlation between the I and Q components of different polarizations. Figure A.5 shows the correlation coefficients of every possible combination of I/Q components of the two X and Y polarizations of the PDM signal as a function of the channel spacing, for the cases of XPM only and XPM+XPolM (top and bottom rows, respectively), before and after the CPE (left and right columns, respectively). From Fig. A.5a) we can observe that the XPM only case presents a strong correlation between the quadrature components of the two polarizations before the CPE, which grow stronger for increasing channel spacing. Such a correlation is ascribed to the strong phase noise composing the XPM, which is a scalar operation thus affecting both polarizations. As we showed previously such a phase noise can be efficiently removed by the CPE, thus removing the correlation between polarizations as showed in Fig. A.5b). By adding XPolM the results change as shown in Fig.3c-d). We observe a smaller correlation between the Q component of X and Y polarization, mainly due to the differential phase induced by XPolM. The CPE now is unable to remove such a correlation since it is working in presence of relevant polarization crosstalk, which behaves as a white noise, as visible in Fig. A.5d). We can also observe in this case a negative correlation induced by the XPolM in the in-phase components of the two polarization. This correlation is left intact by the CPE since it operates only on the phase, i.e., the quadrature component, of the NLI.

Bibliography

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

[2] E. Ip, and J. M. Kahn, "Digital equalization of chromatic dispersion and polarization mode dispersion," *J. Lightw. Technol.*, vol. 25, no. 8, pp. 2033–2043, 2007.

[3] G. Goldfarb, and G. Li, "Chromatic dispersion compensation using digital IIR filtering with coherent detection," *IEEE Photon. Technol. Lett., vol.* 19, no. 13, pp. 969–971, 2007.

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

[5] 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.

[6] C. Francia, "Constant Step-Size Analysis in Numerical Simulation for Correct Four-Wave-Mixing Power Evaluation in Optical Fiber Transmission Systems", *IEEE Photon. Technol. Lett.*, vol. 11, no. 1, pp. 69–71, Jan. 1999.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] G. M. Muslu and H. A. Erbay, "Higher-order split-step Fourier schemes for the generalized nonlinear Schrödinger equation," *Math. Comput. Simul.*, vol. 67, no. 6, pp. 581–595, Jan. 2005.

[13] X. Chen and W. Shieh, "Closed-form expressions for nonlinear transmission performance of densely spaced coherent optical OFDM systems," *Opt. Express*, vol. 18, pp. 19039–19054, 2010.

[14] A. Carena, V. Curri, G. Bosco, P. Poggiolini, and F. Forghieri, "Modeling of the impact of non-linear propagation effects in uncompensated optical coherent transmission links," *J. Lightw. Technol.*, vol. 30, no. 10, pp. 1524–1539, May 2012.

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

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

[17] P. Serena and A. Bononi, "An Alternative Approach to the Gaussian Noise Model and its System Implications," *J. Lightw. Technol.*, vol. 31, no. 21, pp 3489–3499, Nov. 2013.

[18] 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. 16335-16362, 2014.

[19] P. Serena and A. Bononi, "A time-domain extended Gaussian noise model," *J. Lightw. Technol.*, vol. 33, no. 7, pp. 1459–1472, Apr. 2015.

[20] P. Poggiolini and Y. Jiang, "Recent Advances in the Modeling of the Impact of Non-Linear Fiber Propagation Effects on Uncompensated Coherent Transmission Systems," *J. Lightw. Technol.*, vol. 35, no. 3, pp. 458–480, 2016.

[21] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Pulse Collision Picture of Inter-Channel Nonlinear Interference in Fiber-Optic Communications," *J. Lightw. Technol.,* vol. 34, no. 2, pp. 593–607, Jan. 2016.

[22] V. Curri, M. Cantono, and R. Gaudino, "Elastic All-Optical Networks: A New Paradigm Enabled by the Physical Layer. How to Optimize Network Performances?," *J. Lightw. Technol.*, vol. 35, no. 6, pp. 1211-1221, Mar. 2017.

[23] A. Ghazisaeidi *et al.*, "Advanced C + L-band transoceanic transmission systems based on probabilistically shaped PDM-64QAM," *J. Lightw. Technol.,* vol. 35, no. 7, pp. 1291–1299, Apr. 2017.

[24] J.-X. Cai *et al.*, "51.5 Tb/s capacity over 17,107 km in C + L bandwidth using single mode fibers and nonlinearity compensation," *J. Lightw. Technol.,* vol. 36, no. 11, pp. 2135–2141, Jun. 2018.

[25] G. P. Agrawal, *Nonlinear Fiber Optics*, 5nd ed. New York: Academic, 2013.

[26] https://it.m.wikipedia.org/wiki/File:Optical_fiber_transmission_windows.svg

[27] P. P. Mitra and J. B. Stark, "Nonlinear limits to the information capacity of optical fiber communications," *Nature*, vol. 411, no. 6841, pp. 1027–1030, 2001.

[28] R.-J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *J. Lightw. Technol.*, vol. 28, no. 4, pp. 662–701, 2010.

[29] A. D. Ellis, J. Zhao, and D. Cotter, "Approaching the non- linear Shannon limit," J. *Lightw. Technol.*, vol. 28, no. 4, pp. 423–433, 2010.

[30] A. R. Chraplyvy, "Limitations on lightwave communications imposed by optical-fiber nonlinearities," *J. Lightw. Technol.*, vol. 8, no. 10, pp. 1548–1557, Oct. 1990.

[31] E. Agrell, "Conditions for a monotonic channel capacity," *IEEE Trans. Commun.*, vol. 63, no. 3, pp. 738–748, Mar. 2015

[32] T. R. Taha and M. J. Ablowitz, "Analytical and numerical aspects of certain nonlinear evolution equations. II. Numerical, nonlinear Schrodinger equation." *J. Comp. Phys.*, vol. 55, pp. 203-230, 1984.

[33] S. Musetti, P. Serena, and A. Bononi, "On the Accuracy of Split-Step Fourier Simulations for Wideband Nonlinear Optical Communications," accepted for publication on *J. Lightw. Technol., DOI: 10.1109/JLT.2018.2877384.*

[34] R. W. Tkach, A. R. Chraplyvy, Fabrizio Forghieri, A. H. Gnauck, and R. M. Derosier, "Four-Photon Mixing and High-speed WDM Systems," *J. Lightw. Technol.*, vol. 13, no. 5, May 1995.

[35] M. J. Ablowitz, T. Hirooka, and T. Inoue, "Higher-order asymptotic analysis of dispersion-managed transmission systems: Solutions and their characteristics," *J. Opt. Soc. Amer. B*, vol. 19, no. 12, pp. 2876–2885, 2002

[36] N. Rossi, P. Serena and A. Bononi, "Symbol-Rate Dependence of Dominant Nonlinearity and Reach in Coherent WDM Links," *J. Lightw. Technol.*, vol. 33, no. 14, Jul. 2015.

[37] C. D. Poole, R. E. Wagner, "Phenomenological approach to polarization dispersion in long single-mode fibers," *Electronics Letters*, vol. 22, no. 19, Sep. 1986.

[38] C. D. Poole, "Statistical treatment of polarization dispersion in single-mode fiber," *Opt. Lett.*, vol. 13, pp. 687–689, 1988.

[39] C. D. Poole, "Measurement of polarization-mode dispersion in single-mode fibers with random mode coupling," *Opt. Lett.*, vol. 14, pp. 523–525, 1989.

[40] G. J. Foschini and C. D. Poole, "Statistical Theory of Polarization Dispersion in Single Mode Fibers," *J. Light. Technol.*, vol. 9, no. 11, pp. 1439–1456, 1991.

[41] S. V. Manakov, "On the theory of two-dimensional stationary self-focusing of electromagnetic waves," *Zh. Eksp. Tecr. Fiz.*, vol. 65, pp. 505-516, 1973.

[42] 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.

[43] P. K. A. Wai, C. R. Menyuk, and H. H. Chen, "Stability of solitons in randomly varying birefringent fibers," *Opt. Lett.,* vol. 16, pp. 1231–1233, 1991.

[44] S. G. Evangelides Jr., L. F. Mollenauer, J. P. Gordon, and N. S. Bergano, "Polarization multiplexing with solitons," *J. Lightwave Technol*, vol. 10, pp. 28–35, 1992.

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

[46] M. Cantono, D. Pilori, A. Ferrari, A. Carena, and V. Curri, "Observing the Interaction of PMD with Generation of NLI in Uncompensated Amplified Optical Links," in Proc. *OFC 18,* San Diego, CA, USA, paper W1G.4, 2018.

[47] M. Winter, C. A. Bunge, D. Setti, and K. Petermann, "A Statistical Treatment of Cross-Polarization Modulation in DWDM Systems," *J. Lightw. Technol.*, vol. 27, no. 17, pp. 3739–3751, Sep. 2009.

[48] J. P. Gordon and H. Kogelnik, "PMD fundamentals: polarization mode dispersion in optical fibers," PNAS, vol. 97, no. 9, pp. 4541–4550, Apr. 2000.

[49] D. Zwillinger, Handbook of Differential Equations. Boston: Academic, 1989, pp. 528–531.

[50] X. Wei, "Power-weighted dispersion distribution function for characterizing nonlinear properties of long-haul optical transmission links," *Opt. Lett.*, vol. 31, no. 17, pp. 2544–2546, 2006.

[51] A. Vannucci, P. Serena, and A. Bononi, "The RP method: A new tool for the iterative solution of the nonlinear Schroedinger equation," *J. Lightw. Technol.*, vol. 20, no. 7, pp. 1102–1112, 2002.

[52] E. Grellier and A. Bononi, "Quality parameter for coherent transmissions with Gaussian-distributed nonlinear noise," *Opt. Express*, vol. 19, no. 13, pp. 12781-12788, 2011.

[53] 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, 2012.

[54] E. Torrengo, R. Cigliutti, G. Bosco, A. Carena, V. Curri, P. Poggiolini, A. Nespola, D. Zeolla, and F. Forghieri, "Experimental validation of an analytical model for nonlinear propagation in uncompensated optical links," *Opt. Express*, vol. 19, no. 26, pp. B790-B798, Dec. 2011.

[55] 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, 2012

[56] A. Nespola, S. Straullu and A. Carena, "GN-Model Validation Over Seven Fiber Types in Uncompensated P-16QAM Nyquist-WDM Links," *IEEE Photon. Technol. Lett.*, vol. 26, no. 2, pp. 206-209, 2014.

[57] G. H. Weiss and A. A. Maradudin, "The Baker–Hausdorff formula and a problem in crystal physics," *J. Math. Phys.*, vol. 3, no. 4, pp. 771–777, Jul. 1962.

[58] A. Iserles and S. P. Nørsett, "Efficient quadrature of highly oscillatory integrals using derivatives," Proc. *R. Soc. A Math. Phys. Eng. Sci.*, vol. 461, no. 2057, pp. 1383–1399, May 2005.

[59] M. Frigo and S. G. Johnson, "The Design and Implementation of FFTW3," in *Proceedings of the IEEE*, vol. 93 no. 2, pp 216–231, 2005.

[60] J. H. Mathews and K. D. Fink, *Numerical methods using MATLAB.* Prentice Hall, Inc., 1999.

[61] O. Rival and K. Mheidly, "Accumulation rate of inter and intra-channel nonlinear distortions in uncompensated 100G PDM-QPSK systems," in Proc. *OFC 12*, Los Angeles, CA, paper JW2A.52, 2012.

[62] D. Semrau, R. I. Killey and P. Bayvel, "The Gaussian Noise Model in the Presence of Inter-Channel Stimulated Raman Scattering," *J. Lightw. Technol.*, vol. 36, no. 14, Jul. 2018.

[63] K. Atkinson, "Numerical Solution Of Ordinary Differential Equations", Wiley-Interscience, pp. 96-103.

[64] D. T. Neilson, C. R. Doerr, D. M. Marom, R. Ryf, and M. P. Earnshaw, "Wavelength selective switching for optical bandwidth management," *Bell Labs Technical Journal,* vol. 11,no. 2, Summer 2006.

[65] A. Mecozzi, M. Shtaif, "The statistics of polarization-dependent loss in optical communication systems," *IEEE Photon. Technol. Letters*, vol. 14, no. 3, pp. 313–315, Mar. 2002.

[66] A. Galtarossa, and L. Palmieri, "The Exact Statistics of Polarization-Dependent Loss in Fiber-Optic Links," *IEEE Photon. Technol. Lett.*, vol. 15, no. 1, pp. 57-59, Jan. 2003.

[67] L. E. Nelson, M. Birk, P. Magill, A. Schex, and L. Rapp, "Measurements of the polarization dependent loss of multiple WDM channels in an installed, long-haul terrestrial link," in *IEEE Photonics Society Summer Topical 2010*, 19 – 21 July 2010, Playa del Carmen, Mexico, paper MA3.4.

[68] L. E. Nelson, C. Antonelli, A. Mecozzi, M. Birk, P. Magill, A. Schex, and L. Rapp, "Statistics of polarization dependent loss in an installed long-haul WDM system," *Optics Express*, vol. 19, no. 7, pp. 6790-6796, 2011.

[69] 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, 2004.].

[70] T. Duthel, C. R. S. Fludger, J. Geyer, and C. Schulien, "Impact of polarisation dependent loss on coherent POLMUX-NRZ-DQPSK," in *Proc. OFC/NFOEC*, San Diego, CA, 2008, Paper OThU5.

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

[72] 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.

[73] C. Xie, "Polarization-Dependent Loss-Induced Outage Probabilities in Optical Communication Systems," *IEEE Photon. Technol. Lett.*, vol. 20, no. 13, pp. 1091–1093, Jul. 2008.

[74] C. Xie, "Polarization-dependent loss induced penalties in PDM-QPSK coherent optical communication systems," in *Proc. OFC/NFOEC*, San Diego, CA, USA, Mar. 2010, pp. 1–3, paper OWE6.

[75] 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. Letters*, vol. 21, no. 24, pp. 1882–1884, Dec. 2009.

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

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

[78] O. Vassilieva, *et al.*, "Experimental investigation of the statistics of the interplay between nonlinear and PDL effects in coherent polarization multiplexed systems," in *Proc. OFC/NFOEC*, Ananheim, CA, USA, 2013, pp 1–3, paper OM3B.6.

[79] N. Rossi, P. Serena, and A. Bononi, "Polarization-Dependent Loss Impact on Coherent Optical Systems in Presence of Fiber Nonlinearity," *IEEE Photon. Technol. Lett.*, vol. 26, no. 4, pp. 334-337, Feb. 2014.

[80] 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. 4905-4911, Dec. 2014.

[81] 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, Feb. 2017.

[82] A. Andrusier, E. Meron, M. Feder, and M. Shtaif, "Optical implementation of a space–time-trellis code for enhancing the tolerance of systems to polarization-dependent loss," *Opt. Lett.*, vol. 38, no. 2, pp. 118-120, 2013.

[83] Z. Wang, C. Xie, and X. Ren, "PMD and PDL impairments in polarization division multiplexing signals with direct detection," *Opt. Express*, vol. 17, no. 10, pp. 7993-8004, 2009.

[84] N. Rossi, A. Ghazisaeidi, and P. Ramantanis, "Stochastic Nonlinear Interference in Dispersion Managed Coherent Optical Links," in *Proc. ECOC 2016*, p. 1196, Dusseldorf.

[85] A. Bononi, P. Serena, N. Rossi, D. Sperti, "Which is the Dominant Nonlinearity in Long-haul PDM-QPSK Coherent Transmissions?," in *Proc. ECOC 2010*, Torino, Italy, Sep. 2010, pp. 1–3, paper Th.10.E.1.

[86] M. Filer, J. Gaudette, M. Ghobadi, R. Mahajan, T. Issenhuth, B. Klinkers, and J. Cox, "Elastic Optical Networking in the Microsoft Cloud," *J. Opt. Commun. Netw.*, vol. 8, no. 7, p. A45, 2016.

[87] B. D. Taylor et al., "Towards a Route Planning Tool for Open Optical Networks in the Telecom Infrastructure Project," in *Proc. OFC 2018*, San Diego, CA, USA, 2013, pp 1–3, paper Tu3E.4.

[88] V. Curri, A. Carena, P. Poggiolini, G. Bosco, and F. Forghieri. "Extension and validation of the GN model for non-linear interference to uncompensated links using Raman amplification," *Opt. Express*, vol. 21, no. 3, pp. 3308-3317, 2013.

[89] P. Serena, "Nonlinear Signal–Noise Interaction in Optical Links With Nonlinear Equalization," *J. Lightw. Technol.*, vol. 34, no. 6, pp. 1476–1483, Mar. 2016.

[90] I. S. Reed, "On a Moment Theorem for Complex Gaussian Processes," *IRE Trans. Inform. Theory*, vol. 8, no. 3, pp. 194–195, Apr. 1962.

[91] Z. Tao, W. Yan, L. Liu, L. Li, S. Oda, T. Hoshida, and J. C. Rasmussen, "Simple Fiber Model for Determination of XPM Effects," *J. Lightw. Technol.*, vol. 29, no. 7, pp. 974–986, 2011.

[92] Y. Fan, L. Dou, Z. Tao, L. Li, S. Oda, T. Hoshida, and J. C. Rasmussen, "Modulation Format Dependent Phase Noise Caused by Intra-Channel Nonlinearity," in *Proc. ECOC 2012*, 2012, vol. We.2.C.3, pp. 2–4.

[93] 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. 25685--25699, 2013.

[94] R. Dar, M. Feder, A. Mecozzi et al., "Nonlinear phase and polarization rotation noise in fully loaded WDM systems," in *Proc. ECOC 2015*, Valencia, Spain, 2015.

[95] N. Rossi, P. Ramantanis and J-C. Antona, "Nonlinear Interference Noise Statistics in Unmanaged Coherent Networks with Channels Propagating over Different Lightpaths," in *Proc. ECOC 14*, Paper Mo.4.3.4., Cannes, France, 2014.

[96] C. Schmidt-Langhorst, R. Elschner, F. Frey *et al.*, "Experimental Analysis of Nonlinear Interference Noise in Heterogeneous Flex- Grid WDM Transmission," in *Proc. ECOC 2015*, Valencia, Spain, 2015.

[97] T. Pfau, S. Hoffman and R. Noè, "Hardware efficient coherent digital receiver concept with feedforward carrier recovery for M-QAM constellations," *J. Ligthw. Technol.*, vol. 27, n. 8, pp. 989-999, 2009.